The Spectrum of Hepatic Involvement in Pediatric Inflammatory Bowel Disease: An Analysis of the Development of Abnormal Liver Enzymes and Associated Clinical Variables

by

Pamela L. Valentino MD,CM, FRCP(C)

A thesis submitted in conformity with the requirements for the degree of Master of Science in Clinical Epidemiology
Institute of Health Policy, Management and Evaluation
University of Toronto

© Copyright by Pamela L. Valentino 2014
The Spectrum of Hepatic Involvement in Pediatric Inflammatory Bowel Disease: An Analysis of the Development of Abnormal Liver Enzymes and Associated Clinical Variables

Pamela L. Valentino MD,CM, FRCP(C)

Master of Science in Clinical Epidemiology
Institute of Health Policy, Management and Evaluation
University of Toronto
2014

Abstract

Hepatic involvement in pediatric Inflammatory Bowel Diseases (IBD) includes presentation with abnormal liver enzymes (LEs). A proportion of these children do not develop serious hepatic consequences. We sought to describe the natural history of the development of abnormal LEs and associated clinical variables.

A carefully phenotyped cohort of 300 children with IBD (<18 years) was selected. Clinical data, (medications, LEs), were recorded. The Kaplan-Meier method and Cox Proportional Hazards modeling were employed to describe the first episode of abnormal LEs and identify associated variables.

The probability of developing abnormal LEs was 29.1% within 1 year, and 47.6% within 5 years, post IBD diagnosis. Children using corticosteroids or antibiotics were 70% (95% confidence interval [CI] 1.08, 2.66) and 362% (95%CI 2.96, 7.21) more likely to develop abnormal LEs, respectively. Similar associations were observed across various thresholds of LEs abnormalities.

Overall, abnormal LEs are common in IBD and most often associated with medication use.
Acknowledgments

I would like to take this opportunity to express my most sincere gratitude to all the people in my life who have supported me, my career development, and the evolution of this thesis over the past few years.

First and foremost, I would like to thank my thesis committee.

Binita Kamath, as my fellowship supervisor, you have been looking out for me, and guiding me through the jungle that is life in academia. You helped me develop my research ideas and carve a niche to call my own. You have shown me what a successful career in hepatology can look like, given me a path to start on, and a goal to strive for. My awards and success obtained thus far were attained due to your excellent mentorship.

Brian Feldman, as my thesis supervisor in the Clinical Epidemiology department, you have shown me how to get excited about the science behind clinical research. Your insightful suggestions throughout this process have helped me shape this project, and given me a different perspective on how to approach the analysis of pediatric clinical data.

Anne Griffiths, you have been an amazing role model to have in my life. As GI division chief, a research chair, and mother to many, you have been an inspiration, and yet you always found the time to provide counsel on issues related to this project and more.

Eleanor Pullenayegum, you have been extremely supportive in my endeavors to complete all the analysis using advanced software coding. I greatly appreciated your kindness and patience, as well as the time you took to guide me through understanding the statistical methods.

Second, I would like to acknowledge the members of the Gastroenterology, Hepatology, and Nutrition division at the Hospital for Sick Children. Thank you to Simon Ling for your counsel in the development of this project, and for your guidance as my GI fellowship program director. This project would not have been feasible without help from Thomas Walters and Karoline Fiedler; thank you for providing your expertise with the IBD patient database. Thank you to the GI fellows for providing your friendship and collaboration.
Third, I would like to thank the members of the Institute of Health Policy, Management and Evaluation as well as Sharon Dell, Jennifer James and many others in the Clinical Epidemiology Program who provided student support. Thank you to Peter Church for navigating the Clinical Epidemiology program with me and for always providing humor along the way.

Last, but not least, I would like to thank my family and the most important people in my life, my husband Mark and daughter Isabella, for all your endless love and unflinching support. Mark, we have grown up together, and our journey continues to take twists and turns, but I’m so glad to be on the path with you. Isabella, your delightful presence and beautiful smile has brought so much joy to my life; you inspire me to find the energy and courage to tackle any obstacle with grace. I love you both very much.
Table of Contents

Title...........................................................................................................................................i
Acknowledgments......................................................................................................................... iii
Table of Contents ........................................................................................................................... v
List of Tables .................................................................................................................................... viii
List of Figures ................................................................................................................................. x
List of Appendices ........................................................................................................................ xi

Chapter 1 Study Rationale.............................................................................................................. 1
  Rationale ...................................................................................................................................... 1
  Rationale for Analytic Approach ................................................................................................. 3
  Thesis Overview .......................................................................................................................... 3
  Study Questions, Hypotheses and Study Aims ......................................................................... 4
    Study Questions & Hypotheses ................................................................................................. 4
    Study Aims ............................................................................................................................... 5

Chapter 2 Liver disease in Pediatric Inflammatory Bowel Disease ............................................. 6
  Introduction ................................................................................................................................. 6
  Abnormal Liver Chemistry ........................................................................................................ 6
  Primary Sclerosing Cholangitis ................................................................................................ 7
    Diagnosis of Primary Sclerosing Cholangitis .......................................................................... 9
    Treatment of Primary Sclerosing Cholangitis ....................................................................... 12
    Outcomes in Primary Sclerosing Cholangitis ...................................................................... 13
  Other Autoimmune Liver Diseases ......................................................................................... 14
    Autoimmune Hepatitis ......................................................................................................... 14
    Autoimmune Sclerosing Cholangitis ..................................................................................... 17
  Drug Hepatotoxicity ................................................................................................................ 18
Limitations and Future Directions ................................................................. 53
Identification of Liver-Specific Outcomes ....................................................... 55
Summary of Future Directions ........................................................................ 56
References or Bibliography .............................................................................. 57
Appendices ......................................................................................................... 67

Appendix 1: Additional Methodology – Statistical Analysis ................................. 67
Dataset Manipulation – Liver enzymes data ...................................................... 67
Dataset Manipulation – Clinical IBD data ........................................................ 68
Dataset Manipulation – Medication data treated as Time-Varying Covariates .... 69
Dataset Manipulation – Liver enzymes data merged with clinical IBD data ...... 70
Statistical Analysis – Description of cohort ...................................................... 70
Statistical Analysis – Description of liver enzymes abnormalities .................. 71
Statistical Analysis – Time to Event analysis .................................................. 71
Statistical Analysis – Cox Proportional Hazards modeling ............................. 72

Appendix 2: Additional Methodology – SAS/STAT® Software Code .................... 73
Copyright Acknowledgements ........................................................................ 152
List of Tables

Table 2.1: The Majoie Classification of Cholangiographic Findings in Primary Sclerosing Cholangitis\textsuperscript{15} .................................................................................................................. 11

Table 2.2: Ludwig histopathological classification in PSC\textsuperscript{89} ................................................................. 11

Table 2.3: Drug hepatotoxicity reported in different IBD therapies ...................................................... 19

Table 3.1: Demographic data of the IBD cohort at the time of IBD diagnosis ............................................. 34

Table 3.2: Longitudinal clinical data of the IBD cohort .............................................................................. 36

Table 3.3: Univariate Cox Proportional Hazards models for the development of abnormal liver enzymes above the upper limit of normal in patients with IBD* .................................................. 43

Table 3.4: Multivariable Cox Proportional Hazards model for the development of abnormal liver enzymes above the upper limit of normal in patients with IBD .............................................................. 44

Table 3.5: Univariate Cox Proportional Hazards models for the development of abnormal liver enzymes in patients with IBD – 90 days post-IBD diagnosis* ................................................................. 44

Table 3.6: Univariate Cox Proportional Hazards models for the development of abnormal liver enzymes at >2x the upper limit of normal in patients with IBD* ................................................................. 46

Table 3.7: Multivariable Cox Proportional Hazards model for the development of abnormal liver enzymes at >2x the upper limit of normal in patients with IBD .............................................................. 47

Table 3.8: Univariate Cox Proportional Hazards models for the development of persistently, and mildly, abnormal liver enzymes (30-90days, 1-2x ULN) in patients with IBD* ......................... 47

Table 3.9: Multivariable Cox Proportional Hazards model for the development of persistently, and mildly, abnormal liver enzymes (30-90days, 1-2x ULN) in patients with IBD ......................... 48

Table 3.10: Univariate Cox Proportional Hazards models for the development of persistently, and severely, abnormal liver enzymes (30-90days, >2x ULN) in patients with IBD* ................. 48
Table 3.11: Multivariable Cox Proportional Hazards model for the development of persistently, and severely, abnormal liver enzymes (30-90 days, >2x ULN) in patients with IBD............. 49

Appendix Table 1: Dataset Manipulation – Liver enzymes data............................................. 73

Appendix Table 2: Dataset Manipulation – Clinical IBD data............................................. 80

Appendix Table 3: Dataset Manipulation – Medication data treated as Time-Varying Covariates .......................................................................................................................... 93

Appendix Table 4: Dataset Manipulation – Liver enzymes data merged with clinical IBD data .............................................................................................................................. 127

Appendix Table 5: Statistical Analysis – Description of cohort............................................ 133

Appendix Table 6: Statistical Analysis – Description of liver enzymes abnormalities......... 140

Appendix Table 7: Statistical Analysis – Time to Event analysis ........................................ 140

Appendix Table 8: Statistical Analysis – Cox Proportional Hazards modeling.................... 142
List of Figures

Figure 2.1: MRCP of a 17-year-old patient with PSC.................................................. 10

Figure 2.2: Liver histopathology from a 15 year old boy with PSC and UC. ...................... 12

Figure 2.3: Overlap between IBD and chronic liver diseases........................................ 16

Figure 2.4: Diagnostic algorithm for abnormal liver biochemistry in a child with IBD (without previous evidence of liver disease)............................................................ 26

Figure 3.1: Time to event analysis of abnormal liver enzymes (LEs) following IBD diagnosis. Development of the first abnormal LEs > upper limit of normal (ULN)......................... 38

Figure 3.2: Time to event analysis of abnormal liver enzymes (LEs) following IBD diagnosis. Development of the first abnormal LEs >ULN, stratified by phenotype of IBD.................. 39

Figure 3.3: Time to event analysis of abnormal liver enzymes (LEs) following IBD diagnosis. Development of first abnormal LEs >2x ULN ......................................................... 40

Figure 3.4: Time to event analysis of abnormal liver enzymes (LEs) following IBD diagnosis. Development of persistently abnormal LEs (between 30 to 90 days) at a low threshold (1-2x ULN)................................................................................................................................. 41

Figure 3.5: Time to event analysis of abnormal liver enzymes (LEs) following IBD diagnosis. Development of persistently abnormal LEs (between 30 to 90 days) at a high threshold (≥2x ULN)................................................................................................................................. 42
List of Appendices

Appendix 1: Additional Methodology – Statistical Analysis..............................................67

Appendix 2: Additional Methodology – SAS/STAT® Software Code.............................73
Chapter 1
Study Rationale

Rationale

The inflammatory bowel diseases (IBD) primarily affect the gastrointestinal tract; phenotypic classifications include: Crohn Disease (CD), Ulcerative Colitis (UC) and unspecified IBD (IBD-U), wherein a patient develops a colitis that has not yet been classified into one of these subtypes. Chronic liver diseases are the most important extraintestinal manifestations of IBD; as they can lead to end stage liver disease, as well as the need for liver transplantation at the most severe end of the spectrum.

These liver diseases include primary sclerosing cholangitis (PSC), autoimmune hepatitis (AIH), and autoimmune sclerosing cholangitis (ASC), an autoimmune disease with overlapping features of both PSC and AIH. Diagnostic investigations are comprised primarily of liver imaging, namely: ultrasonography, magnetic resonance cholangiography (MRCP) or endoscopic retrograde cholangiopancreatography (ERCP), as well as liver biopsy. The impetus to perform these investigations usually follows an elevation of liver enzymes (LEs). There is a higher index of suspicion for PSC or ASC in children with IBD, compared to the general population, due to the well known increased risk in this group (1.5% prevalence of PSC in one cohort of 1009 children with IBD). Thus, invasive hepatic investigations may be pursued in a more rapid progression in the setting of abnormal LEs in IBD.

Patients with IBD also commonly develop abnormal LEs for reasons other than chronic liver diseases; the majority of which are explored in chapter 2. This thesis investigates the timing and degree of LEs elevation in children with IBD, as well as variables associated with the development of abnormal LEs, with hopes to lay the groundwork for a data-driven algorithm for diagnostic investigation.

The medical literature is limited to few studies of abnormal LEs in pediatric IBD, identifying a clear need for further investigation. Nemeth et. al. reported on 46 children with IBD and found a 52% period prevalence for the development of abnormal LEs over 12 years of follow-up. Due to the small number of patients in this case series, there was insufficient power to perform further tests of association. Hyams et. al. explored the case records of 555 patients and described a 14%
period prevalence of alanine aminotransferase (ALT) abnormality ≥ 80U/L (~2 times the upper limit of normal [ULN] for age) over a mean follow-up of 5.4 years in patients with CD and 6.5 years in patients with UC. Aspartate aminotransferase (AST) and gamma-glutamyl transpeptidase (GGT) were not considered in these analyses. Furthermore, clinical information, such as medication history, was only obtained in patients found to have abnormal LEs. Thus, this study was flawed by its lack of consideration of other LEs testing, its omission to account for patient censoring, and the inability to compare clinical variables between patients with and without liver abnormalities. A significant gap still exists in the literature regarding the natural history of abnormal LEs in pediatric IBD.

The literature in adults with IBD is slightly more extensive; however, less generalizable to the pediatric population. A small selection of older studies in adult patients with IBD, published between 1970 to 2000, examined liver enzyme elevations and reported varying percentages of patients with liver disease. However, the cases in these reports included those with possible viral hepatitis, potentially from transfusion-related infections that were more prevalent at the time, and are thus not generalizable to the present pediatric population. Furthermore, with our current knowledge of liver disease, and with easier access to investigations such as MRCP, it is likely that the work-up for patients with elevated liver enzymes was less complete at that time. More recent studies of adults with IBD report a relatively high prevalence of abnormal LEs and are further described in chapter 2.

At the Hospital for Sick Children (SickKids), we are uniquely poised to retrospectively describe the spectrum of liver involvement in pediatric IBD. A robust prospectively maintained patient database exists that contains pertinent clinical information about the children diagnosed with IBD at our centre. This database stores information about IBD disease phenotype, demographic information, as well as medication data with start and stop dates. LE data are stored in the SickKids data warehouse and are available for download electronically. Thus, a large portion of reliable longitudinal data were feasibly available for study.
Rationale for Analytic Approach

Longitudinal LE data were available for our cohort of children at The Hospital for Sick Children, following their diagnosis with IBD and until their transfer to adult healthcare, via a data warehouse. In order to capture the probability of developing abnormal LEs at different time points following a diagnosis with IBD, a Kaplan-Meier (KM) analysis was used. This method was chosen due to the inevitable presence of right-censoring of data in the pediatric cohort: patients were diagnosed with IBD at different ages, and were followed for variable durations of time prior to transfer to adult healthcare (at approximately 18 years of age). Furthermore, the KM method was able to detect the probability of developing abnormal LEs at different time points following an IBD diagnosis.

The assumptions for KM analysis were met for the computation of the probability of developing abnormal LEs in our cohort. The start time and end points were well defined for both the group who developed abnormal LEs as well as the censored group. Furthermore, the censoring of patients was in no way related to the presence or absence of abnormal LEs. We also made the assumption that the probability of detecting abnormal LEs was similar throughout the course of the time period studied, as there were no new guidelines for routine liver monitoring. Finally, the detection of the abnormal LEs was assumed to be accurate throughout the study, although the frequency at which patients had bloodwork investigations was variable.

IBD phenotypic variables associated with the development of abnormal LEs were identified using Cox proportional hazards (CPH) semi-parametric models. These models are a regression analysis of time to event data and allow for both time dependent, as well as time independent, variables. While the majority of the IBD variables analyzed did not vary over time (e.g. IBD phenotype, disease location at diagnosis, etc.), the medications used in the treatment of IBD fluctuated considerably per patient. As such, the database was constructed using the counting process style of input to account for the start and stop dates of the medications.

Thesis Overview

The spectrum of hepatic involvement in IBD, other than PSC, is relatively under-studied. There was clear justification to systematically study a well-characterized cohort of pediatric IBD
patients to determine the incidence and natural history of LEs abnormalities. Furthermore, we aimed to identify clinical variables associated with the development of abnormal LEs. In chapter 2, a background on the subject of hepatic involvement in pediatric IBD is presented. Following this, chapter 3 describes a study of LEs in a cohort of children with IBD followed at SickKids; greater details of the methodology are outlined in the appendix. Finally, chapter 4 is a discussion of the limitations of this study, as well as a proposal for future investigations. The eventual goal is the development of an evidence-based diagnostic algorithm for the approach to elevated LEs in children with IBD, with the objective of preventing unnecessary, invasive, or costly investigations in children with IBD whilst also not overlooking the first manifestations of serious disease.

Study Questions, Hypotheses and Study Aims

Study Questions & Hypotheses

**Question 1:** What is the duration of time between the diagnosis of IBD in children and the onset of abnormal LEs?

- We hypothesized that patients will present with abnormal LEs early, within the first 3 months, following a diagnosis of IBD.

**Question 2:** How many children develop abnormal LEs at different time points (at 1, 2 and 5 years) following their IBD diagnosis, but prior to transition to adult healthcare?

- When accounting for censoring, and based on adult literature, we hypothesized that overall 50% of children with IBD will develop abnormal LEs prior to transition to adult healthcare.

**Question 3:** How many children develop significantly (≥2xULN), or persistently abnormal LEs following their IBD diagnosis?

- We hypothesized that a considerably smaller proportion of patients with IBD will develop higher thresholds of abnormal LEs at any point in time prior to transition to adult healthcare.
**Question 4:** Which clinical variables are associated with the development of abnormal LEs in children with IBD?

- We hypothesized that known hepatotoxic medications, such as methotrexate, will demonstrate an association with the development of abnormal LEs.

**Study Aims**

**AIM 1:** To estimate the duration of time between the diagnosis of IBD and onset of abnormal LEs, among an inception cohort of children, <18 years of age, with newly diagnosed IBD, between January 2000 – October 2011.

**AIM 2:** To report the prevalence of abnormal LEs at the time of IBD diagnosis as well as the proportion of children who develop abnormal LEs by 1, 2 and 5 years post the diagnosis of IBD.

**AIM 3:** To report the proportion of children who develop higher thresholds, or persistently abnormal LEs by 1, 2 and 5 years post the diagnosis of IBD.

**AIM 4:** To identify specific clinical variables associated with an increased risk for the development of abnormal LEs.
Chapter 2
Liver disease in Pediatric Inflammatory Bowel Disease

Introduction

Hepatic involvement in inflammatory bowel disease (IBD) is a spectrum that ranges from transient abnormalities in laboratory values, to liver diseases that can require transplantation. There is no standard of care in the diagnostic approach to elevated liver biochemistry in IBD. While it is appropriate to minimize invasive investigations, this should not be at the expense of overlooking a diagnosis of chronic liver disease. The chronic hepatic diseases associated with IBD include primary sclerosing cholangitis (PSC), autoimmune hepatitis (AIH) and autoimmune sclerosing cholangitis (ASC, an overlap syndrome in which features of both PSC and AIH are detected on laboratory, histological and cholangiographic investigations\(^9\)). Medications used in IBD treatment can also cause significant hepatotoxicity, and thus require close monitoring during their administration. Other hepatobiliary manifestations reported in the IBD literature include cholangiocarcinoma, choledolithiasis, hepatic abscess, hepatic or portal vein thrombosis, fatty liver, and amyloidosis (all of these have been described in children except the last). In this chapter, we will explore hepatic involvement in pediatric IBD in further detail.

Abnormal Liver Chemistry

Abnormal liver biochemistry, in patients with IBD, is a common finding. In adults with ulcerative colitis (UC), 40% of patients can present with abnormal liver enzymes\(^10\). In one study of 200 adult UC patients, where the outcomes of elevated liver biochemistry were assessed, the
investigators identified non-specific transient increases in 63% of patients, fatty liver disease in 11.2%, PSC in 6.3%, drug toxicity in 6%, and other diagnoses in 13.5%, such as: autoimmune hepatitis, chronic hepatitis C infection and total parenteral nutrition-associated cholestasis. The patients with transient abnormalities in hepatic enzymes all had mild elevations at less than 3 times the upper limit of normal; however, time to resolution was not reported \(^\text{10}\). In another cohort of 786 adult patients with IBD, 15% developed elevated liver enzymes, of which 2.5% developed PSC, 42.3% had hepatotoxicity from azathioprine (AZA) or 6-mercaptopurine (6MP) use, 40.8% had fatty liver disease, 28% had no known associated liver disease \(^\text{11}\).

There has only been one study to date investigating the prevalence of abnormal liver biochemistry in pediatric IBD patients. Nemeth et al retrospectively assessed 46 children with IBD, and identified 24 patients (52%) who developed elevated liver enzymes at some point over a mean follow up of 5.2 years \(^\text{2}\). Of the 9 patients with “severe liver involvement,” defined as liver enzymes greater than three times the upper limit of normal, eight had biopsies with histologic findings that included bile duct proliferation, portal fibrosis, and/or portal inflammation. Four of the eight patients had changes consistent with a diagnosis of small duct PSC. Of those with “severe liver involvement”, more patients were observed to have pancolonic UC as compared to those with distal UC or Crohn disease (CD). Unfortunately, the small number of patients, lack of radiological reporting, and the inability to show statistical differences among the IBD groups, were major limitations in this study.

In general there is a paucity of data regarding the prevalence and natural history of abnormal liver biochemistry in pediatric IBD. The data that exists suggests that up to half of patients with IBD will have abnormal liver enzymes during their course, and in the majority, these are transient changes. The limited literature also seems to indicate that if transaminases are less that 2 times the ULN, for a short period of time (such as less than 4 weeks), then observation alone may be appropriate without resorting to invasive investigations.

**Primary Sclerosing Cholangitis**

PSC is an autoimmune cholestatic liver disease characterized by intrahepatic or extrahepatic pericholangitis with progression to multifocal biliary strictures, liver fibrosis, cirrhosis, and
complications of portal hypertension. A Canadian study of 49 cases of PSC described an incidence rate of 0.23 cases per 100,000 person-years in children and 1.11 cases per 100,000 person-years in adults. The first reported association between PSC and IBD was by Warren et al in 1966; they described 42 cases of PSC, 12 of which had concurrent UC. The prevalence of PSC in pediatric IBD has most recently been reported as 1.5% in a study of 1009 children with IBD across 19 North American centers, and as 2.6% in a study of 786 adult patients with IBD in Spain. Pediatric PSC has an equal sex distribution, with 50% of cases occurring in females, instead of the male overrepresentation seen in adults.

The presentation of PSC can be quite variable: in 47 children with PSC at Mount Sinai Medical Centre in New York, 40% presented with hepatomegaly, 36% had abdominal pain, 23% had splenomegaly, 19% had pruritus, 17% had jaundice, and 19% were asymptomatic. The previously described Canadian incidence study of patients with PSC identified 73.4% of patients to also have IBD (CD: 19/49, UC: 17/49). This is consistent with a subsequent study, where the prevalence of IBD in PSC was 68.9% (20/29). Despite the high prevalence of IBD in PSC patients, the overall frequency of PSC in patients with IBD is quite low. An American multicentre pediatric cohort study of IBD identified PSC in 1.5% of 1009 IBD patients (CD: 7/728, UC: 8/281). Typically, the IBD phenotype associated with PSC is primarily pancolonic UC with low-grade inflammation, rectal sparing, and backwash ileitis. Cases of Crohn Disease (CD) occur as well, and the Canadian PSC incidence study described a variable phenotype: 57.9% (11/19) had enterocolitis, 36.8% (7/19) had isolated colitis, and 5.3% (1/19) had an ileitis without evidence of colonic inflammation. In PSC/IBD patients undergoing proctocolectomy, there is an increased risk of peristomal varices at the ileostomy site, as well as an increased risk of pouchitis after ileal pouch anal anastomosis. Meanwhile, patients with PSC who have underlying mild IBD may be asymptomatic from a GI standpoint. Thus, as stated in the 2010 American Association for the Study of Liver Diseases (AASLD) PSC practice guidelines, a complete screening endoscopy is warranted in all patients with apparently isolated PSC at diagnosis, due to implications for long-term management.
Diagnosis of Primary Sclerosing Cholangitis

There are no widely accepted and validated current diagnostic criteria for pediatric PSC; however, there are typical findings on laboratory investigations, cholangiography, and liver biopsy. This disease can be classified as “large duct” PSC, “small duct” PSC, and ASC. The first indication of PSC is often abnormal liver biochemistry. At the time of PSC diagnosis, in the Mount Sinai Medical Centre study, mean serum biochemical values were: alanine-aminotransferase (ALT) 233 (±327), aspartate-aminotransferase (AST) 236 (±248), gamma-glutamyltranspeptidase (GGT) 553 (±676), alkaline phosphatase (ALP) 610 (±340), total bilirubin 1.3 mg/dL (±1.9), direct bilirubin level 0.7 mg/dL (±1), and albumin level 3.9 g/dL (±0.5). In children, serum GGT levels have been regarded as a better screening tool for biliary disease than ALP due to the elevated ALP levels seen with bone growth. Serology can also be helpful in the diagnosis of PSC as these patients have an increased presence of perinuclear anti-neutrophil cytoplasmic antibody (p-ANCA) positivity with 67-87% prevalence. In a study of 73 adult patients with PSC at the Mayo clinic, there was no significant difference in ANCA positivity between patients with IBD (81% positive) or without IBD (93% positive). When reviewing serology results, it appears that the atypical pattern of p-ANCA fluorescence is frequently found in PSC/IBD, with visualization of fluorescent staining at the periphery of the nucleus and multiple foci within the nucleus as well. In the setting of a child with IBD and raised liver enzymes, ANCA testing may be positive; however, its absence is insufficiently robust for the exclusion of a diagnosis of PSC in patients with IBD.

Imaging investigations are crucial to the diagnosis of PSC. An abdominal ultrasound is usually the first-line investigation to assess the biliary tree, look for biliary dilatation or wall thickening, as well as peri-hilar lymph nodes, and screen for other conditions that could be associated with a raised GGT, such as cholelithiasis. The definitive radiological tests are cholangiographic studies, either magnetic resonance cholangiography (MRCP) or endoscopic retrograde cholangiopancreatography (ERCP) (both of which have superseded the more invasive percutaneous cholangiography). Cholangiography can assess for both intrahepatic and extrahepatic “large duct” involvement, as well as determine if the disease is localized or diffuse (Figure 2.1). The Majoie classification for cholangiography has been used in previous reports to categorize the extent and localization of disease (Table 2.1). Typical findings include biliary duct wall irregularities with stenosis and dilatation, as well as reduction of intraparenchymal
arborisation. In a prospective pediatric study by Ferrara et al in 2002, 21 patients aged 7-14 years, with suspicion of PSC, were all evaluated by ERCP, MRCP, and liver biopsy. Three patients were non-cooperative and could not complete the MRCP protocol. MRCP showed abnormalities consistent with PSC in 13/21 (62%) patients, while ERCP diagnosed PSC in 16/21 (76%) patients. In the 5 patients with normal MRCP imaging, no abnormalities were identified on ERCP. In this study, the specificity of MRCP for PSC was 100% with a sensitivity of 81%. However, the sensitivity was reduced by the inability for the children to complete the exam and not the intrinsic ability for the modality to identify abnormalities. While both MRCP and ERCP are of high quality, ERCP carries the risks of pancreatitis, cholangitis, hemorrhage, duodenal perforation, and death. Thus, MRCP, with sedation when appropriate, can be recommended as a reliable primary diagnostic imaging modality in PSC.

**Figure 2.1: MRCP of a 17-year-old patient with PSC.**

Note the intrahepatic and extrahepatic beading of the biliary tree with sequential strictures and saccular dilatations. The common bile duct wall is also thickened.

A liver biopsy is indicated in a child with IBD and persistently raised liver enzymes, particularly if the GGT is also elevated. The liver biopsy should be performed in this setting regardless of the
radiological findings as PSC may involve the smallest biliary ducts and only be detected microscopically. A biopsy is also needed to identify any evidence of ASC, which requires a different therapeutic approach (see ASC). The histopathology in small duct PSC can reveal acute and/or chronic cholangitis, fibrosis, or cirrhosis (Table 2.2)\textsuperscript{25}. A finding known as “onion skinning” is the most characteristic histopathologic lesion of PSC, which describes concentric periductal fibrosis (Figure 2.2).

Table 2.1: The Majoie Classification of Cholangiographic Findings in Primary Sclerosing Cholangitis\textsuperscript{15}

<table>
<thead>
<tr>
<th>Intrahepatic</th>
<th>Extrahepatic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimal strictures with normal diameter of biliary ducts or minimal dilatation</td>
<td>Duct contours have slight irregularities without stenosis</td>
</tr>
<tr>
<td>Multiple strictures with saccular dilatations and reduction of intraparenchymal arborisation</td>
<td>Segmental stenosis of the biliary duct</td>
</tr>
<tr>
<td>Lack of visualization of one of the main hepatic ducts</td>
<td>Almost the entire biliary duct is stenotic</td>
</tr>
<tr>
<td></td>
<td>Pseudo-diverticular out-pouching of the biliary ducts with irregular duct margins or diameter</td>
</tr>
</tbody>
</table>

Table 2.2: Ludwig histopathological classification in PSC\textsuperscript{89}

<p>| |</p>
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Portal tract fibrosis with cholangitis and portal hepatitis</td>
</tr>
<tr>
<td>Periportal fibrosis or hepatitis</td>
</tr>
<tr>
<td>Septal fibrosis and/or bridging fibrosis</td>
</tr>
<tr>
<td>Cirrhosis</td>
</tr>
</tbody>
</table>
Figure 2.2: Liver histopathology from a 15 year old boy with PSC and UC.

On H&E staining (A) and trichrome staining (B) there is evidence of periductal fibrosis (“onion skinning”, arrow) with degeneration and atrophy of the ductal epithelium of some cholangioles. Also note the severe bridging fibrosis with fibrous septae linking portal tracts.

(Images courtesy of Dr. Ernest Cutz, MD, FRCPC, Hospital for Sick Children, University of Toronto)

Treatment of Primary Sclerosing Cholangitis

Unfortunately, pharmacological treatments for PSC have been disappointing, as they have not shown a decrease in the rate of disease progression or a change in the rate of survival. Treatment with corticosteroids, penicillamine, and azathioprine have not shown long-term success in improving outcomes in PSC. Tacrolimus has been shown to cause a reduction in ALT and ALP but was poorly tolerated in adults. Mycophenolate mofetil, methotrexate, budesonide, etanercept, pentoxifylline (an anti-TNF agent), and pirfenidone (an anti-fibrotic agent) have all not shown any short or long-term benefit. Ursodeoxycholic acid (UCDA), a dihydroxy bile acid, has been the mainstay of medical therapy for PSC, having been shown to decrease the hepatotoxicity of other bile salts, provide a hepato-protective effect, and improve the serum ALT. However, a recent trial of high dose ursodeoxycholic acid (28–30 mg/kg/day) versus placebo in adults revealed an association with higher rates of liver transplantation and death, suggesting toxicity at these doses. The AASLD guidelines recommend against the use of
UCDA in adults with PSC, however there is insufficient evidence to establish firm guidelines for or against the use of this therapy in children\(^9\). Many pediatric hepatologists continue to utilize ursodeoxycholic acid at lower doses (up to 20 mg/kg/day divided twice daily), as there may be potential benefit for improving outcomes. Oral vancomycin, at 50 mg/kg/day, has also been reported to improve liver biochemistry and reduce inflammation on liver biopsy in children with both UC and PSC\(^{30}\). Of note, those with liver cirrhosis on biopsy did not show a statistically significant degree of improvement. Furthermore, survival outcomes after long-term vancomycin use have not been reported. Minocycline and metronidazole have also been shown to reduce abnormal liver biochemistry in PSC\(^{27}\). These data suggest that antibiotics can decrease inflammation in the PSC liver, implicating bacteria in the pathogenesis of PSC. Overall, the management of PSC involves close co-operation with a hepatologist, limitation of exposure to hepatotoxic medications as well as regular monitoring of liver biochemistry.

While MRCP is a non-invasive diagnostic tool, which is desirable in pediatrics, an ERCP has the advantage of providing therapeutic potential. In the setting of a localized large duct stricture, a biliary intervention via ERCP can ameliorate the course of disease. A bile duct stricture in PSC can lead to worsening liver function with progression to cirrhosis as well as increased risk of bacterial cholangitis\(^ {31}\). With ERCP-mediated balloon dilatation ± stenting of a stenotic duct, there can be improvement of liver biochemistry, including bilirubin, decreased jaundice and pruritus (if present), and improved survival\(^ {32}\).

**Outcomes in Primary Sclerosing Cholangitis**

PSC causes liver fibrosis and can progress to liver cirrhosis, as well as hepatic failure. Angulo et. al. in 2002 compared patients with small-duct PSC \((N=18)\) versus classic PSC \((N=36)\), and found that survival with native liver was significantly higher in the group with small duct PSC \((p=0.04)\), suggesting a better prognosis with this pattern of disease\(^ {33}\). A small group of patients \((3/18)\) with small-duct disease were found to develop classic PSC during follow-up. It is unclear if the two patterns of disease are on a continuum or are distinct entities. In a single-center study from Mount Sinai, 19% of children required liver transplantation\(^ {15}\). A recent study from the SPLIT (Studies of Pediatric Liver Transplantation) database assessed the outcomes of children with PSC following liver transplantation\(^ {34}\). The patient survival was 98.7% at 1 year and 86.6%
at 5 years, compared to 94.3% and 88.2% respectively in the non-PSC group. Meanwhile, the graft survival was 93% at 1 year and 76.1% at 5 years, compared to 90% and 79.5% respectively in the non-PSC group. There were no significant differences in patient and graft survival between the two groups. However, with a univariate analysis, mortality was significantly higher in PSC patients with a diagnosis of IBD pre-transplant. Of potential interest, a recurrence of PSC was diagnosed in 6 of 61 patients (9.8%) over a median follow-up time of 36.6 ± 32.7 months, and all who recurred had been diagnosed with concurrent IBD pre-transplant. In a prospective comparison of adult PSC patients with a dominant stricture of the biliary duct, the transplant-free survival after 18 years of follow-up was poorer in the group of patients with concomitant UC (23% versus 77.8% respectively, \( p=0.045 \))\(^{35} \). Proctocolectomy has not shown any difference in survival or decrease in the complications of portal hypertension, in adult patients with concomitant PSC/UC\(^{36} \). There is insufficient evidence to recommend colectomy in children with PSC/UC to prevent progression of PSC disease or liver transplantation.

Another important complication of PSC is an increased risk of cholangiocarcinoma. In a study of 1274 adult patients with UC, the overall prevalence of this malignancy was 0.3%, however, in patients with concurrent PSC, the prevalence was 13%\(^{4} \). Furthermore, patients with PSC and a dominant bile duct stricture have a statistically significant increased risk of cholangiocarcinoma, as well as colorectal carcinoma, if they have concurrent IBD\(^{35} \). These cancers are rare in children, and the AASLD guidelines do not recommend increased surveillance in the pediatric population with PSC/IBD, while adult screening with colonoscopy is recommended every 1-2 years after diagnosis\(^{9} \).

**Other Autoimmune Liver Diseases**

**Autoimmune Hepatitis**

AIH is an autoimmune disorder characterized by hepatic inflammation, and it can be categorized according to the type of autoantibodies present. In AIH type 1, antinuclear antibodies and/or smooth muscle antibodies (ANA/SMA) are frequently positive, while in AIH type 2 antibodies to liver/kidney microsomal type 1 (LKM1) can be positive. Children can be seronegative, or have antibodies titers that are lower than in adults such that 1:20 for ANA or SMA, and 1:10 for
anti-LKM1 are significant levels in children, while in adults 1:40 is clinically relevant, as per AASLD guidelines. Both types of AIH favor a female predominance (75%), have elevated immunoglobulin G (IgG) titers, and in children, progress with a similar disease course. However, patients with type 2 disease frequently present at a younger age (median age: 7.4 versus 10.5 years). The clinical presentation can include non-specific symptoms such as: fatigue, nausea, abdominal pain, and arthralgia, while others may present with jaundice, acute hepatitis, or even liver failure. A liver biopsy is necessary to confirm the diagnosis as well as to rule out other causes of liver biochemistry derangement. Findings include: portal tract inflammation, lobular hepatitis, interface hepatitis, and cirrhosis. Plasma cell infiltration of the portal tracts are commonly associated with the periportal hepatitis, however, in some patients with AIH this may be absent. Autoantibodies are integral to making a diagnosis of AIH as described above. However, ANCA can also be frequently positive in AIH type 1. In one study within adults, 65% of 46 patients with AIH type 1 were ANCA-positive, however none of the 19 patients with AIH type 2 were positive. The International Autoimmune Hepatitis Group (IAIHG) developed and revised a scoring system for the diagnosis of AIH in 1999 that incorporates many of these parameters. Some indices included in the score are: liver biochemistry, autoantibodies, viral serology status, drug or alcohol use, and findings on liver histology. With a total AIH score above 15 points, the diagnosis of AIH is considered definite, while with a score between 10-15 points, AIH is probable. Ebbeson and Schreiber retrospectively evaluated this tool in 2004, in children with diagnoses of AIH (N=21) and sclerosing cholangitis (N=7). They found 85.7% (18/21) of patients clinically diagnosed with AIH to have “definite AIH” according to the IAIHG tool, while 14.3% (3/21) had “probable AIH”. Of the patients with sclerosing cholangitis, all 4 with isolated cholangitis were found to have a score of <10 points, while the three patients with overlap disease (ASC) had scores consistent with definite AIH. This study evaluated a small group of patients, however, it suggests that patients with hepatic inflammation consistent with AIH may be reliably differentiated from isolated PSC with this tool. However, this tool is not specific enough to rule out ASC in patients who appear to have AIH.

Immunosuppressive therapy is the standard of care for AIH in children due to the increased disease severity at onset as well as the aggressive course with poorer long-term outcomes observed if treatment is delayed. Prednisone (1–2 mg/kg/day) is used to induce remission and can be continued at a lower dose (2.5–5 mg/day) for long-term maintenance. Azathioprine is
usually added for maintenance treatment, as recommended in AASLD guidelines. In children who fail to respond to these medications, additional options include MMF, cyclosporine, and tacrolimus, although the evidence for use of these agents is based on case series and not randomized trials comparing to standard treatment regimens. After successful therapy for 2–3 years, with persistent normalization of liver biochemistry, immunosuppression discontinuation may be attempted after a liver biopsy has confirmed complete histologic remission of disease.

Other autoimmune diseases, such as hypothyroidism, occur in 20–22% of patients with AIH. Of particular relevance here is the overlap with IBD. Within the IBD population, AIH is an uncommon finding: 0.9% of 1009 children in a multicenter study with IBD, and 1.4% with UC had a “chronic active hepatitis”1. However, within the pediatric AIH population, IBD is more frequently diagnosed with a prevalence of 12–20% of children with AIH in one study, advocating a low index of suspicion for IBD in this population (Figure 2.3).

**Figure 2.3: Overlap between IBD and chronic liver diseases**

This Venn diagram represents the relationship between IBD and chronic liver diseases, as well as the perceived prevalence (not drawn to scale).
Autoimmune Sclerosing Cholangitis

Autoimmune sclerosing cholangitis occurs almost exclusively in children and young adults\(^{25,42,43}\). It is diagnosed when features of both PSC and AIH are present on liver biopsy and imaging, however, as in PSC, there are no specific diagnostic criteria. These patients have an increased frequency of concomitant IBD (44%) compared with both PSC and AIH\(^{38}\). In one study at King’s College in London, children referred with biochemical evidence of liver disease as well as positive autoantibodies were prospectively assessed\(^{25}\). Of the 76 patients referred between 1984 and 1997, 55 patients, who were investigated with both a liver biopsy as well as cholangiography, were included in the study. A diagnosis of ASC was attributed to patients who had histological evidence of AIH and cholangiographic evidence of intrahepatic or extrahepatic sclerosing cholangitis. Of note, this definition of ASC may have precluded a diagnosis of small duct sclerosing cholangitis, overlapping with AIH, which would only be diagnosed histologically. Nonetheless, 27/55 (49%) patients were diagnosed with ASC while 28/55 (51%) of patients were diagnosed with AIH. There were no significant differences in the median age at diagnosis (10.5 years ASC, 11.8 years AIH), or female preponderance (55% ASC, 79% AIH).

Signs and symptoms were also similar at diagnosis with the majority presenting with jaundice, hepatomegaly, and splenomegaly, except for pruritus, which was significantly increased in AIH (25% vs. 7% in ASC). Laboratory investigations at baseline were mostly similar between ASC and AIH except for median AST (102 IU/L in ASC, 333 IU/L in AIH, \(p=0.002\)), median total bilirubin (1.2 mg/dL or 20 \(\mu\)mol/L in ASC, 2.0 mg/dL or 35 \(\mu\)mol/L in AIH, \(p=0.04\)), and ANCA positivity (74% in ASC, 36% in AIH, \(p=0.009\)). In children with ASC, ANA and SMA titers were positive in 20/27 (74%) of patients and LKM1 antibodies were present in 1/27 (4%) of patients. Autoimmune disorders in first-degree relatives were also significantly increased in AIH (71% vs. 37% in ASC). On histology, there was increased acute and/or chronic cholangitis in patients with ASC compared to AIH (35% vs. 12% respectively, \(p=0.049\)) but decreased portal tract inflammation (58% vs. 92% respectively, \(p=0.004\)).

Routine sigmoidoscopy with rectal biopsies was also performed in all patients within this cohort except for one who refused. IBD was diagnosed in 44% of patients with ASC. This was significantly different from the 18% of patients with AIH who were also diagnosed with IBD. There was potential for the underestimation of the prevalence of IBD in this population as the
colitis associated with PSC frequently has rectal sparing and only rectal biopsies were taken in this cohort.

On treatment with immunosuppression, 100% of patients with AIH experienced normalization of liver biochemistry, while in ASC, 83% had normal AST levels, and 73% had normal bilirubin. Meanwhile, there was no significant difference in rate of relapse with 36% of AIH patients versus 33% of ASC patients experiencing at least one episode of relapse. Of those patients with ASC who had follow-up cholangiography, 8/17 had progression of intrahepatic and extrahepatic biliary disease. Upon follow-up, all patients were alive and no patients with AIH required liver transplantation in this cohort. Meanwhile, 4/27 (15%) of children with ASC required liver transplantation. This study suggests that children with AIH/IBD should be screened with cholangiographic imaging to assess for PSC, as this could have implications for a poorer prognosis. This is clearly stated in the AASLD practice guideline in 2010. Conversely, children with cholangiographic evidence of PSC should undergo liver biopsy and autoantibody detection to screen for evidence of overlap, as ASC requires immunosuppressive treatment.

According to AASLD guidelines, though based on limited evidence, treatment with corticosteroids (±azathioprine) may be attempted in PSC patients with elevated IgG, positive autoimmune antibodies, and interface hepatitis on liver biopsy.

**Drug Hepatotoxicity**

**Methotrexate**

Methotrexate, an anti-inflammatory that inhibits synthesis of purines and pyrimidines via inhibition of the folate metabolism, has been found to maintain remission in IBD. Previous use in rheumatological diseases has identified a hepatotoxic effect of methotrexate, with reports of elevation of liver biochemistry and hepatic fibrosis (15% prevalence). The use of methotrexate in psoriasis came with the recommendation in 1996 to perform protocol liver biopsies after administration of a cumulative dose of 1,500 mg. With further investigation, this practice was changed. Updated publications for psoriasis and juvenile idiopathic arthritis have included recommendations for biopsies in cases of refractory abnormal liver biochemistry despite decreased or held doses of methotrexate. A biopsy can also be considered with higher cumulative
doses of methotrexate (3.5–4.0 g)\textsuperscript{47,48}. Subsequent studies in adults and children with IBD have also not supported protocol liver biopsies although the numbers of patients in each study were limited. The 2 adult retrospective studies that included the most number of liver biopsies after methotrexate administration (Fournier 2010: $N=17/87$ and Te 2000: $N=20/32$) demonstrated that an increased cumulative dose of methotrexate did not correlate with worse liver histopathology\textsuperscript{49,50}. In the pediatric IBD literature, retrospective studies of methotrexate use have described abnormal liver biochemistry requiring dose reductions or drug discontinuation (Turner 2007: $N=7/60$ and Uhlen 2006: $N=2/61$)\textsuperscript{51,52}. However, liver biopsies were not performed in these cases making it difficult to interpret the exact effect of methotrexate in children with IBD. Previous studies of methotrexate hepatotoxicity have also described hepatic steatosis on liver biopsy. These findings were significantly worse in psoriatic adult patients with obesity ± diabetes mellitus compared to psoriatic patients without these risk factors (and matched for cumulative methotrexate doses)\textsuperscript{45,53}. It is rare for patients to receive baseline liver biopsies prior to methotrexate use. Therefore, the contribution of methotrexate to the degree of steatosis is unclear. Nonetheless, patients with obesity or diabetes mellitus require close monitoring as methotrexate could further exacerbate underlying non-alcoholic fatty liver disease (NAFLD).

From the available limited evidence, when starting methotrexate, liver biochemistry should be obtained at baseline and in follow-up: weekly after the initial dose or after changes in doses for the first month, and every 2–3 months thereafter. With development of persistently abnormal liver biochemistry, in consultation with a hepatologist, the dose of methotrexate can be adjusted, or temporarily held with mild to moderate abnormalities (e.g., up to 2–3x ULN, the upper limit of normal), or completely discontinued if highly abnormal (e.g., >5x ULN). Liver biopsies can be reserved for cases of persistently abnormal liver biochemistry or if discontinuation of the methotrexate would be deleterious for the IBD management. Caution should be exerted in prescribing methotrexate to patients with underlying liver disease and avoided in children with PSC unless under exceptional circumstances (Table 2.3).

Table 2.3: Drug hepatotoxicity reported in different IBD therapies

<table>
<thead>
<tr>
<th>Drug</th>
<th>Liver Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methotrexate</td>
<td>Hepatic Fibrosis</td>
</tr>
<tr>
<td></td>
<td>Hepatic Steatosis</td>
</tr>
<tr>
<td>Azathioprine /</td>
<td>Hepatitis</td>
</tr>
<tr>
<td>Drug</td>
<td>Side Effects</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>6-Mercaptopurine</td>
<td>Peliosis hepatis, Veno-occlusive disease, Severe cholestasis, Nodular regenerative hyperplasia, Portal hypertension</td>
</tr>
<tr>
<td>5-aminosalicylic acid</td>
<td>Fever, rash, jaundice, hepatomegaly, Acute liver failure</td>
</tr>
<tr>
<td>Anti-TNF, biologic agents</td>
<td>Hepatitis, Hepatitis B reactivation, Hepatosplenic T cell Lymphoma</td>
</tr>
<tr>
<td>Cyclosporine</td>
<td>Mild elevated liver biochemistry, Biliary sludge</td>
</tr>
</tbody>
</table>

**Azathioprine/6-Mercaptopurine**

The thiopurine immunomodulators provide an anti-inflammatory effect, and maintenance of remission in IBD, by conversion to 6-thioguanine nucleotides, which insert into the DNA of rapidly dividing cells and suppress replication\(^5^4\). While, adverse effects can occur in 15-30% of patients with IBD on thiopurines, hepatotoxicity is rare\(^5^5\). Adult retrospective studies have shown frequencies of Azathioprine (AZA)/6-Mercaptopurine (6MP) hepatotoxicity ranging from 0 to 5.2%\(^5^6\)\(^-\)\(^6^1\). The majority of patients improved with dose reductions or drug discontinuation. One patient developed peliosis hepatis, blood-filled cavities on histopathology, which resolved with drug withdrawal. Other reports have identified isolated cases of veno-occlusive disease and severe cholestasis, of which both resolved upon drug discontinuation\(^6^2\)\(^,\)\(^6^3\). A case-control study of adults with IBD treated with azathioprine identified 37 cases of nodular regenerative hyperplasia, and found a positive association with male gender and stricturing IBD disease. These cases were identified either due to abnormal liver biochemistry, versus symptoms or abdominal imaging consistent with portal hypertension (PHT). Of these cases, 15 required treatment for PHT and 1 patient received a liver transplant\(^6^4\). Meanwhile, studies in children have assessed smaller cohorts of patients: Cuffari et al. in 1996 found 1 in 25 adolescents (4%) developed hepatitis, confirmed with liver biopsy\(^6^5\). Grossman et al. in 2008 assessed a cohort of 30 children < 6 years of age who were prescribed elevated doses of azathioprine (3mg/kg/day)
due to suspected differences in pharmacokinetics in this age group. Two patients (6.7%) developed abnormal liver biochemistry that resolved without dose discontinuation. Both these patients had normal levels of thiopurine methyltransferase (TPMT), the enzyme implicated in toxicity when deficient. Another study of 22 children with IBD by Kader et al. in 2000 also found that TPMT levels did not correlate with hepatotoxicity in the 2 children who developed hepatitis. A dose monitoring strategy for thiopurine administration may include liver biochemistry at baseline, and then weekly for the first 4 weeks of therapy, twice monthly for the following second and third months, and then monthly.

Other IBD Therapies and the Liver

Other IBD treatments have reported rare cases of hepatic disturbance. Within the 5-aminosalicylic acid class of medications, sulfasalazine has been associated with significant cases of fever, rash, jaundice, hepatomegaly that either progressed to acute liver failure, or resolved with corticosteroid administration. These cases were suspected to be idiosyncratic reactions to the sulfa moiety within this compound.

Hepatic steatosis can occur with glucocorticoid administration due to increased lipogenesis and decreased free fatty acid oxidation pathways. Children with IBD who are treated with corticosteroids are at risk of this complication. However, despite frequent use of this medication, there are few, if any, biopsy-proven case reports of this complication in adults or children. This likely indicates that the short courses of steroids used in IBD do not frequently cause significant liver disease.

Hepatotoxicity with anti-TNF use is rare. However, after 35 cases of hepatitis, cholestasis, and acute liver failure were reported to the FDA medwatch program, in patients using anti-TNF agents, a black-box warning was issued to alert physicians to the possible associated hepatotoxicity. Also concerning were the findings of increased hepatitis B reactivation in patients on these biologic agents. It is possible that endogenous TNF promotes viral clearance and its loss contributes to the increased viral replication. Cyclosporine was studied in 111 patients with IBD in 2008 and 21 (19%) were found to develop mild elevations in liver
biochemistry that did not cause any adverse effects. Only one patient was investigated further and was found to have biliary sludge on ERCP\textsuperscript{74}.

Finally, total enteral nutrition (TEN) is a therapy that has been used to both induce and maintain remission in Crohn disease. This treatment is preferred to the use of total parenteral nutrition (TPN) due to a more physiologic delivery of nutrition as well as a lower side effect profile thus avoiding the feared intestinal failure-associated liver disease. Two previous studies have studied the liver biochemistry profile in TEN for IBD: Dolz et al. in 1989 did not find any statistical differences between IBD patients on TEN versus not on TEN (though this was not randomized). Meanwhile, Schotorjje and Hoekstra in 2010 found a transient increase in liver biochemistry in the first 6 weeks of treatment, that subsequently normalized spontaneously without progression to liver disease, over a mean follow-up period of 2.1 years\textsuperscript{75, 76}.

**Hepatosplenic T-cell Lymphoma**

Hepatosplenic T-cell Lymphoma (HSTCL) is a rare malignancy that has been associated with immunocompromized patients. In this disease, activated T cells appear to infiltrate the sinusoids of the liver, spleen, and bone marrow. Though rare, this lymphoma is aggressive, extraordinarily difficult to treat and almost uniformly fatal. In general, the median age at diagnosis is 32 years (range 10–80 years), with a 70% male predominance\textsuperscript{77}. Retrospective analysis of 2 case series (up to 61 patients total) demonstrated that 98% of patients presented with splenomegaly, 77% with hepatomegaly, 70% with systemic “B” symptoms, and 29% with jaundice\textsuperscript{77}. Investigations performed on these patients revealed bone marrow infiltration in 100%, liver enzyme abnormalities in 40%, thrombocytopenia in 89%, and anemia in 80%. In the IBD population, reports of HSTCL have been published in children and young adults who were treated with concomitant infliximab and 6-MP/AZA therapy\textsuperscript{77}. A retrospective review of all cases of HSTCL in patients on anti-TNF agents that were reported to the US Food and Drug Administration between 2004 and 2009 was recently published\textsuperscript{78}. A total of 25 cases were found, 18 of which were patients under 50 years of age. Of the 18, the mean age was 23.6 years (±7.1 years), 17/18 (94.4%) were male, all had IBD and all were co-treated with 6-MP/AZA. Unfortunately, 14/18 (77.7%) were confirmed to have died at the time of publication. This risk of malignancy has led to a change in practice, as many doctors and families are now reluctant to use combinations of
biologic agents with 6-MP/AZA. Clearly, in a patient where this regimen has been employed and the symptoms listed above develop, further investigation and an oncology consultation is warranted.

Other liver diseases and IBD

Cholelithiasis

There is an increased prevalence of gallstone disease in patients with IBD. Though rare in pediatrics, 2.3% of 1649 children with IBD in an American consortium developed gallstones compared to a population prevalence of approximately 0.88%–0.99% in a British prospective cohort (<30 years of age).\(^{79, 80}\) Previous adult studies have described a 13–28% prevalence rate in CD and 4.6% in UC\(^{81-83}\). Significant associations with an increased rate of cholelithiasis include: increased age, increased duration of disease, increased extent of disease, and prior intestinal surgery. The etiology of this problem is unclear and is hypothesized to be due to increased enterohepatic recirculation of bilirubin as elevated levels of bilirubin (increased three to tenfold) can be found in the bile of patients with ileal disease, or past resections, compared to patients with colitis\.\(^{84}\) In contrast, a study in Crohn disease found an increased prevalence of gallstones with colonic disease compared to disease limited to the ileum\.\(^{82}\) Akerlund et al., in 2000, assessed the bile biochemical composition in patients with UC status-post colectomy undergoing ileal pouch re-anastomosis. They found the bile composition in patients with IBD to be different from non-IBD patients undergoing cholecystectomy for cholelithiasis, but similar to healthy patients who had cholecystectomy for other reasons\.\(^{85}\) While the pathophysiology remains uncertain in IBD, once gallstones have been identified, cholecystectomy may be indicated as these patients have a higher propensity to develop further stones.

Liver Abscesses

Cases of liver abscesses in both children and adults with IBD have been reported. They can arise at presentation of IBD or while on immunosuppressant therapy. A 17-year-old adolescent with ileocecal CD, who presented with right upper quadrant abdominal pain, hepatomegaly, and chills, was found to have multiple liver abscesses. These lesions required antibiotic treatment
over 6 weeks, but subsequently resolved\textsuperscript{86}. In the adult literature, liver abscesses arise more commonly in Crohn disease than in ulcerative colitis\textsuperscript{87, 88}. An elevated index of suspicion is required with the aforementioned symptoms as the diagnosis can usually be made relatively easily on ultrasonography.

**Vascular Lesions**

There is an increased propensity for developing vascular thrombotic lesions in the IBD population. Hepatic vein thrombosis (HVT: Budd-Chiari syndrome) has also been reported in IBD, and requires a high index of suspicion as diagnosis on imaging can be challenging. HVT causes an impaired vascular outflow from the liver with subsequent hepatomegaly, ascites, and portal hypertension\textsuperscript{89}. One case report describes a 12-year-old girl who presented to hospital 6 months after her diagnosis with IBD. She had uncontrolled IBD with bloody diarrhea and anemia. She developed abnormal liver biochemistry, and an abdominal ultrasound demonstrated hepatomegaly. A Doppler examination was performed and the hepatic vein had a decreased diameter with old organized thrombi in the lumen\textsuperscript{89}. Diuretics and treatment for her underlying IBD were administered and she subsequently improved (without anticoagulant therapy). In 1986, 2 cases of HVT in adults with IBD were described, and both patients presented with ascites and hepatomegaly. One patient was diagnosed on necropsy, while the other had evidence of sinusoidal dilatations on liver biopsy that was consistent with HVT. He received treatment with diuretics as well as anticoagulation therapy and he improved with time\textsuperscript{90}. HVT is uncommon; however, it occurs more frequently in patients with IBD and requires particular consideration. Initially, investigations with Doppler ultrasonography can be helpful in identifying HVT. If there is uncertainty, computed tomography, magnetic resonance imaging, liver biopsy, or percutaneous venography can be used for diagnostic confirmation. The therapeutic goal is to prevent progression of disease with thrombolytic therapy, anticoagulant therapy, angioplasty or vascular stents, although caution should be exerted in patients with active bleeding\textsuperscript{89}. Symptomatic treatment of HVT includes management of ascites with diuretics and paracentesis.
Non-Alcoholic Fatty Liver Disease

There are no current reports in the literature of non-alcoholic fatty liver disease (NAFLD) in pediatric IBD, however, with evidence of increased obesity in developed countries, this may change in the near future. In a multicenter study assessing BMI of children at time of diagnosis with IBD, 10% of 456 children with CD and 20-30% of 156 children with UC were found to be overweight or at risk of being overweight.

**A clinical approach to children with IBD and liver abnormalities**

Patients with IBD who develop abnormal liver biochemistry or physical stigmata of liver disease, may be manifesting signs or symptoms of a range of potential diagnoses, as described in this chapter. Based on the available (limited) literature we suggest the following algorithm to approach hepatic involvement in pediatric IBD (Figure 2.4). All children with IBD should have routine liver biochemistry with liver enzymes, albumin, and bilirubin at 6–12 month intervals when the child is well. This schedule may be increased when the child is unwell or if they are being treated with hepatotoxic medications. If abnormalities in liver enzymes are detected and are low-grade, the tests may be repeated in 1–2 weeks, to ensure they are not increasing acutely, and subsequently followed for the first few months. With higher elevations, or if other stigmata of liver disease appear, including hepatomegaly, splenomegaly, or jaundice, further investigations should be considered including: autoimmune antibody, viral hepatitis and celiac testing, ceruloplasmin and alpha-1-antitrypsin level assessment, abdominal ultrasound, MRCP, and/or liver biopsy. The cut-off points between low and high-grade liver biochemistry disturbance is often a point of contention between hepatologists. With limited evidence in children with IBD, we suggest that 2–3 times above the upper limit of normal may be considered a significant elevation that requires further investigation. Consultation with a hepatologist may also be appropriate at this stage of evaluation. Subsequent evaluation may also include an MRCP and/or liver biopsy. Further studies are required to devise an evidence-based diagnostic algorithm for the approach to abnormal liver biochemistry in children with IBD.
This proposed algorithm is based on our current understanding of liver diseases in IBD. Further research is required for its validation.
Chapter 3
Abnormal Liver Enzymes Are Common In Pediatric Inflammatory Bowel Disease: Prevalence, Natural History and Associations

Summary

Background & Aims: Abnormalities in liver enzymes (LEs) associated with pediatric inflammatory bowel diseases (IBD) are understudied. We undertook to describe the development of abnormal LEs in pediatric IBD.

Methods: We ascertained a cohort of 300 carefully phenotyped children with IBD, and collected retrospective clinical and laboratory data. A Kaplan-Meier analysis determined the time to development of different thresholds of abnormal LEs. Associations between clinical variables and the development of abnormal LEs were determined.

Results: The probability of developing the first episode of abnormal LEs, prior to adult healthcare transition, was 58.1% (16.3% by 1 month and 29.5% by 1 year post-IBD diagnosis). The probability of developing a higher threshold of elevated LEs (>2x upper limit of normal [ULN]) before transition was 31.3%, while persistently abnormal LEs (30-90 days) was 23.0% at 1-2xULN, and 10.1% at >2xULN. Corticosteroids, antibiotics, and exclusive enteral nutrition presented strongly positive associations with development of abnormal LEs (hazard ratio [HR] 1.70 [95%CI {confidence interval} 1.08, 2.66], HR 4.62 [95%CI 2.96, 7.21], HR 3.65 [95%CI 1.39, 9.61] respectively). Findings were similar at different thresholds of abnormal LEs. Methotrexate and adalimumab were positively associated with abnormal LEs when including LEs data >90 days post-IBD diagnosis (HR 2.32 [95%CI 1.30, 4.15] p=0.005, and HR 3.43 [95%CI 1.50, 7.84] p=0.004 respectively).

Conclusions: Abnormal LEs are common in pediatric IBD, occur early following IBD diagnosis, and are likely associated with uncontrolled IBD, as medications used to induce remission had a strongly positive HR. Abnormal LEs are typically low-grade and infrequently require invasive investigations as chronic liver diseases are relatively uncommon.
Background & Aims

Analyses of Ontario administrative data available since 1986 reveal that the incidence of inflammatory bowel disease (IBD), developing prior to adulthood, has been steadily increasing over the past two decades. Liver disease causes significant morbidity and mortality in a subset of children and adults with IBD, including the need for liver transplantation. The spectrum of chronic liver diseases associated with IBD includes primary sclerosing cholangitis (PSC), autoimmune hepatitis (AIH), the overlap syndrome autoimmune sclerosing cholangitis (ASC), drug-induced liver injury (DILI), and several rarer conditions (e.g., vascular lesions, nodular regenerative hyperplasia, hepato-splenic T cell lymphoma). Other common liver diseases may coexist with IBD, including non-alcoholic fatty liver disease and chronic viral hepatitis.

Abnormal liver enzymes (LEs) in children with IBD may either indicate the presence of significant hepatic disease or a non-specific transient increase, without any long-term hepatic consequences. It may be difficult to exclude serious chronic liver diseases without biliary imaging (e.g., magnetic resonance cholangiopancreatography), or invasive investigations, such as a liver biopsy. The determination of when and how to investigate abnormal LEs in pediatric IBD has been understudied.

Studies addressing the prevalence of abnormal LEs in adults with IBD have reported rates ranging between 15-63%. While a small group of patients had higher elevations of LEs, (>5 times the upper limit of normal [ULN]), the majority had milder changes, the significance of which was not clear in at least 28% of patients. There is conflicting evidence as to the association of active bowel inflammation with the development of abnormal LEs, as not all IBD cohorts observed transient disturbances in LEs in patients with uncontrolled IBD.

The pediatric data regarding abnormal LEs in IBD are limited. Nemeth et al reported a 52% prevalence of abnormal LEs, over 12 years of follow-up, in a small cohort of 46 children with IBD. Of the 20% with LEs elevated >3xULN, half were diagnosed with small duct PSC on liver biopsy. In a cohort of 555 children with IBD, Hyams et al identified 75 (14%) with abnormal LEs defined as alanine aminotransferase (ALT) ≥ 80 U/L (~2xULN). Persistent elevations, over a minimum of 6 months, were observed in 14 patients (2.5%), all of which were diagnosed with a chronic liver disease (PSC, AIH). A small group of patients (12, 2.2%) developed elevations that persisted over a shorter period of time (<6 months); these events were attributed to
medication use, uncontrolled IBD, bowel obstruction, viral illness, and blood transfusion. Recently, a study of the Pediatric IBD Collaborative Research Group Registry data reported that elevated LEs (ALT and gamma glutamyltranspeptidase [GGT] > 50U/L), at the time of diagnosis of IBD, were significantly associated with a diagnosis of chronic liver diseases\textsuperscript{99}. However, longitudinal LEs data were not available in this database, highlighting the need for further study of hepatic abnormalities in IBD.

We sought to systematically study a well-characterized cohort of pediatric IBD patients to determine the prevalence and natural history of the first episode of abnormal LEs. Furthermore, we endeavored to identify specific clinical variables associated with the time to development of the first abnormal LEs event. This study is an essential stepping-stone towards the development of an evidence-based algorithm in the management of hepatic abnormalities in children with IBD.

**Methods**

**Study Design and Population**

The Hospital for Sick Children (SickKids), in Toronto, is a major referral base for the Greater Toronto Area (GTA) and provides care from the time of first presentation for children and adolescents aged less than 18 years with IBD. All patients are systematically examined via ileocolonoscopy, upper endoscopy, and small bowel imaging at diagnosis, for accurate IBD phenotyping according to the Paris classification.

This was a retrospective study of a random sample of children all newly diagnosed with IBD (i.e. an inception cohort) at age less than 18 years at SickKids Hospital in Toronto between January 2000 and October 2011. This study was approved by the SickKids Hospital and the University of Toronto research ethics boards. The study cohort was ascertained from a pediatric IBD database that was prospectively created at SickKids. In order to ensure that all patients with IBD were captured by the database, the clinical charts were searched via the SickKids medical records department for the Crohn disease (CD) and ulcerative colitis (UC) ICD-9 & ICD-10 codes (K50 and K51 respectively).
To avoid other causes of elevated LEs as confounders, we excluded patients who were on hepatotoxic medications prior to, and unrelated to, their diagnosis of IBD. We excluded patients who had been previously diagnosed with non-IBD-related chronic liver diseases before clinical presentation with IBD symptoms. However, we did include those patients who had developed IBD-related chronic liver diseases (e.g., PSC, AIH or ASC) prior to a diagnosis of IBD so that we could describe the prevalence of these chronic liver diseases in our cohort. These patients were excluded from subsequent analyses of tests of association. Finally, we also excluded patients who were diagnosed at other institutions, and subsequently referred to SickKids, since LEs data were not consistently available from the onset of the IBD diagnosis.

Due to the identification of a significantly large number of cases (>1000), a random representative sample of the patients was generated to improve feasibility. After including the patients with IBD identified from medical records, the sample (1117) was randomized using an online random list generator. The patients on the randomized list were sequentially reviewed according to inclusion and exclusion criteria to generate a cohort of 300 cases.

**Collection of Longitudinal Liver Biochemistry**

Using the electronic laboratory database, we collected all available LE results, until August 2012. Many patients also had LEs investigations performed at private laboratories. Manual data collection of these LE results from scanned reports was performed and managed using REDCap (Research Electronic Data Capture), a secure electronic data capture tool hosted at SickKids. To assess the quality of the manual data entry, errors were verified using a visual record check in a random sample of 5% of the study cohort (list generated using the same online random list generator); the error rate percentage was recorded.

**Other Data Collection**

General demographic, as well as disease-specific information was obtained from the SickKids pediatric IBD database collected and managed via Microsoft Access® software (Microsoft Corporation, Redmond, WA, USA). This included: age, sex, IBD diagnosis, date of IBD diagnosis, IBD disease location & behavior, weight & height at time of IBD diagnosis, past &
family history of autoimmune diseases, and extraintestinal manifestations of IBD. Patients with the IBD-U designation were treated as having UC in all analyses, due to the small number of patients in this sub-group. All medical charts were thoroughly reviewed to confirm the presence or absence of a diagnosis of chronic liver disease (e.g. PSC, AIH, ASC), both on histopathology and liver imaging, based on current recommended criteria. All IBD-related medications (including nutrition [EEN]), laboratory investigations (including: ALT, GGT, anti-neutrophil cytoplasmic antibodies [ANCA], & viral serology), and severity of IBD at the time of IBD diagnosis (as measured by validated CD & UC severity scores) were recorded. In patients with CD, the Pediatric Crohns Disease Activity Index [PCDAI] and mathPCDAI (an abbreviated PCDAI) scores were calculated, while Pediatric Ulcerative Colitis Activity Index [PUCAI] scores were calculated in patients with UC or IBD-U. Finally, liver investigations such as findings on diagnostic imaging (ultrasound, MRCP, ERCP), and histopathology were obtained from original reports in order to confirm diagnoses of liver disease.

**Statistical Analysis**

The analysis of the demographic, IBD phenotypic data and the presence of liver-related outcomes included means with standard deviations for normally distributed factors, or medians with interquartile ranges for non-normally distributed factors. Normality was assessed using the Shapiro-Wilk test. Comparisons utilized either Wilcoxon Rank-Sum test or logistic regression where appropriate, for continuous or categorical dependent variables. Follow-up time was calculated from the time of IBD diagnosis to the last LEs measurement.

A time to event analysis, using the Kaplan-Meier (KM) method, was performed on the LEs data (ALT and GGT) to describe the incidence rate of first occurrence of abnormal LEs in this population. The diagnosis of IBD with endoscopic biopsies (or less commonly, intestinal imaging) was used as **time-zero** for the analysis. Biochemistry measurements obtained within the previous three months prior to the IBD diagnosis were included and reset to the date of the IBD diagnosis for the purposes of the analysis. Patients were censored after the time of their last LEs measurement.
The primary endpoint for this analysis was any abnormality above ULN for age in ALT (generally >40 U/L), or GGT (varied by age and sex). This low threshold was chosen in light of recent literature, demonstrating that healthy children, without liver disease, typically have lower aminotransferase levels than previously believed\textsuperscript{106,107}. Thus, any abnormalities above the current ULN for age could be indicative of hepatic involvement. The median time to the development of any abnormality in ALT or GGT was determined using the KM analysis, as well as the proportion of children who developed abnormal LEs at 1, 2 and 5 years post IBD diagnosis. We also performed the analysis with stratification according to the type of IBD, as differences in the natural history of liver involvement could exist between CD and UC.

Some medications, such as methotrexate and thiopurines, are only utilized for the maintenance of remission of IBD. Instances where patients develop abnormal LEs within the first month of diagnosis with IBD, will typically be in the setting of agents utilized for induction of remission, such as: corticosteroids, exclusive enteral nutrition (EEN), or biologic agents for CD, and corticosteroids or sulfasalazine/mesalazine for UC. Thus, an analysis was also performed with the exclusion of LEs data from the first 90 days following the IBD diagnosis to attempt to capture the hepatotoxic effects of these maintenance medications.

Higher thresholds of LEs elevation, as well as persistence of these abnormalities, may characterize a more severe phenotype of perturbed LEs. Therefore the KM method was utilized to describe the time to development of higher LEs elevations >2xULN after a diagnosis of IBD. In order to describe persistent abnormal LEs in the IBD population, two abnormal LEs measurements were required to be present with a minimum interval of 30 days and a maximum interval of 90 days. Chronic liver diseases, such as PSC or drug hepatotoxicity, can lead to elevations that can persist for weeks to months. Thus, it was important to set a maximum limit as two isolated abnormalities occurring within a longer interval (e.g. separated by one year) might reflect two different underlying disease processes. Analyses were performed for the development of persistence of abnormal LEs (30-90 days) at low-grade elevations (1-2x ULN) and higher threshold elevations ≥2x ULN.

We undertook to assess if there were associations between specific IBD phenotypic clinical variables and the development of abnormal LEs. Univariate Cox-proportional hazards models were built including the variables: IBD phenotype (CD, UC), age at diagnosis, sex, growth
parameters at diagnosis, extra intestinal manifestations of IBD, disease severity at diagnosis, history of autoimmune diseases, family history of IBD or other autoimmune diseases, specific laboratory investigations and medications. Medications were treated as time varying covariates in the analysis, and end dates were extended by 7 days to account for the persistence of hepatotoxic effects following discontinuation. Although we examined the association of abnormal LEs with extra intestinal manifestations of IBD, the development of PSC/ASC was excluded from the statistical models, as the diagnosis of these chronic liver diseases depend on the presence of abnormal LEs, which is a circular argument and invalid in such an analysis. Tests of associations with other thresholds of abnormal LEs (≥2x ULN, 1-2x ULN and persistent over 30-90 days, >2x ULN and persistent over 30-90 days) were also performed. The Cox-proportional hazards multivariable models were built in a forward stepwise fashion including IBD phenotype (CD versus UC) as well as variables that were statistically significant in the univariate models (p<0.1). These models were tested for violations of assumptions via verification of: random censoring, model over-specification, and multicollinearity. All statistical analyses were completed using SAS/STAT® software, version 9.3 (SAS Institute, Cary, North Carolina, USA).

Results

Study Cohort

According to the search criteria, a total of 1193 patients with IBD were identified. One patient with prior liver disease, and 143 patients who were diagnosed with IBD at other institutions were excluded from our study. Following random selection, 300 children with IBD were included as a sample of convenience in our study cohort (Table 3.1). Of this group, 299 patients had laboratory investigations and 296 patients had LEs (ALT, GGT) measured at some point in time after their diagnosis of IBD. As a marker of longitudinal assessments, 288 (96%) had at least 2 LEs measurements, while 243 (81%) had at least 5 measurements. Overall, the median number of LEs measurements was 11 with an interquartile range (IQR) of 5, 21. Manual data entry was performed for 25.9% of LEs laboratory results with a data entry error rate of approximately 2.4% (0.62% of the overall LEs data).
CD phenotype was present at a rate of 54.3% in the study cohort, while 33.3% of patients were diagnosed with UC (Table 3.1). There was a slight preponderance of males (59%), and patients were diagnosed in early adolescence at a median age of 12.2 years (IQR 9.4, 14.3). Exclusive colonic disease was present in 20.8% of children with CD, while 84.1% had at least some small bowel disease. The majority of patients with UC had pancolonic disease (56.3%). Corticosteroids and antibiotics were utilized for the treatment of all IBD phenotypes at high rates (>70%), while exclusive enteral nutrition, methotrexate, and biologics were more commonly taken by patients with CD (20-40%) compared to those with UC (<13% [Table 3.2]).

Table 3.1: Demographic data of the IBD cohort at the time of IBD diagnosis

<table>
<thead>
<tr>
<th>Variables</th>
<th>Overall Cohort (N=300)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex (Male) [N, %]</td>
<td>177, 59%</td>
</tr>
<tr>
<td>Age At Diagnosis (years) [median, IQR]</td>
<td>12.2, [9.4, 14.3]</td>
</tr>
</tbody>
</table>

IBD Diagnosis

<table>
<thead>
<tr>
<th>Location of disease</th>
<th>CD* [N, %]</th>
<th>UC [N, %]</th>
<th>IBD-U [N, %]</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1 [N, %]</td>
<td>43, [27.0%]</td>
<td>6, [4.4%]</td>
<td>-</td>
</tr>
<tr>
<td>L2 [N, %]</td>
<td>33, [20.8%]</td>
<td>39, [28.9%]</td>
<td>14, [10.4%]</td>
</tr>
<tr>
<td>L3 [N, %]</td>
<td>30, [18.9%]</td>
<td>14, [10.4%]</td>
<td>76, [56.3%]</td>
</tr>
<tr>
<td>[N, %]</td>
<td>-</td>
<td>39, [28.9%]</td>
<td>76, [56.3%]</td>
</tr>
</tbody>
</table>

Growth Parameters at Diagnosis

<table>
<thead>
<tr>
<th>CD</th>
<th>weight z-score [mean, SD]</th>
<th>height z-score [median, IQR]</th>
<th>BMI z-score [median, IQR]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.734, [1.42]</td>
<td>-0.374, [-1.01, 0.319]</td>
<td>-0.571, [-1.65, 0.304]</td>
</tr>
<tr>
<td>UC &amp; IBD-U</td>
<td>weight z-score [mean, SD]</td>
<td>height z-score [median, IQR]</td>
<td>BMI z-score [median, IQR]</td>
</tr>
<tr>
<td></td>
<td>-0.248, [1.35]</td>
<td>-0.182, [-0.834, 0.860]</td>
<td>-0.071, [-0.916, 0.893]</td>
</tr>
</tbody>
</table>

IBD Severity at Diagnosis

<table>
<thead>
<tr>
<th>CD</th>
<th>PCDAI [mean, SD]</th>
<th>mathPCDAI [mean, SD]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>36.5, [15.9]</td>
<td>58.9, [26.0]</td>
</tr>
<tr>
<td>UC &amp; IBD-U</td>
<td>PUCAI [median, IQR]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>55.0 [35.0, 65.0]</td>
<td></td>
</tr>
</tbody>
</table>
### Family History of IBD

<table>
<thead>
<tr>
<th></th>
<th>CD</th>
<th>UC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall [N, %]</td>
<td>92, [30.7%]</td>
<td>44, [27.0%]</td>
</tr>
<tr>
<td>First degree relative [N, %]</td>
<td>41, [24.2%]</td>
<td>20, [23.0%]</td>
</tr>
<tr>
<td>Second degree relative [N, %]</td>
<td>25, [21.4%]</td>
<td>13, [20.3%]</td>
</tr>
</tbody>
</table>

### Family History of Autoimmune disease

<table>
<thead>
<tr>
<th></th>
<th>CD</th>
<th>UC</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIH [N, %]</td>
<td>1, [0.33%]</td>
<td>0</td>
</tr>
<tr>
<td>Juvenile Idiopathic Arthritis [N, %]</td>
<td>1, [0.33%]</td>
<td>1, [0.61%]</td>
</tr>
<tr>
<td>Rheumatoid Arthritis [N, %]</td>
<td>1, [0.33%]</td>
<td>0</td>
</tr>
<tr>
<td>Arthritis (specified) [N, %]</td>
<td>3, [1.0%]</td>
<td>2, [1.2%]</td>
</tr>
<tr>
<td>Systemic Lupus Erythematosus [N, %]</td>
<td>1, [0.33%]</td>
<td>0</td>
</tr>
<tr>
<td>Alopecia [N, %]</td>
<td>1, [0.33%]</td>
<td>0</td>
</tr>
<tr>
<td>Thyroid disease [N, %]</td>
<td>16, [5.3%]</td>
<td>9, [5.5%]</td>
</tr>
<tr>
<td>Celiac Disease [N, %]</td>
<td>3, [1.0%]</td>
<td>3, [1.8%]</td>
</tr>
<tr>
<td>Diabetes Mellitus [N, %]</td>
<td>9, [3.0%]</td>
<td>5, [3.1%]</td>
</tr>
<tr>
<td>Atopy [N, %]</td>
<td>26, [8.7%]</td>
<td>18, [11.0%]</td>
</tr>
</tbody>
</table>

N: number of patients, IQR: Interquartile range, IBD: Inflammatory Bowel Disease, CD: Crohn disease, UC: Ulcerative Colitis, IBD-U: Inflammatory Bowel Disease–Undetermined, SD: standard deviation

†Ulcerative Colitis Paris classification: E1: Ulcerative proctitis, E2: Left sided disease, E3: Extensive disease (proximal to the splenic flexure, but distal to the hepatic flexure), E4: Pancolitis

There were 13 patients (4.3%) with PSC, plus 5 patients (1.7%) with ASC in this cohort diagnosed during the study follow-up period (Table 3.2). All 5 children with ASC presented with elevated LEs, and abnormal liver biopsies on further investigation, at the time of IBD diagnosis. Three patients had developed PSC prior to their diagnosis of IBD and were excluded from further analyses. Another 5 of the 13 PSC patients had elevated LEs at the time of presentation with symptoms of IBD, and were diagnosed with PSC and IBD concurrently. The remainder (5/13) developed signs of PSC following their diagnosis of IBD. The vast majority of patients had exclusively colonic inflammation, and were labeled as UC or IBD-U.
Table 3.2: Longitudinal clinical data of the IBD cohort

<table>
<thead>
<tr>
<th>Variables</th>
<th>Overall Cohort (N=300)</th>
<th>CD</th>
<th>UC &amp; IBD-U</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Extra-intestinal Manifestations of IBD</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSC [N, %]</td>
<td>10, [3.3%]</td>
<td>1, [0.61%]</td>
<td>9, [6.6%]</td>
</tr>
<tr>
<td>AIH [N, %]</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ASC [N, %]</td>
<td>5, [1.7%]</td>
<td>0</td>
<td>5, [5.0%]</td>
</tr>
<tr>
<td>Large Joint Arthritis [N, %]</td>
<td>14, [4.7%]</td>
<td>12, [7.4%]</td>
<td>2, [1.5%]</td>
</tr>
<tr>
<td>Small Joint Arthritis [N, %]</td>
<td>7, [2.3%]</td>
<td>6, [3.7%]</td>
<td>1, [0.7%]</td>
</tr>
<tr>
<td>Ankylosing Spondylitis [N, %]</td>
<td>2, [0.67%]</td>
<td>2, [1.2%]</td>
<td>0</td>
</tr>
<tr>
<td>Sacroileitis [N, %]</td>
<td>5, [1.7%]</td>
<td>3, [1.8%]</td>
<td>2, [1.5%]</td>
</tr>
<tr>
<td>Arthralgia [N, %]</td>
<td>3, [1.0%]</td>
<td>2, [1.2%]</td>
<td>1, [0.7%]</td>
</tr>
<tr>
<td>Psoriasis [N, %]</td>
<td>1, [0.33%]</td>
<td>1, [0.61%]</td>
<td>0</td>
</tr>
<tr>
<td>Erythema Nodosum [N, %]</td>
<td>14, [4.7%]</td>
<td>13, [8.0%]</td>
<td>1, [0.7%]</td>
</tr>
<tr>
<td>Pyoderma Gangrenosum [N, %]</td>
<td>3, [1.0%]</td>
<td>2, [1.2%]</td>
<td>1, [0.7%]</td>
</tr>
<tr>
<td>Oral Ulcers [N, %]</td>
<td>2, [0.67%]</td>
<td>2, [1.2%]</td>
<td>0</td>
</tr>
<tr>
<td>Orofacial Granulomatosis [N, %]</td>
<td>1, [0.33%]</td>
<td>1, [0.61%]</td>
<td>0</td>
</tr>
<tr>
<td>Iritis/Uveitis [N, %]</td>
<td>1, [0.33%]</td>
<td>1, [0.61%]</td>
<td>0</td>
</tr>
<tr>
<td><strong>Medications</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methotrexate [N, %]</td>
<td>64, [39.2%]</td>
<td>6, [4.4%]</td>
<td></td>
</tr>
<tr>
<td>Azathioprine [N, %]</td>
<td>37, [22.7%]</td>
<td>29, [21.2%]</td>
<td></td>
</tr>
<tr>
<td>6-Mercaptopurine [N, %]</td>
<td>2, [1.2%]</td>
<td>2, [1.5%]</td>
<td></td>
</tr>
<tr>
<td>Biologics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infliximab [N, %]</td>
<td>61, [37.4%]</td>
<td>18, [13.1%]</td>
<td></td>
</tr>
<tr>
<td>Adalimumab [N, %]</td>
<td>24, [14.7%]</td>
<td>4, [2.9%]</td>
<td></td>
</tr>
<tr>
<td>Certolizumab [N, %]</td>
<td>1, [0.61%]</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Golimumab [N, %]</td>
<td>2, [1.2%]</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>5-Acetylsalicylic Acid [N, %]</td>
<td>54, [33.1%]</td>
<td>77, [56.2%]</td>
<td></td>
</tr>
<tr>
<td>Sulfasalazine [N, %]</td>
<td>43, [26.4%]</td>
<td>76, [55.5%]</td>
<td></td>
</tr>
<tr>
<td>Corticosteroids</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prednisone PO [N, %]</td>
<td>116, [71.2%]</td>
<td>105, [76.6%]</td>
<td></td>
</tr>
<tr>
<td>Budesonide PO [N, %]</td>
<td>32, [19.6%]</td>
<td>2, [1.5%]</td>
<td></td>
</tr>
<tr>
<td>Methylprednisolone IV [N, %]</td>
<td>41, [25.2%]</td>
<td>52, [38.0%]</td>
<td></td>
</tr>
<tr>
<td>Hydrocortisone IV [N, %]</td>
<td>1, [0.61%]</td>
<td>2, [1.5%]</td>
<td></td>
</tr>
<tr>
<td>Metronidazole [N, %]</td>
<td>115, [70.6%]</td>
<td>48, [35.0%]</td>
<td></td>
</tr>
<tr>
<td>Ciprofloxacin [N, %]</td>
<td>68, [41.7%]</td>
<td>22, [16.1%]</td>
<td></td>
</tr>
<tr>
<td>Vancomycin [N, %]</td>
<td>1, [0.61%]</td>
<td>2, [1.5%]</td>
<td></td>
</tr>
<tr>
<td>Tacrolimus [N, %]</td>
<td>1, [0.61%]</td>
<td>5, [3.6%]</td>
<td></td>
</tr>
<tr>
<td>Ursodeoxycholic Acid [N, %]</td>
<td>0</td>
<td>11, [8.0%]</td>
<td></td>
</tr>
<tr>
<td>Exclusive Enteral Nutrition (Tolerex® or Modulen IBD®) [N, %]</td>
<td>37, [22.7%]</td>
<td>1, [0.7%]</td>
<td></td>
</tr>
<tr>
<td><strong>History of other Autoimmune diseases</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thyroid disease [N, %]</td>
<td>5, [1.7%]</td>
<td>2, [1.2%]</td>
<td>3, [2.2%]</td>
</tr>
<tr>
<td>Celiac Disease [N, %]</td>
<td>1, [0.33%]</td>
<td>0</td>
<td>1, [0.7%]</td>
</tr>
<tr>
<td>IDDM [N, %]</td>
<td>2, [0.67%]</td>
<td>0</td>
<td>2, [1.5%]</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------</td>
<td>---</td>
<td>-----------</td>
</tr>
<tr>
<td>Vitiligo [N, %]</td>
<td>1, [0.33%]</td>
<td>1, [0.61%]</td>
<td>0</td>
</tr>
<tr>
<td>Asthma [N, %]</td>
<td>3, [1.0%]</td>
<td>2, [1.2%]</td>
<td>1, [0.7%]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Autoantibody positivity</th>
<th>CD</th>
<th>UC</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANCA or ANA or ASMA or ALKM [N, %]</td>
<td>82, [36.3%]</td>
<td>20, [17.9%]</td>
</tr>
<tr>
<td>ANCA [N, %]</td>
<td>87, [41.6%]</td>
<td>12 [12.0%]</td>
</tr>
<tr>
<td>ASCA [N, %]</td>
<td>13, [23.2%]</td>
<td>13, [36.1]</td>
</tr>
</tbody>
</table>


Other liver investigations performed in the overall cohort included: testing for viral hepatitis B and C infection (0/24 abnormal results), and alpha-1-antitrypsin deficiency (0/107 abnormal results). Abdominal ultrasonography demonstrated signs of fatty infiltration of the liver in 18 of the 129 patients examined. In general, the BMI z scores were low in our IBD population with a median z score of -0.571, [IQR -1.65, 0.304] for CD and a median z score of -0.071, [IQR 0.916, 0.893] for UC.

**The development of abnormal liver enzymes**

The median duration of follow-up for this cohort, as defined by the interval of time between the IBD diagnosis and the last LEs measurement, was 2.8 years (IQR 1.5, 4.8 years). Using a time to event analysis, 119 children with IBD (probability of 58.1%) developed abnormal LEs for the first time during follow-up (Figure 3.1). Approximately half of all patients with IBD first developed abnormal LEs within the first 6 years, with a median time to development of abnormal LEs of 76.2 months (95% CI – lower limit: 41.9 months). After accounting for censoring, the probabilities of developing abnormal LEs were 16.3%, 29.5%, and 47.6% at one month, one year and five years, respectively. ALT was the most commonly elevated laboratory test at the time of the first abnormal results (119/165 abnormal results, 72.1%), compared with GGT, which was elevated in 27.9% of abnormal results. When stratified by the type of IBD, there were no statistically significant differences in the development of abnormal LEs (Figure 3.2, log-rank p=0.732). When excluding all patients diagnosed with chronic liver diseases (PSC/ASC) at any
point in time, the probability of developing abnormal LEs was almost unchanged at 55.8%, with a median time to development of 82.4 months (95% CI – lower limit: 55.6 months).

**Figure 3.1: Time to event analysis of abnormal liver enzymes (LEs) following IBD diagnosis. Development of the first abnormal LEs > upper limit of normal (ULN)**
The degree of severity of LEs abnormalities in patients with IBD was assessed to determine if there were specific thresholds of severity commonly experienced by patients with IBD. In this cohort, when accounting for censoring, the probability of developing elevated LEs for the first time above 2xULN was 31.3% over 150 months of follow-up (Figure 3.3). By one and five years, the probability of developing higher thresholds of abnormal LEs was 15.9% and 22.3% respectively. The median time to development of >2xULN abnormal LEs was undefined as less than 50% of the cases reached this endpoint. ALT continued to be more commonly elevated than GGT: 64.4% (47/73 results) and 35.6% (26/73 results) respectively. There were no statistically significant differences in time to development of first abnormal LEs >2xULN between the types of IBD.
Figure 3.3: Time to event analysis of abnormal liver enzymes (LEs) following IBD diagnosis. Development of first abnormal LEs >2x ULN

“Persistence” of abnormal LEs, as defined by two or more LEs measurements at least 30 days apart and up to 90 days, was found to be present at a low threshold (1-2x ULN) in 37 patients (23.0% probability over 150 months), and at a higher threshold (>2xULN) in 21 patients (10.1% probability over 150 months). Time to event analyses were also performed (Figure 3.4, Figure 3.5). Median times to the persistence of abnormal LEs were undefined as less than 50% of the patients reached these end points.
Figure 3.4: Time to event analysis of abnormal liver enzymes (LEs) following IBD diagnosis. Development of persistently abnormal LEs (between 30 to 90 days) at a low threshold (1.2x ULN)
While all patients with PSC and ASC developed abnormal LEs, 14/15 (93.3%) PSC/ASC patients developed higher elevations >2xULN. Eleven (73%) of the patients with PSC/ASC were found to have persistent results at >2xULN at the time of their first episode of abnormal LEs, while 2 (20%) only had persistent results at 1-2xULN.

Variables associated with the development of abnormal liver enzymes

Compared to other clinical variables tested, medications used for the treatment of uncontrolled IBD were most strongly associated with the development of abnormal LEs for the first time (Table 3.3, Table 3.4). Specifically, corticosteroids, antibiotics, and EEN presented a statistically significant, and strongly positive, association. However, sulfasalazine and mesalazine had a significantly reduced risk of development of liver abnormalities. Children with CD phenotype, older age and females were 189% (HR 2.89 [95%CI 1.65, 5.05]), 11% (HR 1.11 [95%CI 1.03,
1.19], and 44% (HR 0.56 for males versus females [95%CI 0.35, 0.88]) more likely to develop abnormal LEs, respectively (Table 3.4). Among the other items tested in the Cox-proportional hazards model, no other specific clinical variables demonstrated a statistically significant association.

Table 3.3: Univariate Cox Proportional Hazards models for the development of abnormal liver enzymes above the upper limit of normal in patients with IBD*

<table>
<thead>
<tr>
<th>Variables</th>
<th>Hazard Ratio Estimates (95% Confidence Interval)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBD Diagnosis: CD versus UC</td>
<td>0.84, (0.54, 1.30)</td>
<td>0.428</td>
</tr>
<tr>
<td>Age</td>
<td>1.14, (1.07, 1.22)</td>
<td>0.0001</td>
</tr>
<tr>
<td>Sex (males versus females)</td>
<td>0.66, (0.42, 1.06)</td>
<td>0.084</td>
</tr>
<tr>
<td>PCDAI at diagnosis in CD patients</td>
<td>1.01, (0.99, 1.02)</td>
<td>0.586</td>
</tr>
<tr>
<td>PUCAI at diagnosis in UC patients</td>
<td>1.00, (0.98, 1.02)</td>
<td>0.945</td>
</tr>
<tr>
<td>Extra-intestinal manifestations of IBD**</td>
<td>0.55, (0.25, 1.20)</td>
<td>0.132</td>
</tr>
<tr>
<td>History of Autoimmune disease</td>
<td>0.77, (0.22, 2.69)</td>
<td>0.675</td>
</tr>
<tr>
<td><strong>Family history of IBD</strong></td>
<td>0.61, (0.38, 1.00)</td>
<td>0.049</td>
</tr>
<tr>
<td>Family history of autoimmune disease</td>
<td>1.39, (0.83, 2.35)</td>
<td>0.216</td>
</tr>
<tr>
<td>Weight z-score at IBD diagnosis</td>
<td>1.10, (0.91, 1.34)</td>
<td>0.323</td>
</tr>
<tr>
<td>Height z-score at IBD diagnosis</td>
<td>1.11, (0.93, 1.33)</td>
<td>0.239</td>
</tr>
<tr>
<td>BMI z-score at IBD diagnosis</td>
<td>1.07, (0.89, 1.28)</td>
<td>0.476</td>
</tr>
<tr>
<td>Autoimmune antibody positivity</td>
<td>1.04, (0.61, 1.80)</td>
<td>0.878</td>
</tr>
<tr>
<td>(ANCA/ANA/ASMA/ALKM)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANCA positive</td>
<td>1.29, (0.74, 2.25)</td>
<td>0.370</td>
</tr>
<tr>
<td>ASCA positive</td>
<td>1.04, (0.39, 2.77)</td>
<td>0.936</td>
</tr>
<tr>
<td>Medications</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biologics (infliximab/adalimumab/certolizumab/golimumab)</td>
<td>0.75, (0.39, 1.47)</td>
<td>0.409</td>
</tr>
<tr>
<td>Infliximab</td>
<td>0.62, (0.29, 1.36)</td>
<td>0.236</td>
</tr>
<tr>
<td>Adalimumab</td>
<td>1.44, (0.45, 4.63)</td>
<td>0.543</td>
</tr>
<tr>
<td>Methotrexate</td>
<td>0.85, (0.43, 1.68)</td>
<td>0.646</td>
</tr>
<tr>
<td>6-Mercaptopurine or Azathioprine</td>
<td>0.56, (0.29, 1.10)</td>
<td>0.091</td>
</tr>
<tr>
<td>Corticosteroids (oral or intravenous)</td>
<td>1.70, (1.08, 2.66)</td>
<td>0.021</td>
</tr>
<tr>
<td>Metronidazole or Ciprofloxacin</td>
<td>4.62, (2.96, 7.21)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Sulfasalazine</td>
<td>0.25, (0.13, 0.48)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Mesalazine</td>
<td>0.55, (0.32, 0.96)</td>
<td>0.035</td>
</tr>
<tr>
<td>Exclusive Enteral Nutrition (Tolerex® or Modulen IBD®)</td>
<td>3.65, (1.39, 9.61)</td>
<td>0.009</td>
</tr>
</tbody>
</table>

*Abnormal liver enzymes: 119/300 patients
**Excluding PSC/ASC

Table 3.4: Multivariable Cox Proportional Hazards model for the development of abnormal liver enzymes above the upper limit of normal in patients with IBD

<table>
<thead>
<tr>
<th>Variables</th>
<th>Hazard Ratio Estimates (95% Confidence Interval)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood Ratio (Chi Sq [df])</td>
<td>127.4 [9]</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>IBD Diagnosis: CD versus UC</td>
<td>2.89, (1.65, 5.05)</td>
<td>0.0002</td>
</tr>
<tr>
<td>Age</td>
<td>1.11, (1.03, 1.19)</td>
<td>0.005</td>
</tr>
<tr>
<td>Sex (males versus females)</td>
<td>0.56, (0.35, 0.88)</td>
<td>0.013</td>
</tr>
<tr>
<td>Medications</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-Mercaptopurine or Azathioprine</td>
<td>0.44, (0.21, 0.94)</td>
<td>0.033</td>
</tr>
<tr>
<td>Corticosteroids (oral or intravenous)</td>
<td>1.12, (0.67, 1.86)</td>
<td>0.666</td>
</tr>
<tr>
<td>Metronidazole or Ciprofloxacin</td>
<td>3.97, (2.48, 6.35)</td>
<td>0.0001</td>
</tr>
<tr>
<td>Sulfasalazine</td>
<td>0.18, (0.08, 0.40)</td>
<td>0.0001</td>
</tr>
<tr>
<td>Mesalazine</td>
<td>0.29, (0.14, 0.59)</td>
<td>0.0007</td>
</tr>
<tr>
<td>Exclusive Enteral Nutrition (Tolerex® or Modulen IBD®)</td>
<td>3.01, (1.22, 7.40)</td>
<td>0.017</td>
</tr>
</tbody>
</table>

IBD: Inflammatory Bowel Disease, CD: Crohn disease, UC: Ulcerative Colitis

With the removal of LEs obtained within the first 90 days of IBD diagnosis, corticosteroids (Hazard Ratio [HR] 2.09 [95% Confidence Interval [CI] 1.31, 3.33]; p = 0.002) and antibiotics (HR 4.39 [95% CI 2.76, 6.97] p < 0.0001) continued to demonstrate significant positive associations with the development of liver abnormalities (Table 3.5). Importantly, methotrexate and adalimumab also displayed statistically significant increased associations with abnormal LEs (HR 2.32 [95% CI 1.30, 4.15] p = 0.005 & HR 3.43 [95% CI 1.50, 7.84] p = 0.004 respectively). Sulfasalazine, on the other hand, continued to demonstrate an association with a reduced risk of development of abnormal LEs (HR 0.49 [95% CI 0.27, 0.89] p = 0.019).

Table 3.5: Univariate Cox Proportional Hazards models for the development of abnormal liver enzymes in patients with IBD – 90 days post-IBD diagnosis*

<table>
<thead>
<tr>
<th>Variables</th>
<th>Hazard Ratio Estimates (95% Confidence Interval)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBD Diagnosis: CD versus UC</td>
<td>0.79, (0.51, 1.24)</td>
<td>0.307</td>
</tr>
<tr>
<td>Age</td>
<td>1.08, (1.01, 1.16)</td>
<td>0.017</td>
</tr>
<tr>
<td>Sex (males versus females)</td>
<td>0.74, (0.46, 1.18)</td>
<td>0.208</td>
</tr>
<tr>
<td>PCDAI at diagnosis in CD patients</td>
<td>1.00, (0.98, 1.02)</td>
<td>0.807</td>
</tr>
<tr>
<td>PUCAI at diagnosis in UC patients</td>
<td>1.00, (0.98, 1.02)</td>
<td>0.734</td>
</tr>
<tr>
<td>Extra-intestinal manifestations of IBD**</td>
<td>0.49, (0.21, 1.14)</td>
<td>0.097</td>
</tr>
<tr>
<td>History of Autoimmune disease</td>
<td>1.08, (0.37, 3.16)</td>
<td>0.895</td>
</tr>
<tr>
<td>Family history of IBD</td>
<td>0.46, (0.27, 0.78)</td>
<td>0.004</td>
</tr>
<tr>
<td>Family history of autoimmune disease</td>
<td>1.29, (0.78, 2.15)</td>
<td>0.324</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>-------------------</td>
<td>-------</td>
</tr>
<tr>
<td>Weight z-score at IBD diagnosis</td>
<td>1.14, (0.96, 1.36)</td>
<td>0.147</td>
</tr>
<tr>
<td>Height z-score at IBD diagnosis</td>
<td>1.17, (0.97, 1.41)</td>
<td>0.097</td>
</tr>
<tr>
<td>BMI z-score at IBD diagnosis</td>
<td>1.10, (0.93, 1.30)</td>
<td>0.252</td>
</tr>
<tr>
<td>Autoimmune antibody positivity</td>
<td>0.92, (0.55, 1.55)</td>
<td>0.759</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Medications</th>
<th>HR</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biologics (infliximab/adalimumab/certolizumab/golimumab)</td>
<td>1.60, (0.90, 2.87)</td>
<td>0.112</td>
<td></td>
</tr>
<tr>
<td>Infliximab</td>
<td>1.11, (0.54, 2.26)</td>
<td>0.785</td>
<td></td>
</tr>
<tr>
<td>Adalimumab</td>
<td>3.43, (1.50, 7.84)</td>
<td>0.004</td>
<td></td>
</tr>
<tr>
<td>Methotrexate</td>
<td>2.32, (1.30, 4.15)</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td>6-Mercaptopurine or Azathioprine</td>
<td>0.79, (0.41, 1.52)</td>
<td>0.476</td>
<td></td>
</tr>
<tr>
<td>Corticosteroids (oral or intravenous)</td>
<td>2.09, (1.31, 3.33)</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>Metronidazole or Ciprofloxacin</td>
<td>4.39, (2.76, 6.97)</td>
<td>&lt;0.0001</td>
<td></td>
</tr>
<tr>
<td>Sulfasalazine</td>
<td>0.43, (0.23, 0.80)</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>Mesalazine</td>
<td>0.87, (0.51, 1.50)</td>
<td>0.621</td>
<td></td>
</tr>
<tr>
<td>Exclusive Enteral Nutrition (Tolerex® or Modulen IBD®)</td>
<td>1.19, (0.31, 4.50)</td>
<td>0.798</td>
<td></td>
</tr>
</tbody>
</table>

*Abnormal liver enzymes: 88/300 patients
**Excluding PSC/ASC/AIH


Medications, specifically corticosteroids and antibiotics, and CD phenotype continued to display the strongest positive associations with liver abnormalities when tested against higher thresholds for LEs results, as well as persistence of the abnormalities (Table 3.6, Table 3.7, Table 3.8, Table 3.9, Table 3.10, and Table 3.11). Biologics were also associated with the development of persistent low-grade abnormal LEs (HR 2.19 [95% CI 1.03, 4.66] p = 0.042, Table 3.8). There were no statistically significant associations between the development of higher thresholds, or persistent, abnormal LEs and other medications, such as methotrexate or azathioprine.
Table 3.6: Univariate Cox Proportional Hazards models for the development of abnormal liver enzymes at >2x the upper limit of normal in patients with IBD*

<table>
<thead>
<tr>
<th>Variables</th>
<th>Hazard Ratio Estimates (95% Confidence Interval)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBD Diagnosis: CD versus UC</td>
<td>0.84, (0.41, 1.74)</td>
<td>0.643</td>
</tr>
<tr>
<td>Age</td>
<td>1.09, (0.96, 1.25)</td>
<td>0.188</td>
</tr>
<tr>
<td>Sex (males versus females)</td>
<td>0.82, (0.39, 1.71)</td>
<td>0.592</td>
</tr>
<tr>
<td>PCDAI at diagnosis in CD patients</td>
<td>1.01, (0.98, 1.03)</td>
<td>0.735</td>
</tr>
<tr>
<td>PUCAI at diagnosis in UC patients</td>
<td>1.00, (0.98, 1.03)</td>
<td>0.886</td>
</tr>
<tr>
<td><strong>Extra-intestinal manifestations of IBD</strong></td>
<td><strong>0.13, (0.02, 0.94)</strong></td>
<td><strong>0.043</strong></td>
</tr>
<tr>
<td>History of Autoimmune disease</td>
<td>0.15, (0.02, 1.35)</td>
<td>0.090</td>
</tr>
<tr>
<td>Family history of IBD</td>
<td>0.81, (0.37, 1.80)</td>
<td>0.608</td>
</tr>
<tr>
<td>Family history of autoimmune disease</td>
<td>0.91, (0.41, 2.04)</td>
<td>0.816</td>
</tr>
<tr>
<td>Weight z-score at IBD diagnosis</td>
<td>0.76, (0.55, 1.05)</td>
<td>0.101</td>
</tr>
<tr>
<td>Height z-score at IBD diagnosis</td>
<td>0.98, (0.69, 1.39)</td>
<td>0.909</td>
</tr>
<tr>
<td><strong>BMI z-score at IBD diagnosis</strong></td>
<td><strong>0.73, (0.56, 0.95)</strong></td>
<td><strong>0.017</strong></td>
</tr>
<tr>
<td>Autoimmune antibody positivity (ANCA/ANA/ASMA/ALKM)</td>
<td>0.69, (0.30, 1.59)</td>
<td>0.383</td>
</tr>
<tr>
<td><strong>ANCA positive</strong></td>
<td><strong>1.52, (0.66, 3.49)</strong></td>
<td><strong>0.328</strong></td>
</tr>
<tr>
<td><strong>ASCA positive</strong></td>
<td><strong>2.49, (0.75, 8.22)</strong></td>
<td><strong>0.135</strong></td>
</tr>
<tr>
<td><strong>Medications</strong></td>
<td><strong>Biologics (infliximab/adalimumab/certolizumab/golimumab)</strong></td>
<td><strong>0.61, (0.26, 1.42)</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Infliximab</strong></td>
<td><strong>0.59, (0.23, 1.52)</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Adalimumab</strong></td>
<td><strong>0.81, (0.24, 2.74)</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Methotrexate</strong></td>
<td><strong>0.55, (0.22, 1.39)</strong></td>
</tr>
<tr>
<td></td>
<td><strong>6-Mercaptopurine or Azathioprine</strong></td>
<td><strong>0.76, (0.38, 1.55)</strong></td>
</tr>
<tr>
<td><strong>Corticosteroids (oral or intravenous)</strong></td>
<td><strong>5.85, (3.59, 9.54)</strong></td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td><strong>Metronidazole or Ciprofloxacin</strong></td>
<td><strong>5.48, (3.41, 8.79)</strong></td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td><strong>Sulfasalazine</strong></td>
<td><strong>0.32, (0.13, 0.79)</strong></td>
<td><strong>0.014</strong></td>
</tr>
<tr>
<td><strong>Mesalazine</strong></td>
<td><strong>0.47, (0.21, 1.07)</strong></td>
<td><strong>0.073</strong></td>
</tr>
<tr>
<td><strong>Exclusive Enteral Nutrition (Tolerex® or Modulen IBD®)</strong></td>
<td><strong>2.72, (0.99, 7.48)</strong></td>
<td><strong>0.052</strong></td>
</tr>
</tbody>
</table>

*Abnormal liver enzymes: 58/300 patients
**Excluding PSC/ASC

Table 3.7: Multivariable Cox Proportional Hazards model for the development of abnormal liver enzymes at >2x the upper limit of normal in patients with IBD

<table>
<thead>
<tr>
<th>Variables</th>
<th>Hazard Ratio Estimates (95% Confidence Interval)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood Ratio (Chi Sq [df])</td>
<td>59.1 [7]</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>IBD Diagnosis: CD versus UC</td>
<td>1.73, (0.84, 3.58)</td>
<td>0.141</td>
</tr>
<tr>
<td>Extra-intestinal manifestations of IBD**</td>
<td>0.14, (0.02, 1.02)</td>
<td>0.052</td>
</tr>
<tr>
<td>BMI z-score at IBD diagnosis</td>
<td>0.72, (0.59, 0.88)</td>
<td>0.001</td>
</tr>
<tr>
<td>Medications</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corticosteroids (oral or intravenous)</td>
<td>3.57, (2.18, 5.86)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Metronidazole or Ciprofloxacin</td>
<td>4.56, (2.81, 7.40)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Sulfasalazine</td>
<td>0.27, (0.11, 0.68)</td>
<td>0.005</td>
</tr>
<tr>
<td>Mesalazine</td>
<td>0.36, (0.16, 0.80)</td>
<td>0.013</td>
</tr>
</tbody>
</table>

**Excluding PSC/ASC
IBD: Inflammatory Bowel Disease, CD: Crohn disease, UC: Ulcerative Colitis

Table 3.8: Univariate Cox Proportional Hazards models for the development of persistently, and mildly, abnormal liver enzymes (30-90days, 1-2x ULN) in patients with IBD*

<table>
<thead>
<tr>
<th>Variables</th>
<th>Hazard Ratio Estimates (95% Confidence Interval)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBD Diagnosis: CD versus UC</td>
<td>1.24, (0.64, 2.40)</td>
<td>0.533</td>
</tr>
<tr>
<td>Age</td>
<td>1.12, (1.00, 1.26)</td>
<td>0.062</td>
</tr>
<tr>
<td>Sex (males versus females)</td>
<td>0.79, (0.40, 1.57)</td>
<td>0.508</td>
</tr>
<tr>
<td>PCDAI at diagnosis in CD patients</td>
<td>0.99, (0.96, 1.02)</td>
<td>0.435</td>
</tr>
<tr>
<td>PUCAI at diagnosis in UC patients</td>
<td>0.98, (0.96, 1.00)</td>
<td>0.079</td>
</tr>
<tr>
<td>Extra-intestinal manifestations of IBD**</td>
<td>0.48, (0.11, 2.03)</td>
<td>0.316</td>
</tr>
<tr>
<td>History of Autoimmune disease</td>
<td>0.78, (0.11, 5.68)</td>
<td>0.804</td>
</tr>
<tr>
<td>Family history of IBD</td>
<td>0.65, (0.30, 1.39)</td>
<td>0.267</td>
</tr>
<tr>
<td>Family history of autoimmune disease</td>
<td>1.72, (0.81, 3.64)</td>
<td>0.159</td>
</tr>
<tr>
<td>Weight z-score at IBD diagnosis</td>
<td>0.95, (0.67, 1.33)</td>
<td>0.744</td>
</tr>
<tr>
<td>Height z-score at IBD diagnosis</td>
<td>0.89, (0.63, 1.25)</td>
<td>0.501</td>
</tr>
<tr>
<td>BMI z-score at IBD diagnosis</td>
<td>1.23, (0.95, 1.58)</td>
<td>0.118</td>
</tr>
<tr>
<td>Autoimmune antibody positivity (ANCA/ANA/ASMA/ALKM)</td>
<td>1.55, (0.76, 3.19)</td>
<td>0.229</td>
</tr>
<tr>
<td>ANCA positive</td>
<td>2.12, (1.00, 4.49)</td>
<td>0.050</td>
</tr>
<tr>
<td>ASCA positive</td>
<td>1.07, (0.26, 4.37)</td>
<td>0.927</td>
</tr>
<tr>
<td>Medications</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biologics (infliximab/adalimumab/ certolizumab/golimumab)</td>
<td>1.96, (0.90, 4.25)</td>
<td>0.090</td>
</tr>
<tr>
<td>Infliximab</td>
<td>1.82, (0.79, 4.18)</td>
<td>0.159</td>
</tr>
<tr>
<td>Adalimumab</td>
<td>1.75, (0.56, 5.50)</td>
<td>0.339</td>
</tr>
</tbody>
</table>
Methotrexate 0.88, (0.32, 2.42) 0.803
6-Mercaptopurine or Azathioprine 0.33, (0.07, 1.61) 0.169
**Corticosteroids (oral or intravenous)** 2.98, (1.45, 6.13) 0.003
Metronidazole or Ciprofloxacin 3.06, (1.40, 6.68) 0.005
Sulfasalazine 0.60, (0.26, 1.39) 0.232
Mesalazine 0.36, (0.11, 1.15) 0.083
Exclusive Enteral Nutrition (Tolerex® or Modulen IBD®) 0.80, (0.11, 5.87) 0.825

*Abnormal liver enzymes: 37/300 patients
**Excluding patients with PSC/ASC

Table 3.9: Multivariable Cox Proportional Hazards model for the development of persistently, and mildly, abnormal liver enzymes (30-90 days, 1-2x ULN) in patients with IBD

<table>
<thead>
<tr>
<th>Variables</th>
<th>Hazard Ratio Estimates (95% Confidence Interval)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Likelihood Ratio (Chi Sq [df])</strong></td>
<td>27.7 [6]</td>
<td>0.0001</td>
</tr>
<tr>
<td>IBD Diagnosis: CD versus UC</td>
<td>2.39, (1.15, 4.97)</td>
<td>0.020</td>
</tr>
<tr>
<td>Age</td>
<td>1.12, (0.99, 1.27)</td>
<td>0.068</td>
</tr>
<tr>
<td>Medications</td>
<td>Biologics (infliximab/adalimumab/certolizumab/golimumab)</td>
<td>2.19, (1.03, 4.66)</td>
</tr>
<tr>
<td></td>
<td>Corticosteroids (oral or intravenous)</td>
<td>2.19, (0.99, 4.81)</td>
</tr>
<tr>
<td></td>
<td>Metronidazole or Ciprofloxacin</td>
<td>2.79, (1.11, 7.00)</td>
</tr>
<tr>
<td></td>
<td>Mesalazine</td>
<td>0.34, (0.10, 1.15)</td>
</tr>
</tbody>
</table>

IBD: Inflammatory Bowel Disease, CD: Crohn disease, UC: Ulcerative Colitis

Table 3.10: Univariable Cox Proportional Hazards models for the development of persistently, and severely, abnormal liver enzymes (30-90 days, >2x ULN) in patients with IBD

<table>
<thead>
<tr>
<th>Variables</th>
<th>Hazard Ratio Estimates (95% Confidence Interval)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBD Diagnosis: CD versus UC</td>
<td>2.30, (0.92, 5.75)</td>
<td>0.075</td>
</tr>
<tr>
<td>Age</td>
<td>1.10, (0.91, 1.33)</td>
<td>0.337</td>
</tr>
<tr>
<td>Sex (males versus females)</td>
<td>0.78, (0.30, 2.00)</td>
<td>0.606</td>
</tr>
<tr>
<td>PCDAI at diagnosis in CD patients</td>
<td>1.02, (1.00, 1.05)</td>
<td>0.112</td>
</tr>
<tr>
<td>PUCAI at diagnosis in UC patients</td>
<td>0.96, (0.93, 0.99)</td>
<td>0.022</td>
</tr>
<tr>
<td>Variables</td>
<td>Hazard Ratio Estimates (95% Confidence Interval)</td>
<td>p value</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>--------------------------------------------------</td>
<td>---------</td>
</tr>
<tr>
<td><strong>Likelihood Ratio (Chi Sq [df])</strong></td>
<td>26.7 [4]</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>IBD Diagnosis: CD versus UC</td>
<td>3.42, (1.34, 8.70)</td>
<td>0.010</td>
</tr>
<tr>
<td><strong>Family history of IBD</strong></td>
<td>0.10, (0.01, 0.80)</td>
<td>0.030</td>
</tr>
<tr>
<td>Medications</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steroids (oral or intravenous)</td>
<td>2.23, (0.80, 6.26)</td>
<td>0.127</td>
</tr>
<tr>
<td>Metronidazole or Ciprofloxacin</td>
<td>4.45, (1.44, 13.78)</td>
<td>0.010</td>
</tr>
</tbody>
</table>

*Abnormal liver enzymes: 21/300 patients

**Excluding patients with PSC/ASC


Table 3.11: Multivariable Cox Proportional Hazards model for the development of persistently, and severely, abnormal liver enzymes (30-90 days, >2x ULN) in patients with IBD
Conclusions

In this study, we describe the prevalence and natural history of the development of the first episode of abnormal LEs in a large, carefully characterized, cohort of children with IBD. These data demonstrate that the probability of developing abnormal LEs is common, at 58.1% of children, by 150 months following their diagnosis with IBD. A large proportion presented with these abnormalities within the first month of diagnosis with IBD (16.3%). Higher elevations of LEs (>2xULN) were also quite frequent (31.3%) compared to a previous report in pediatric IBD (14%)\(^3\). However, less than half of all patients with abnormal LEs developed persistence of these findings. Meanwhile, the development of chronic liver diseases in this cohort was a relatively infrequent event (PSC 4.3%, ASC 1.7%) indicating that abnormal LEs in pediatric IBD is usually a benign phenomenon. The most common association with the development of abnormal LEs in the whole cohort was medication use and CD phenotype.

The development of abnormal LEs at both lower and higher thresholds, as well as with persistence of these abnormalities, was associated with the use of corticosteroids and antibiotics. While it is possible that these medications could be hepatotoxic, we hypothesize that these medications were actually acting as surrogate markers for bowel inflammatory activity in the analysis, such that it is the inflammatory activity of IBD that is actually associated with the abnormal LEs. At SickKids, patients who are newly diagnosed with moderate to severe UC are routinely treated with corticosteroids, whereas those with moderate to severe new onset CD in this time period may have received either corticosteroids or EEN. Biologics were specifically associated with low-grade, persistent elevations of LEs, as was also recognized in our recent single centre report of 195 CD patients treated with infliximab for inflammatory luminal CD\(^{108}\). 5-aminosalicylic acid-containing medications had a negative association with the development of abnormal LEs, and are used primarily to maintain remission of ulcerative colitis or to treat mildly to moderately active ulcerative colitis. Patients with milder intestinal disease may be less prone to elevation of LEs.

Multiple theories have linked the diseased intestine with liver inflammation. Alterations in the microbiome have been identified in both IBD and PSC, and have been suggested as an important contributor to the etiology of PSC\(^{109}\). Increased intestinal wall permeability in the setting of inflammation, leading to translocation of bacteria, or the pathogen-associated molecular patterns
(PAMPs), may trigger the liver’s innate immune system and result in hepatic inflammation. Decreased farnesoid X receptor (FXR) activity in active IBD is another hypothesis for the development of liver damage, as there is impaired enterohepatic recirculation of bile salts. Thus it is biologically plausible that children with IBD and active intestinal inflammation can develop transient hepatic abnormalities.

The potential for medication hepatotoxicity in the treatment of IBD is well-known. Specifically, corticosteroids can cause a macrovesicular steatohepatitis, and metronidazole can cause hepatocellular injury; however, both are rarely reported in pediatric IBD, hence our hypothesis that these associations with abnormal LEs are actually surrogates for intestinal inflammation. Methotrexate, however, is often associated with abnormal LEs (1 in 10 children with IBD) and rarely with the development of hepatic fibrosis in the adult rheumatology and IBD literature. In a recent multi-centre retrospective review of methotrexate in children with Crohn disease, LEs were elevated at least once during the first year of use in 89 (39%) of 226. Forty (18% of the entire cohort) required dose reduction or treatment cessation. In our analysis, exclusion of the biochemistry data around the time of initial diagnosis with IBD, confirmed a significantly positive association with the development of abnormal LEs. Similar associations were found with biologic agents. These data highlight the potential hepatotoxic effects of methotrexate and biologics in our cohort.

Previous studies of abnormal LEs in pediatric IBD were limited. They primarily focused on higher thresholds of abnormal LEs; however, recent literature has demonstrated that healthy children, without liver disease, typically have lower aminotransferase levels than previously believed. Thus, those patients with low-grade abnormal LEs, who were not included in the Hyams study, may have had hepatic involvement that required further evaluation. Furthermore, the previous studies did not account for patient censoring, which could have led to bias and an underestimation of the true burden of disease. Most importantly, additional information, such as medication use, was only collected on those patients who had developed high-grade LEs abnormalities. Thus, there was a significant gap in the literature examining associations between clinical variables important in IBD and the development of abnormal LEs.

Our study should be interpreted in the light of potential limitations. Firstly, due to the retrospective nature of this study design, and the lack of LEs screening recommendations in the
literature, LEs were not routinely obtained at specific time intervals. Also, with the detection of abnormal LEs, further investigations, such as assessment for viral hepatitis, varied as well. This could have led to a detection bias, with a skew towards a certain clinical phenotype of IBD or medications. Our analysis was able to account for censoring that arose as adolescents graduated to adult care at different intervals from the time of diagnosis; however, if the frequency of enzymes procurement was related to the patient’s status, there is potential for some remaining bias. Secondly, Sickkids is a major tertiary referral center for pediatric IBD. Our study is still likely generalizable as the rates of PSC/ASC in this cohort were similar to those reported in other studies of pediatric IBD. Finally, patients newly diagnosed with IBD may have presented after variable intervals of time from the initial onset of inflammation. Although the date of positive IBD diagnostic investigations may have been an artificial starting point of disease for the analysis, it was not arbitrary, such that the monitoring of liver impairment, and subsequent investigations, only occur once the diagnosis of IBD is established. Thus, this study provides useful information that can be utilized in a clinical setting about the natural history of the time to development of abnormal LEs observed in children following a diagnosis of IBD.

This study systematically characterizes the frequency and pattern of the first development of abnormal LEs in a carefully phenotyped cohort of children with IBD. Elevated LEs can be expected to occur following a diagnosis of IBD in greater than half of children, typically early (in the first few months) following the initial diagnosis. This phenomenon is usually not associated with the development of chronic liver disease and therefore rarely requires invasive investigations. Abnormal LEs in pediatric IBD is most strikingly associated with medication use, which may reflect the effect of uncontrolled inflammation on the liver in IBD or medication hepatotoxicity. These data are a first essential step in the development of a diagnostic algorithm for abnormal LEs in pediatric IBD. We recommend that statistical modeling methods be employed to determine the specific patterns of LEs elevations in children with IBD to identify associations with specific clinical outcomes (e.g. benign elevations versus chronic liver disease) in future studies.
Chapter 4
Discussion

Summary of Conclusions

The development of abnormal LEs in IBD was a common occurrence, at a probability of 58.1% following a median duration of follow-up of 2.8 years in this cohort. More specifically, a large proportion of children developed the abnormalities early following the diagnosis with IBD (16.3% by 1 month, and 29.5% by 1 year, post-IBD diagnosis). Chronic liver diseases, on the other hand, were relatively infrequent in comparison (4.3% of patients with PSC, 1.7% with ASC). CD phenotype and medications (corticosteroids, antibiotics, and exclusive enteral nutrition) were strongly associated with the development of abnormal LEs. While certain medications, such as methotrexate, have been reported to commonly cause LEs abnormalities, corticosteroids and antibiotics are not typically considered hepatotoxic. These medications are possibly serving as indirect markers of uncontrolled IBD as they are commonly employed in inducing remission. The majority of first episodes of elevated LEs were likely related to uncontrolled IBD, and not to significant underlying hepatic illness, rendering invasive investigations unnecessary in most instances of the first episode of abnormal LEs.

Limitations and Future Directions

A potential limitation to this study included the possibility of a detection bias. Due to the retrospective nature of the study methods, LEs were not routinely obtained. There is a paucity of evidence in the literature to recommend the optimal frequency of LEs testing in this population. Thus, LEs were obtained at irregular intervals, and in the event of an abnormal result, testing was repeated at variable frequencies. Thus, it was possible that patients with well-controlled IBD may have had less LEs measurements, while those with active intestinal disease had investigations more often. Furthermore, there was a high likelihood that patients on well-known hepatotoxic medications also had more frequent LEs monitoring. Overall, there was a possibility for a bias towards the detection of abnormal LEs in patients with uncontrolled IBD or on hepatotoxic medications, purely due to an increased frequency of testing.
Another type of possible information bias was the detection of chronic liver disease in the study cohort. Apart from those patients with PSC or ASC, no other children had liver biopsies. Thus, patients with chronic liver diseases could potentially have been missed. However, these diseases typically cause significantly elevated LEs that persist for many months, which would have naturally prompted further investigations. So only atypical presentations with elevations at lower thresholds in LEs would have been missed, which was likely quite rare.

The risk of a selection bias was likely quite low as the study cohort ascertainment was extremely rigorous. The IBD clinical database was verified for completeness by cross-referencing ICD-9 and ICD-10 codes of the hospital medical records. Furthermore, we suspected that only a minority of children with IBD from the greater Toronto area were diagnosed outside of the Hospital for Sick Children as there were few community pediatric gastroenterologists. Additionally, all patients referred to our tertiary care centre with severe, uncontrolled intestinal disease were excluded, which significantly reduced the risk of a referral bias in the study cohort.

This study was designed to investigate the first episode of abnormal LEs in children with IBD. A high proportion of children with IBD were found to have developed these abnormalities. More importantly, this study determined that a significant percentage developed the abnormal LEs early in the IBD disease course. This information is helpful in the management of children with IBD, as only a small minority were diagnosed with a chronic liver disease. Thus, in the majority of cases, further invasive investigations are unnecessary following the development of abnormal LEs. However, this study was not designed to investigate the patterns of abnormal LEs. Specifically, some patients may develop spikes of abnormality that resolve quickly, and may, or may not, recur. On the other hand, some patients may develop episodes of elevated LEs that may persist, at either low-grade or high-grade thresholds. This study attempted to assess the prevalence of persistence of abnormal LEs, and determine the association of IBD-specific clinical variables. However, only the first episode of persistent LEs was identified and not the overall pattern of abnormalities longitudinally. A diagnostic algorithm for the approach to abnormal LEs in pediatric IBD was presented in chapter 2 (Figure 2.4). This study was not designed to provide evidence of its validity; however, it did identify a significant number of patients with abnormal LEs who were not diagnosed with significant liver disease. This confirmed that only a small subset of children with IBD and abnormal LEs require further investigations.
Further research is required to model the trajectories of LEs over the course of IBD in childhood, to determine if discrete patterns of LEs can be identified and associated with specific liver-related outcomes. By modeling the data in this fashion it will be possible to describe the degree of abnormality, the pattern of the abnormalities, as well as the duration of time that the LEs remain abnormal. Further analyses of these patterns of LEs trajectories could also identify associations with IBD phenotypic clinical variables as well as liver-specific outcomes, such as: PSC, AIH, ASC, hepatotoxicity from specific IBD medications (methotrexate, mercaptopurines, biologic agents, corticosteroids), uncontrolled IBD, intercurrent infections, cholelithiasis, non-alcoholic fatty liver disease (NAFLD), etc.

The goal for these analyses would be to validate the diagnostic algorithm presented in Figure 2.4, and determine the ideal interval for the reassessment of abnormal LEs in children with IBD. Furthermore, it would be possible to identify a threshold and pattern of LEs elevation that may select, with high sensitivity, the subset of individuals that require more invasive investigations in the assessment of acute or chronic liver diseases. Finally, this may also identify a pattern of abnormal LEs that will predict a benign course, without liver disease, and without the need for invasive investigations.

**Identification of Liver-Specific Outcomes**

The SickKids IBD patient database contains information about extra-intestinal manifestations of IBD. For this study, charts were also reviewed by the investigators to determine if any other liver-specific outcomes had developed, but were not recorded, at the time of abnormal LEs. Apart from those patients with PSC/ASC, no patients in this cohort received a liver biopsy. Furthermore, liver ultrasonography and laboratory testing, typically performed in children to investigate the differential diagnoses of elevated LEs, were only obtained in a subset of patients with abnormalities. While it would be unlikely for a serious chronic liver disease to remain undetected, this could be possible without a complete work-up. Standardized investigations should be well-defined in future prospective studies of LEs in IBD, including routine LEs (at least annually) as well as more extensive testing with any abnormalities. These should include liver ultrasonography, with doppler assessment of hepatic vessels, as well as a minimum panel of
bloodwork (e.g. ALT, AST, GGT, fractionated bilirubin, viral hepatic serologies, ceruloplasmin, alpha-1-anti-trypsin, celiac serologies, and thyroid hormones).

Summary of Future Directions

We have identified a critical need for investigation into the pattern and natural history of elevated LEs in children with IBD. This clinical problem requires an evidence-based diagnostic algorithm, and to date, few data exist to support a standardized approach. This study was essential as a stepping-stone to future prospective work that will lead to the development and validation of a diagnostic algorithm that will standardize care in the management of children with IBD and elevated LEs.
References or Bibliography


Appendices

Appendix 1: Additional Methodology – Statistical Analysis

All data manipulation and statistical analysis was performed with SAS/STAT® software, version 9.3 (SAS Institute, Cary, North Carolina, USA).

Dataset Manipulation – Liver enzymes data

Liver biochemistry data were obtained from the SickKids data warehouse on all patients (>1000) who met criteria for inclusion into the study. All biochemistry data was received in one column. The specific biochemistry data that were analyzed and extracted included: ALT and GGT. The results and upper limit of normal (ULN) reference range were verified for data integrity such that character strings were converted to numerical entries. Missing ULN values were corrected to default values: 40U/L for ALT, 45U/L for GGT in children < 15 years, 55U/L for GGT in females ≥ 15 years, and 75U/L for GGT in males ≥ 15 years. These values were chosen after reviewing the ULN ranges for a large portion of data over the course of the decade used in the analysis (Appendix Table 1).

Manual data entry was performed to record the LEs obtained at laboratories outside of SickKids. This data was stored in the RedCap database (Research Electronic Data Capture) and similarly verified for data integrity. The LEs datasets were then merged (Appendix Table 1).

The LEs data were merged with the IBD diagnosis date and IBD phenotype data for the cohort of 300 patients included in the study. Using SAS proc sql (left join) code, this deleted all the LEs results for patients not included in the study. All LEs obtained prior to 90 days before the IBD diagnosis date was also deleted. However, the biochemistry data from the 90 days preceding the IBD diagnosis were included to capture the findings at the first clinic visit or emergency department visit. This was important as patients do not routinely have laboratory investigations repeated on the date of endoscopy. Collection of laboratory data ended on August 21, 2012; no data obtained past this date were included (Appendix Table 1).

A dataset was created with the last LEs measurement for each patient. The date of collection (“Date_Collection”) of these biochemistries was considered the last date of follow-up for each patient. After this date, all patients who did not reach the primary endpoint in the study were
censored. If the biochemistry data were obtained in the 90 days preceding the diagnosis with IBD, then the “Date_Collection” was reset to the date of IBD diagnosis (Appendix Table 1).

A dummy variable, “Abn,” was created to describe if a patient had ever developed an abnormal LEs (ALT or GGT > ULN for age). The “Date_Collection” of this abnormal result became the date of the primary endpoint for the main Kaplan-Meier analysis. This dataset was merged with the previous dataset containing the censoring information using the “update” SAS command which allowed the LEs result, as well as the date of collection, of the first abnormal result to replace the data for the last biochemistry result (the results and date of collection of the censoring date). Demographic data for the three patients who had never had any LEs performed were added into this dataset. The variable "abn_censor_duration" was created to determine the duration of time either to the primary endpoint or to the date of censoring (Appendix Table 1).

A second dataset was created with the same SAS code except for the dummy variable “Abn_2ULN”, which was set to describe biochemistry results that were ≥ 2xULN.

The persistence of abnormal LEs were evaluated by detecting a minimum of two abnormal LEs results spaced by a minimum interval of 30 days and a maximum interval of 90 days. One dataset was created with a dummy variable set to detect abnormalities between 1-2x ULN, while a second dataset was set to detect abnormalities >2x ULN. The SAS “retain” statement was used to identify two episodes of abnormality according to these criteria (Appendix Table 1).

**Dataset Manipulation – Clinical IBD data**

Clinical IBD data was extracted from the IBD database managed in Microsoft Office Access 2007® (Microsoft Corporation, Redmond, WA, USA). The data for intestinal disease location (upper GI, esophagus, stomach, etc.) at IBD diagnosis were converted from one column to many for further analysis (including determination of Paris classification for CD and UC location phenotype). Disease severity scores (PCDAI, mathPCDAI, PUCAI) were calculated from disease symptoms at the time of IBD diagnosis. Data on extraintestinal manifestations of IBD in the index patient, or their family members, were converted from one column to many for further analysis in the Cox models (Appendix Table 2: Dataset Manipulation – Clinical IBD data).
Growth data, including height and weight, were further analyzed to determine height percentile, weight percentile and body mass index percentile. This was performed using SAS programs (gc-setup-BIV and gc-calculate-BIV) offered by the Center for Disease Control and Prevention website (Appendix Table 2: Dataset Manipulation – Clinical IBD data).

Other laboratory investigations specific to autoimmunity, including ANCA and ASCA, were assessed to determine if patients were ever positive for these tests. Thresholds for positivity were set according to the laboratory norms for age. Meanwhile, CRP results were obtained; however, they were not included in the Cox models as they were obtained at variable intervals for each patient (Appendix Table 2: Dataset Manipulation – Clinical IBD data).

**Dataset Manipulation – Medication data treated as Time-Varying Covariates**

Medication data were treated as time-varying covariates in the Cox models. As such, the data were manipulated in order to develop a dataset where each patient had multiple rows of data with start and stop date intervals recorded. In this dataset, a new row was added each time there was a change to the medications (started or stopped). Thus, some medications, taken long-term, were taken over multiple sequential intervals, while other medications, taken short-term, were potentially only present in one interval. In each row, dummy variables were created to describe whether the medications were being taken during that specific time interval. Thus, in the Cox models, medications taken at time “t” (“t” = duration of time between the date of IBD diagnosis and the date of the development of abnormal LEs), were compared against medications not taken at time “t” to determine if there was an association with the development of abnormal LEs. Thus, only medications that were actually being taken by the patients at the time of abnormality or censoring were taken into account in the analysis. However, to account for pharmacokinetics, and a possible lingering hepatotoxic effect, 7 days were added to the end dates of all medications taken by patients in the cohort (Appendix Table 3: Dataset Manipulation – Medication data treated as Time-Varying Covariates).
Dataset Manipulation – Liver enzymes data merged with clinical IBD data

The LEs data were merged with the IBD clinical data. Following this, only the laboratory investigations (e.g. ANCA, ASCA) that were obtained prior to the abnormal LEs (or censoring) were included. Finally, the medication data were imported and concatenated to the dataset. The time intervals for the medication data were corrected so that the start date of the first interval coincided with the date of IBD diagnosis, and the end date of the last interval coincided with either the date of abnormal LEs or the date of censoring. All data before or after these dates were deleted. A dummy variable was added to describe if each medication was being taken during the time intervals (Appendix Table 4: Dataset Manipulation – Liver enzymes data merged with clinical IBD data).

In the case of multiple changes to medications in the same day, start dates for two consecutive time intervals were occasionally the exact same, which would have caused computational errors in the Cox modeling. Thus, a time factor was added to the start and end dates so that there was a 12-hour time difference between these intervals (Appendix Table 4: Dataset Manipulation – Liver enzymes data merged with clinical IBD data).

Similar SAS code was used with the datasets containing information for higher elevations of abnormal LEs (>2xULN) and for the persistence of abnormal LEs (30-90d, 1-2xULN or > 2xULN).

Statistical Analysis – Description of cohort

The number of LEs test procurements per patient was calculated to determine the availability of longitudinal assessments. The cohort was described by presenting means or medians of specific IBD clinical features. This included: age, sex, IBD phenotype, IBD disease location, IBD disease severity, EIM history, family history of IBD and autoimmune diseases, growth percentiles, autoimmune antibody positivity, and frequency or medication use (Appendix Table 5: Statistical Analysis – Description of cohort).

For continuous variables, the Shapiro-Wilk test was used to determine if the data was normally distributed since the cohort contained < 2000 cases. For normally distributed data, means with
standard deviations were calculated; while for non-normally distributed data, medians with interquartile ranges were calculated. For categorical variables, percentages were presented.

Note: for these analyses, a dataset with one row per patient was created.

**Statistical Analysis – Description of liver enzymes abnormalities**

The LEs tests examined in these analyses were ALT and GGT. While all tests were treated equally in the time to abnormality analyses, the frequency of abnormality for each test was presented separately as a percentage. This was calculated for any abnormality in LEs as well as 2xULN (the SAS code was similar for both (Appendix Table 6: Statistical Analysis – Description of liver enzymes abnormalities).

**Statistical Analysis – Time to Event analysis**

Due to the nature of the longitudinal data, Kaplan-Meier analyses were performed to determine the time to the development of abnormal LEs in the pediatric IBD cohort. Specific endpoints included any abnormality in LEs, >2xULN, and persistence of abnormalities (between 30-90 days) at either 1-2xULN or >2xULN (Appendix Table 7: Statistical Analysis – Time to Event analysis).

The dataset was organized such that patients had multiple rows of data to account for changes in medications over time. The last row of data had the end date of when patients either developed the outcome of interest (dummy variable “Abn”=1) or were censored (dummy variable “Abn”=0). In all previous time intervals the dummy variable “Abn” was set to 50 to be treated as censored in the Cox proportional hazard modeling, so these rows were deleted for the Kaplan-Meier analysis. The “Event time” for the Kaplan-Meier analysis was calculated from the date of IBD diagnosis and the end date (either outcome of interest or censoring date).

The cohort and LB data met all key assumptions for analysis with the Kaplan-Meier method. First, the date of IBD diagnosis provided a well-defined starting point for the analysis. Second, the first episode of abnormal LB, and the last available LB procurement, served as a clear
endpoint and censor date respectively. Third, patients were only censored due to transfer to an adult IBD clinic at 18 years of age, or due to change of address. Patients who developed important liver-related outcomes remained at our institution as it is a tertiary care centre. Thus, patients were not lost to follow-up due to factors related to the LB results. Finally, the chances of identifying LB abnormalities in this cohort were likely relatively high and uniform throughout the study period. Guidelines for the monitoring of LB in children with IBD were not available throughout the course of the time period assessed in this study. Although LB procurement practices may have varied depending on the individual IBD specialist’s preferences, the vast majority of patients had at least one LB measurement and >80% of patients had five measurements following their IBD diagnosis.

**Statistical Analysis – Cox Proportional Hazards modeling**

Cox Proportional Hazards models (univariate and multivariable) were constructed to test for associations between specific IBD clinical variables and the development of abnormal LEs. As described for the time to event analysis, the dummy variable “Abn” was set to “1” when the outcome of interest occurred, set to “0” for censoring, and set to “50” for all intervals of time when there were medication changes prior to the last time interval. The SAS code is presented here, and it was similar for the testing of any abnormality in LEs, >2xULN, and persistence of abnormalities (between 30-90 days) at either 1-2xULN or >2xULN. A univariate analysis was also performed on a dataset that excluded the LEs from the first 90 days post IBD diagnosis (Appendix Table 8: Statistical Analysis – Cox Proportional Hazards modeling).

The multivariable models were verified for key assumptions. First, the models were assessed for over-specification by comparing the number of independent variables to the number of events and assuring that the ration was less than 1:10. Second, the models were verified for multicollinearity by assessing the variance inflation factors (if VIF>10 then collinearity likely), the tolerance (if tolerance <0.25 then collinearity likely), and the condition indices (if CI>10 then collinearity possible). After already accounting for the time-varying factor of the medications, no violations to the assumptions were identified in any of the models (Appendix Table 8: Statistical Analysis – Cox Proportional Hazards modeling).
Appendix 2: Additional Methodology – SAS/STAT® Software Code

Appendix Table 1: Dataset Manipulation – Liver enzymes data

/*Fix ALT; ULN @ 40 if missing*/
data ALT;
   set DataBiox;
   if test = 'ALT,blood';
      if Result = 'Not Available' then delete;
      if Result = '<2' then Result = 2;
      if Result = '<3' then Result = 3;
      if Result = '<6' then Result = 6;
      if Result = '<10' then Result = 10;
      if Result = 'Not sufficient quantity.' then delete;
      if Result = 'ERROR.' then delete;
      if Result = 'Canceled: Incorrect Entry of Collect Date or Time.' then delete;
      if Result = 'moderately hemolysed' then delete;
      if Result = 'slightly hemolysed' then delete;
   /* Result = Result + 0; /*to ensure the variable is numerical */
   if ULN= then ULN=40.0;
   ULN = ULN + 0; /*to ensure the variable is numerical */
run;
/*Need to add Sex in order to specify ULN for GGT*/
proc sql;
   create table DataBiox1 as
      select * from DataBiox as MRN left join DataDOB as MRN
      on DataDOB.MRN eq DataBiox.MRN;
quit;
/*Fix GGT; ULN @ 45 for <15yrs, ULN @ 55 for girls, @75 for boys if missing */
data GGT_M;
   set DataBiox1;
   if test = 'GGT,blood';
   if Sex = 1;
      if Result = 'Not Available' then delete;
      if Result = '<2' then Result = 2;
      if Result = '<3' then Result = 3;
      if Result = '<6' then Result = 6;
      if Result = '<10' then Result = 10;
      if Result = 'Not sufficient quantity.' then delete;
      if Result = 'ERROR.' then delete;
      if Result = 'Canceled: Incorrect Entry of Collect Date or Time.' then delete;
      if Result = 'moderately hemolysed' then delete;
      if Result = 'slightly hemolysed' then delete;
   /* Result = Result + 0; /*to ensure the variable is numerical */
   if ULN=. AND (Date_Collected-DOB)>=5475 then ULN=75.0;
   /*ULN for 15 years old and up*/
   if ULN=. AND (Date_Collected-DOB)<5475 then ULN=45.0;
   /*ULN for <15 years old*/
   ULN = ULN + 0; /*to ensure the variable is numerical */
run;
data GGT_F;
   set DataBiox1;
   if test = 'GGT,blood';
if Sex = 2;
if Result = 'Not Available' then delete;
   if Result = '<2' then Result = 2;
   if Result = '<3' then Result = 3;
   if Result = '<6' then Result = 6;
   if Result = '<10' then Result = 10;
   if Result = 'Not sufficient quantity.' then delete;
   if Result = 'ERROR.' then delete;
   if Result = 'Canceled: Incorrect Entry of Collect Date or Time.' then delete;
   if Result = 'moderately hemolysed' then delete;
   if Result = 'slightly hemolysed' then delete;
/* Result = Result + 0; /*to ensure the variable is numerical */
if ULN=. AND (Date_Collected-DOB)>=5475 then ULN=55.0;
/*ULN for 15 years old and up*/
if ULN=. AND (Date_Collected-DOB)<5475 then ULN=45.0;
/*ULN for <15 years old*/
ULN = ULN + 0; /*to ensure the variable is numerical*/
run;
/*FIX RedCap Liver Data*/
proc sort data=DataLIV out=DataLIV1;
   by MRN Date_Collected;
run;
proc sort data=dataDOB out=dataDOB1;
   by MRN;
run;
data DataLIV2;
   merge DataLIV1 dataDOB1;
   by MRN;
run;
data DataLIV3;
   set DataLIV2;
   if Test="Bilirubin-Conjugated,blood" then delete;
   if test="ALT,blood" AND ULN=. then ULN=40.0;
   if test="AST,blood" AND ULN=. then delete;
   if test="GGT,blood" AND ULN=. AND (Date_Collected-DOB)>=5475 then
      ULN=55.0; /*ULN for 15 years old and up*/
   if test="GGT,blood" AND ULN=. AND (Date_Collected-DOB)<5475 then
      ULN=45.0; /*ULN for <15 years old*/
   ULN = ULN + 0; /*to ensure the variable is numerical*/
run;
data DataLIV4;
   set DataLIV3;
   keep MRN Date_Collected Test NumResult ULN;
run;
/*MERGE OF ALT/GGT DATA SETS*/
/*Concatenate the liver biochem from the kidcare dataset, and ensure all results are numerical="NumResult" */
data KidcareMerged;
   set ALT GGT_M GGT_F;
   NumResult = result * 1;
run;
/*Concatenate the datasets: kidcare dataset, with the RedCap dataset with liver biochem only */
data Merged1;
   set KidcareMerged DataLIV4;
run;
/**MERGE the biochem data with the IBD diagnosis data*/
/*Cody Program 26-3 Using PROC SQL to create a SAS data set*/
/*Cody Program 26-2 Using an asterisk to select all the variables in a data set*/
/*Cody Program 26-8 Demonstrating a left join for merging. This will allow us to keep only the data for the included study cohort (N=300) */

```sql
proc sql;
  create table Merged2 as
    select *
    from DataDx as StudyID_Info left join
      Merged1 as MRN
    on Merged1.MRN eq DataDx.StudyID_Info;
quit;

data Merged3;
  set Merged2;
  DROP MRN DOB Sex;
run;

proc sql;
  create table Merged4 as
    select *
    from Merged3 as StudyID_Info left join
      DataDOB as MRN
    on DataDOB.MRN eq Merged3.StudyID_Info;
quit;

/* Remove biochemistry measurements that preceded a diagnosis of IBD
(INCLUDE the 3 months preceding the IBD diagnosis as the clinic assessment
can be considered close enough to the "time of diagnosis") */

data Merged5;
  set Merged4;
  if DateDx > (Date_Collected+90) then delete;
  DROP MRN Result;
run;

/*Fix Result & StudyID variable name*/

data Merged6;
  set Merged5;
  Result = NumResult;
  DROP NumResult;
  StudyID = StudyID_info;
  DROP StudyID_info;
  orig_coll_date=Date_Collected;
  format orig_coll_date date9.;
run;

/* Sort the liver biochem data by patient and then by of lab test;
NOTE: KIDCARE database of labs has data until August 21, 2012*** */
proc sort data=Merged6 out=All_Liv_Biochem;
  by StudyID Date_Collected;
  where Date_Collected <= '21aug2012'd;
run;

/*ADD A VARIABLE FOR THE TOTAL DURATION OF FOLLOW-UP (based on last liver biochem measurement)*/
/* create a dataset with last liver biochem measurement */
data Last_liver_test1;
  set All_Liv_Biochem;
  by StudyID;
  if last.StudyID=0 then delete; /*the last biochem measurement should be last.StudyID=1*/
run;
/*calculate the duration of follow-up from the time of diagnosis - last date of follow-up=last biochem measurement*/
data Last_liver_test;
   set Last_liver_test1;
   if orig_coll_date < DateDx then Date_Collected = DateDx; /*if the biochem data was only obtained around the time of diagnosis, then F/U time = 0*/
   if Date_Collected=. then Date_Collected = DateDx;
   duration_of_FU = Date_Collected - DateDx; /*output is in days*/
run;
data All_Liv_Biochem_totFU;
   set All_Liv_Biochem;
run;
data All_Liv_Biochem_totFU;
   merge All_Liv_Biochem_totFU Last_liver_test;
   by StudyID;
run;
 לקוחות ל diss

/*********************/
/*add dummy variable for 1st abnormal biochemistry. Abn=0 (normal), Abn=1 (abnormal)*/
data Abn_LivBiochem;
   set All_Liv_Biochem_totFU;
   Abn = 0;
   if Result > ULN then Abn = 1;
run;
/*TIME TO EVENT DATASET*/
/* data set with only abnormal results and then remove all normal results*/
data Abn_LivBiochem_only;
   set Abn_LivBiochem;
   if Abn=1;
run;
/* create a dataset with time to ABNORMAL event*/
/*ensure the dataset is sorted by StudyID and date collected*/
proc sort data=Abn_LivBiochem_only out=Abn_LivBiochem_only_sorted;
   by StudyID Date_Collected;
run;
data First_Abn_liver_test;
   set Abn_LivBiochem_only_sorted;
   by StudyID;
   if first.StudyID=1;
run;
/*Add a variable "abn_censor_duration" that represents the duration of time to the first abn biochem result or total duration of follow-up if patient was censored*/
data First_Abn_liver_test1;
   set First_Abn_liver_test;
   abn_censor_duration=(Date_Collected-DateDx);
   if orig_coll_date < DateDx then abn_censor_duration = 0;
run;
/*DATASET = 'time_to_1st_event': DATASET WITH FIRST ABNORMAL TEST *AND* LAST CENSORED TEST FOR THOSE WITH ONLY NORMAL TESTS*/
data time_to_1st_event;
   set Last_liver_test;
   abn_censor_duration=duration_of_FU;
run;
data time_to_1st_event;
update time_to_lst_event First_Abn_liver_test1;
by StudyID;
run;

/*CREATE A DATABASE "all_ptn_time_to_event" WITH ALL IBD PATIENTS INCLUDING THOSE WITHOUT LIVER BIOCHEM DATA*/
data all_ptn_time_to_event1;
set DataDx;
StudyID = StudyID_info;
DROP StudyID_info;
run;
proc sort data=all_ptn_time_to_event1 out=all_ptn_time_to_event1;
by StudyID;
run;
data all_ptn_time_to_event1;
update all_ptn_time_to_event1 time_to_1st_event;
by StudyID;
run;

/*Ensure the 3 patients with no liver biochem have DOB & Sex data*/
data DataDOBTEMP;
set DataDOB;
StudyID = MRN;
DROP MRN;
run;
proc sort data=DataDOBTEMP out=DataDOBTEMP;
by StudyID;
run;
data all_ptn_time_to_event1;
update all_ptn_time_to_event1 DataDOBTEMP;
by StudyID;
run;
data all_ptn_time_to_event;
set all_ptn_time_to_event1;
if DateDx = . then delete;
run;

/*CREATE A DATABASE WITH ALL IBD PATIENTS INCLUDING THOSE WITHOUT LIVER BIOCHEM DATA ***AND*** Follow-up duration*/
data Only_FU_time;
set All_Liv_Biochem_totFU;
by StudyID;
if first.StudyID=1;
keep StudyID duration_of_FU;
run;
data all_ptn_time_to_event_FU;
merge all_ptn_time_to_event Only_FU_time;
by StudyID;
if duration_of_FU = . then duration_of_FU = 0;
if abn_censor_duration = . then abn_censor_duration = 0;
if Abn = . then Abn = 0;
AgeDxYrs = (DateDx - DOB)/365; /*age in years*/
run;
proc print data=all_ptn_time_to_event_FU (obs=50); Title "all_ptn_time_to_event_FU: dataset with all patients in DataInfo and time to event AND FU";
r
/*----------------------------------------------
**adding dummy variable for persistently abnormal biochemistry over ***30-90 DAYS***. Abn_pers90=0 (normal), Abn_pers90=1 (abnormal)*/
data Abn_LivBiochem_pers1;
    set All_Liv_Biochem_totFU;
    Abn = 0;
    Abn_pers90 = 0;
    if (Result > ULN AND Result <= (2*ULN)) then Abn = 1;
run;

data Abn_LivBiochem_pers2;
    set Abn_LivBiochem_pers1;
    if Abn =1;
run;
/*Keep only biochem measurement per day - Ref: the little SAS book pg 108*/
proc sort data=Abn_LivBiochem_pers2 out=Abn_LivBiochem_pers3 NODUPKEY;
    by StudyID Date_Collected;
run;
/*Using the retain statement (Cody, Learning SAS by Example: A Programmer’s Guide, pg 517, Ch 24.8)*/
/*This code will detect all episodes of persistence!!!*/
data Abn_LivBiochem_pers;
    set Abn_LivBiochem_pers3;
    by StudyID;
    retain Abn_pers90 First_Abn_date;
    if first.StudyID=1 then First_Abn_date = Date_Collected; /*Set the date for the first abnormal result*/
    format First_Abn_date date9.;
    if Abn_pers90 = 0 then do;
        if abn=1 AND Date_Collected NE First_Abn_date AND Date_Collected-First_Abn_date >= 30 AND Date_Collected-First_Abn_date <= 90 then do; /*Assessing the next biochem measurement for persistence of abnormal liver biochem for 30-90 DAYS*/
            Abn_pers90=1;
            Persis_Abn_date=Date_Collected;
            format Persis_Abn_date date9.;
            output;
        end;
    else do;
        First_Abn_date = Date_Collected;
        format First_Abn_date date9.;
        /*if first.StudyID=1 then delete;*/
        output;
    end;
run;
/*TIME TO EVENT DATASET*/
/*Data set with only persistently abnormal results and then remove all normal results*/
data Abn_LivBiochem_only_pers;
    set Abn_LivBiochem_pers;
    if Abn_pers90=1;
run;
/*Create a dataset with time to persistently ABNORMAL event*/
/*Ensure the dataset is sorted by StudyID and date collected*/
proc sort data=Abn_LivBiochem_only_pers out=Abn_LivBiochem_only_sorted_pers;
    by StudyID Date_Collected;
run;
data First_Abn_liver_test_pers;
    set Abn_LivBiochem_only_sorted_pers;
    by StudyID;
    if first.StudyID=1;
run;
/* Add a variable "abn_censor_duration" that represents the duration of time to the first abn biochem result or total duration of follow-up if patient was censored*/
data First_Abn_liver_test_pers_1;
   set First_Abn_liver_test_pers;
   abn_censor_duration_pers=(First_Abn_date-DateDx);
   if First_Abn_date < DateDx then abn_censor_duration_pers = 0;
run;
/* DATASET = 'time_to_1st_event_pers': DATASET WITH FIRST PERSISTENTLY ABNORMAL TEST *AND* LAST CENSORED TEST FOR THOSE WITH ONLY NORMAL TESTS*/
data time_to_1st_event_pers;
   set Last_liver_test;
   abn_censor_duration_pers=duration_of_FU;
run;
data time_to_1st_event_pers;
   update time_to_1st_event_pers First_Abn_liver_test_pers_1;
   by StudyID;
run;
/*CREATE A DATABASE "all_ptn_time_to_event_pers" WITH ALL IBD PATIENTS INCLUDING THOSE WITHOUT LIVER BIOCHEM DATA*/
data all_ptn_time_to_event1_pers;
   set DataDx;
   StudyID = StudyID_info;
   DROP StudyID_info;
run;
proc sort data=all_ptn_time_to_event1_pers out=all_ptn_time_to_event1_pers;
by StudyID;
run;
data all_ptn_time_to_event1_pers;
   update all_ptn_time_to_event1_pers time_to_1st_event_pers;
   by StudyID;
run;
/*Ensure the 3 patients with no liver biochem have DOB & Sex data*/
data DataDOBTEMP;
   set DataDOB;
   StudyID = MRN;
   DROP MRN;
run;
proc sort data=DataDOBTEMP out=DataDOBTEMP;
by StudyID;
run;
data all_ptn_time_to_event1_pers;
   update all_ptn_time_to_event1_pers DataDOBTEMP;
   by StudyID;
run;
data all_ptn_time_to_event1_pers;
   set all_ptn_time_to_event1_pers;
   if DateDx = . then delete;
run;
/*CREATE A DATABASE WITH ALL IBD PATIENTS INCLUDING THOSE WITHOUT LIVER BIOCHEM DATA ***AND*** Follow-up duration*/
data Only_FU_time;
   set All_Liv_Biochem_totFU;
   by StudyID;
   if first.StudyID=1;
   keep StudyID duration_of_FU;
run;
data all_ptn_time_to_eventFU_pers;
    merge all_ptn_time_to_event_pers Only_FU_time;
    by StudyID;
    if duration_of_FU = . then duration_of_FU = 0;
    if abn_censor_duration_pers = . then abn_censor_duration_pers = 0;
    if Abn_pers90 = . then Abn_pers90 = 0;
    AgeDxYrs = (DateDx - DOB) / 365; /*age in years*/
run;

proc print data=all_ptn_time_to_eventFU_pers (obs=10); Title
"all_ptn_time_to_eventFU_pers: dataset with all patients in DataInfo and time to event for persistent abn biochem 30-90 days AND FU";
run;

Appendix Table 2: Dataset Manipulation – Clinical IBD data

data IBD_Data1;
    set DataIBD;
run;
/*Convert IBD site & involvement from 2 columns into many - according to site involved with IBD*/
data IBD_Data2;
    set IBD_Data1;
    if site = "Upper GI" AND Involvement = "Macrosopic Disease" then UpperGI = "macro";
    if site = "Upper GI" AND Involvement = "Microscopic Only" then UpperGI = "micro";
    if site = "Upper GI" AND Involvement = "Not Involved" then UpperGI = "none";
    if site = "Upper GI" AND Involvement = "Not assessed" then UpperGI = "NA";
    if site = "Oral" AND Involvement = "Macrosopic Disease" then Oral = "macro";
    if site = "Oral" AND Involvement = "Microscopic Only" then Oral = "micro";
    if site = "Oral" AND Involvement = "Not Involved" then Oral = "none";
    if site = "Oral" AND Involvement = "Not assessed" then Oral = "NA";
    if site = "Esophagus" AND Involvement = "Macrosopic Disease" then Esophagus = "macro";
    if site = "Esophagus" AND Involvement = "Microscopic Only" then Esophagus = "micro";
    if site = "Esophagus" AND Involvement = "Not Involved" then Esophagus = "none";
    if site = "Esophagus" AND Involvement = "Not assessed" then Esophagus = "NA";
    if site = "Stomach" AND Involvement = "Macrosopic Disease" then Stomach = "macro";
    if site = "Stomach" AND Involvement = "Microscopic Only" then Stomach = "micro";
    if site = "Stomach" AND Involvement = "Not Involved" then Stomach = "none";
    if site = "Stomach" AND Involvement = "Not assessed" then Stomach = "NA";
    if site = "Duodenum" AND Involvement = "Macrosopic Disease" then Duodenum = "macro";
if site = "Duodenum" AND Involvement = "Microscopic Only" then Duodenum = "micro";
if site = "Duodenum" AND Involvement = "Not Involved" then Duodenum = "none";
if site = "Duodenum" AND Involvement = "Not assessed" then Duodenum = "NA";

if site = "Jejunum/ileum" AND Involvement = "Macroscopic Disease" then JejIleum = "macro";
if site = "Jejunum/ileum" AND Involvement = "Microscopic Only" then JejIleum = "micro";
if site = "Jejunum/ileum" AND Involvement = "Not Involved" then JejIleum = "none";
if site = "Jejunum/ileum" AND Involvement = "Not assessed" then JejIleum = "NA";

if site = "Terminal Ileum" AND Involvement = "Macroscopic Disease" then TI = "macro";
if site = "Terminal Ileum" AND Involvement = "Microscopic Only" then TI = "micro";
if site = "Terminal Ileum" AND Involvement = "Not Involved" then TI = "none";
if site = "Terminal Ileum" AND Involvement = "Not assessed" then TI = "NA";

if site = "Cecum" AND Involvement = "Macroscopic Disease" then Cecum = "macro";
if site = "Cecum" AND Involvement = "Microscopic Only" then Cecum = "micro";
if site = "Cecum" AND Involvement = "Not Involved" then Cecum = "none";
if site = "Cecum" AND Involvement = "Not assessed" then Cecum = "NA";

if site = "Ascending colon" AND Involvement = "Macroscopic Disease" then AscColon = "macro";
if site = "Ascending colon" AND Involvement = "Microscopic Only" then AscColon = "micro";
if site = "Ascending colon" AND Involvement = "Not Involved" then AscColon = "none";
if site = "Ascending colon" AND Involvement = "Not assessed" then AscColon = "NA";

if site = "Transverse Col" AND Involvement = "Macroscopic Disease" then TransCol = "macro";
if site = "Transverse Col" AND Involvement = "Microscopic Only" then TransCol = "micro";
if site = "Transverse Col" AND Involvement = "Not Involved" then TransCol = "none";
if site = "Transverse Col" AND Involvement = "Not assessed" then TransCol = "NA";

if site = "Left Colon" AND Involvement = "Macroscopic Disease" then LtColon = "macro";
if site = "Left Colon" AND Involvement = "Microscopic Only" then LtColon = "micro";
if site = "Left Colon" AND Involvement = "Not Involved" then LtColon = "none";
if site = "Left Colon" AND Involvement = "Not assessed" then LtColon = "NA";
if site = "Proctitis" AND Involvement = "Macroscopic Disease" then Proctitis = "macro";
if site = "Proctitis" AND Involvement = "Microscopic Only" then Proctitis = "micro";
if site = "Proctitis" AND Involvement = "Not Involved" then Proctitis = "none";
if site = "Proctitis" AND Involvement = "Not assessed" then Proctitis = "NA";

DROP site involvement;

if TI="macro" then Rt_side=1;
if Cecum="macro" then Rt_side=1;
if AscColon="macro" then Rt_side=1;

run;

data IBD_Data2a;
set IBD_Data2;
keep StudyID UpperGI Oral Esophagus Stomach Duodenum JejIleum TI Cecum AscColon TransCol LtColon Proctitis Rt_side;
run;

/* List of just the 300 patients in the cohort*/
data IBD_Data2b;
set DataDx;
StudyID = StudyID_Info;
drop DateDx StudyID_Info;
run;

data IBD_Data2b;
update IBD_Data2b IBD_Data2a;
by StudyID;
run;

/********************* NEED TO ADD IBD_Data2b BACK TO IBD_Data1*****************************/
data IBD_Data4;
set IBD_Data1;
drop Site Involvement Medication_Name Med_Startdate Med_Enddate;
run;

data IBD_Data5;
set IBD_Data2b;
run;

data IBD_Data5;
update IBD_Data5 IBD_Data4;
by StudyID;
run;

/********************* Calculate PUCAI, PCDAI, mathPCDAI from raw data *****************************/
proc sort data=DataIBDU out=DataIBDU;
by StudyID;
run;

data IBD_Data6;
set IBD_Data5;
merge IBD_Data2b DataIBDU;
by StudyID;

if Rt_side~=1 then Rt_side=0;
PCDAI = .;
PUCAI = .;
```plaintext
mathPCDAI = .;
/*PCDAI for CD*/
  if InitDx = 1 then PCDAI = P_AbPain + P_StoolNumCat + P_AbdoTender +
                  P_General_well_being + P_Weight + P_Height + P_Perirectal_Disease + P_EIM +
                  P_Lab_HCT + P_Lab_ESR + P_Lab_Alb;
  /*Convert PCDAI to mathPCDAI (a weighted and validated scoring tool)*/
  if InitDx = 1 AND P_AbPain = 0 then math_AbPain = 0;
  if InitDx = 1 AND P_AbPain = 5 then math_AbPain = 10;
  if InitDx = 1 AND P_AbPain = 10 then math_AbPain = 20;
  if InitDx = 1 AND P_StoolNumCat = 0 then math_StoolNumCat = 0;
  if InitDx = 1 AND P_StoolNumCat = 5 then math_StoolNumCat = 10;
  if InitDx = 1 AND P_StoolNumCat = 10 then math_StoolNumCat = 20;
  if InitDx = 1 AND P_General_well_being = 0 then math_General_well_being = 0;
  if InitDx = 1 AND P_General_well_being = 5 then math_General_well_being = 10;
  if InitDx = 1 AND P_General_well_being = 10 then math_General_well_being = 20;
  if InitDx = 1 AND P_Weight = 0 then math_Weight = 0;
  if InitDx = 1 AND P_Weight = 5 then math_Weight = 5;
  if InitDx = 1 AND P_Weight = 10 then math_Weight = 10;
  if InitDx = 1 AND P_Perirectal_Disease = 0 then math_Perirectal_Disease = 0;
  if InitDx = 1 AND P_Perirectal_Disease = 5 then math_Perirectal_Disease = 7.5;
  if InitDx = 1 AND P_Perirectal_Disease = 10 then math_Perirectal_Disease = 15;
  if InitDx = 1 AND P_Lab_ESR = 0 then math_Lab_ESR = 0;
  if InitDx = 1 AND (P_Lab_ESR = 2 OR P_Lab_ESR = 2.5) then math_Lab_ESR = 7.5;
  if InitDx = 1 AND P_Lab_Alb = 0 then math_Lab_Alb = 0;
  if InitDx = 1 AND P_Lab_Alb = 5 then math_Lab_Alb = 10;
  if InitDx = 1 AND P_Lab_Alb = 10 then math_Lab_Alb = 20;
  if InitDx = 1 then mathPCDAI = math_AbPain + math_StoolNumCat +
                          math_General_well_being + math_Weight + math_Perirectal_Disease +
                          math_EIM + math_Lab_HCT + math_Lab_ESR + math_Lab_Alb;
/*PUCAI for UC*/
  if InitDx = 2 then PUCAI = P_AbPain + P_RectalBleed + P_StoolCon +
                       P_StoolNumCat + P_Nocturnal + P_Activity;
/*PCDAI for IBD-U*/
  if InitDx = 3 then PCDAI = IBDU_abdo_pain1 + IBDU_stool___cat +
                       IBDU_activity1 + IBDU_wt + IBDU_ht + IBDU_abdo_tender + IBDU_perianal +
                       IBDU_EIM + IBDU_HCT + IBDU_ESR + IBDU_ALB;
/*PUCAI for IBD-U*/
  if InitDx = 3 then PUCAI = IBDU_abdo_pain + IBDU_rect_bld +
                     IBDU_stl_consis + IBDU___stool + IBDU_noct + IBDU_activity;
```
/*Convert PCDAI to mathPCDAI for IBD-U (a weighted and validated scoring tool)*/
if InitDx = 3 AND IBDU_abdo_pain1 = 0 then math_AbPain = 0;
  if InitDx = 3 AND IBDU_abdo_pain1 = 5 then math_AbPain = 10;
  if InitDx = 3 AND IBDU_abdo_pain1 = 10 then math_AbPain = 20;

if InitDx = 3 AND IBDU_stool__cat = 0 then math_StoolNumCat = 0;
  if InitDx = 3 AND IBDU_stool__cat = 5 then math_StoolNumCat = 10;
  if InitDx = 3 AND IBDU_stool__cat = 10 then math_StoolNumCat = 20;

if InitDx = 3 AND IBDU_activity1 = 0 then math_General_well_being = 0;
  if InitDx = 3 AND IBDU_activity1 = 5 then math_General_well_being = 10;
  if InitDx = 3 AND IBDU_activity1 = 10 then math_General_well_being = 20;

if InitDx = 3 AND IBDU_wt = 0 then math_Weight = 0;
  if InitDx = 3 AND IBDU_wt = 5 then math_Weight = 5;
  if InitDx = 3 AND IBDU_wt = 10 then math_Weight = 10;

if InitDx = 3 AND IBDU_perianal = 0 then math_Perirectal_Disease = 0;
  if InitDx = 3 AND IBDU_perianal = 5 then math_Perirectal_Disease = 7.5;
  if InitDx = 3 AND IBDU_perianal = 10 then math_Perirectal_Disease = 15;

if InitDx = 3 AND IBDU_EIM = 0 then math_EIM = 0;
  if InitDx = 3 AND IBDU_EIM = 5 then math_EIM = 10;
  if InitDx = 3 AND IBDU_EIM = 10 then math_EIM = 10;

if InitDx = 3 AND IBDU_ESR = 0 then math_Lab_ESR = 0;
  if InitDx = 3 AND IBDU_ESR = 2.5 then math_Lab_ESR = 7.5;
  if InitDx = 3 AND IBDU_ESR = 5 then math_Lab_ESR = 15;

if InitDx = 3 AND IBDU_ALB = 0 then math_Lab_Alb = 0;
  if InitDx = 3 AND IBDU_ALB = 5 then math_Lab_Alb = 10;
  if InitDx = 3 AND IBDU_ALB = 10 then math_Lab_Alb = 20;

if InitDx = 3 then mathPCDAI = math_AbPain + math_StoolNumCat +
  math_General_well_being + math_Weight + math_Perirectal_Disease + math_EIM +
  math_Lab_ESR + math_Lab_Alb;
run;

/* CD location by Paris classification*/
/* L1: Distal 1/3 ileum ± limited cecal disease*/
/* L2: Colonic*/
/* L3: Ileocolonic*/
/* L4a: Upper disease, proximal to ligament of Treitz*/
/* L4b: Upper disease, proximal to ligament of Treitz + proximal to distal 1/3 ileum*/
data IBD_Data7;
  set IBD_Data6;
  if InitDx = 1 AND (TI = “macro” AND (AscColon = “macro” OR TransCol = “macro” OR LtColon = “macro” OR Proctitis = “macro”)) then L3 = 1;
  else if InitDx = 1 AND (TI = “macro”) then L1 = 1;
else
  if InitDx = 1 AND (Cecum = "macro" OR AscColon = "macro" OR TransCol = "macro" OR LtColon = "macro" OR Proctitis = "macro") then L2 = 1;

  if InitDx = 1 AND (Oral = "macro" OR Esophagus = "macro" OR Stomach = "macro" OR Duodenum = "macro" OR UpperGI = "macro") then L4a = 1;
else
  if InitDx = 1 AND ((Oral = "macro" OR Esophagus = "macro" OR Stomach = "macro" OR Duodenum = "macro" OR UpperGI = "macro") AND JejIleum = "macro") then L4b = 1;
run;

data IBD_Data7a;
  set IBD_Data7;
  /* if L1=1 AND L4a=1 then CD_Location = 6;*/
  /* else if L1=1 AND L4b=1 then CD_Location = 7;*/
  /* else if L2=1 AND L4a=1 then CD_Location = 8;*/
  /* else if L2=1 AND L4b=1 then CD_Location = 9;*/
  /* else if L3=1 AND L4a=1 then CD_Location = 10;*/
  /* else if L3=1 AND L4b=1 then CD_Location = 11;*/
  if L1=1 then CD_Location = 1;
  else if L2=1 then CD_Location = 2;
  else if L3=1 then CD_Location = 3;
  /* else if L4a=1 then CD_Location = 4;*/
  /* else if L4b=1 then CD_Location = 5;*/
run;
/* CD behavior by Paris classification*/
/* B1: Nonstricturing non penetrating*/
/* B2: Stricturing*/
/* B3: Penetrating*/
/* B2B3: Both stricturing and penetrating*/
/* p: Perianal disease modifier*/
/*BEHAVIOR Not included in results due to too much missing data*/
data IBD_Data8;
  set IBD_Data7a;
  if Stricturing = 1 then B2 = 1;
  if Penetrating = 1 then B3 = 1;
  if Stricturing = 1 AND Penetrating = 1 then B2B3 = 1;
  if (Stricturing = 0 OR Stricturing = ".") AND (Penetrating = 0 OR Penetrating = ".") then B1 = 1;
run;
/* UC location by Paris classification*/
/* E1: Ulcerative proctitis*/
/* E2: Left sided disease*/
/* E3: Extensive disease (proximal to the splenic flexure, but distal to the hepatic flexure)*
/* E4: Pancolitis*/
/* U_Ex - IBD-U classification*/
data IBD_Data9;
  set IBD_Data8;
  if InitDx = 2 AND (Cecum = "macro" OR AscColon = "macro") then E4 = 1;
else
  if InitDx = 2 AND TransCol = "macro" then E3 = 1;
else
  if InitDx = 2 AND LtColon = "macro" then E2 = 1;
else
  if InitDx = 2 AND Proctitis = "macro" then E1 = 1;
if InitDx = 3 AND (Cecum = “macro” OR AscColon = “macro”) then U_E4 = 1;
else
  if InitDx = 3 AND TransCol = “macro” then U_E3 = 1;
  else
    if InitDx = 3 AND LtColon = “macro” then U_E2 = 1;
    else
      if InitDx = 3 AND Proctitis = “macro” then U_E1 = 1;
run;

data IBD_Data9a;
  set IBD_Data9;
  if E1=1 then UC_Location = 1;
  else if E2=1 then UC_Location = 2;
  else if E3=1 then UC_Location = 3;
  else if E4=1 then UC_Location = 4;
/* if U_E1=1 then IBDU_Location = 1;*/
/* else if U_E2=1 then IBDU_Location = 2;*/
/* else if U_E3=1 then IBDU_Location = 3;*/
/* else if U_E4=1 then IBDU_Location = 4;*/
if U_E1=1 then UC_Location = 1;
else if U_E2=1 then UC_Location = 2;
else if U_E3=1 then UC_Location = 3;
else if U_E4=1 then UC_Location = 4;
run;

/* EIM (extraintestinal manifestations of IBD) list */
/*Title “List of EIM in our cohort”;*/
/*proc freq Data=DataIBD;*/
/*  tables StudyID * EIM;*/
/*run;*/
/*ADHD Ankylosing Spondylitis Arthralgia (bilateral*/
/*wrists) Asthma Autoimmune Hepatitis Autoimmune Thyroid Disease Autimmune Thyroid Disease */
/*Celiac Disease Diabetes Erythema nodosum Hypothyroidism IDDM Iritis/Uveitis Large Joint Arthritis Oral Ulcers Orofacial Granulomatosis */
/*Other (arthralgias) */
/*Other (specify) */
/*Other Eye complaints */
/*Other Skin (folliculitis) */
/*Other Skin (oral ulcers) */
/*Primary Sclerosing Cholangitis Psoriasis Pyoderma Gangrenosum Sacro-Ileitis Small Joint Arthritis Tubular Interstitial Nephritis Vitiligo */
data IBD_Data10;
  set IBD_Data1;
  if EIM = “Autoimmune Hepatitis” then AIH = 1;
  if EIM = “Autoimmune Hepatitis” then AIH_date = EIM_Date_fix;
  if EIM = “Primary Sclerosing Cholangitis” then PSC=1;
  if EIM = “Primary Sclerosing Cholangitis” then PSC_date = EIM_Date_fix;
  if EIM = “Ankylosing Spondylitis” then AnkSpond = 1;
  if EIM = “Ankylosing Spondylitis” then AnkSpond_date = EIM_Date_fix;
  if EIM = “Arthralgia” then Arthralgia = 1;
  if EIM = “Arthralgia” then Arthralgia_date = EIM_Date_fix;
if EIM = “Asthma” then Asthma = 1;
if EIM = “Asthma” then Asthma_date = EIM_Date_fix;

if EIM = “Autoimmune Thyroid Disease” OR EIM=”Autoimmune Thyroid Disease” OR EIM=”Hypothyroidism” then Thyroid = 1;
if EIM = “Autoimmune Thyroid Disease” OR EIM=”Autoimmune Thyroid Disease” OR EIM=”Hypothyroidism” then Thyroid_date = EIM_Date_fix;

if EIM = “Celiac Disease” then Celiac = 1;
if EIM = “Celiac Disease” then Celiac_date = EIM_Date_fix;

if EIM = “Erythema nodosum” then EN = 1;
if EIM = “Erythema nodosum” then EN_date = EIM_Date_fix;

if EIM = “Diabetes” or EIM=”IDDM” then Diabetes = 1;
if EIM = “Diabetes” or EIM=”IDDM” then Diabetes_date = EIM_Date_fix;

if EIM = “Iritis/Uveitis” then uveitis = 1;
if EIM = “Iritis/Uveitis” then uveitis_date = EIM_Date_fix;

if EIM = “Large Joint Arthritis” then L_Arthritis = 1;
if EIM = “Large Joint Arthritis” then L_Arthritis_date = EIM_Date_fix;

if EIM = “Small Joint Arthritis” then S_Arthritis = 1;
if EIM = “Small Joint Arthritis” then S_Arthritis_date = EIM_Date_fix;

if EIM = “Oral Ulcers” OR EIM=”Other Skin (oral ulcers)” then o_ulcers = 1;
if EIM = “Oral Ulcers” OR EIM=”Other Skin (oral ulcers)” then o_ulcers_date = EIM_Date_fix;

if EIM = “Orofacial Granulomatosis” then OFG = 1;
if EIM = “Orofacial Granulomatosis” then OFG_date = EIM_Date_fix;

if EIM = “Psoriasis” then Psoriasis = 1;
if EIM = “Psoriasis” then Psoriasis_date = EIM_Date_fix;

if EIM = “Pyoderma Gangrenosum” then PG = 1;
if EIM = “Pyoderma Gangrenosum” then PG_date = EIM_Date_fix;

if EIM = “Sacro-Ileitis” then SI = 1;
if EIM = “Sacro-Ileitis” then SI_date = EIM_Date_fix;

if EIM = “Tubular Interstitial Nephritis” then nephritis = 1;
if EIM = “Tubular Interstitial Nephritis” then nephritis_date = EIM_Date_fix;

if EIM = “Vitiligo” then Vitiligo = 1;
if EIM = “Vitiligo” then Vitiligo_date = EIM_Date_fix;

format AIH_date PSC_date AnkSpond_date Arthralgia_date Asthma_date Thyroid_date Celiac_date EN_date Diabetes_date uveitis_date L_Arthritis_date S_Arthritis_date o_ulcers_date OFG_date Psoriasis_date PG_date SI_date nephritis_date Vitiligo_date date9.;

run;
data IBD_Data10a;
set IBD_Data10;
keep StudyID AIH AIH_date PSC PSC_date AnkSpond AnkSpond_date Arthralgia Arthralgia_date Asthma Asthma_date Thyroid Thyroid_date Celiac Celiac_date EN EN_date Diabetes Diabetes_date uveitis uveitis_date L_Arthritis L_Arthritis_date S_Arthritis S_Arthritis_date o_ulcers o_ulcers_date OFG OFG_date Psoriasis Psoriasis_date PG PG_date SI SI_date nephritis nephritis_date Vitiligo Vitiligo_date;
run;

/* List of just the 300 patients in the cohort*/
data IBD_Data10b;
  set DataDx;
  StudyID = StudyID_Info;
  drop DateDx StudyID_Info;
run;

data IBD_Data10b;
  update IBD_Data10b IBD_Data10a;
  by StudyID;
run;

data IBD_Data10b;
  update IBD_Data10b IBD_Data9a;
  by StudyID;
run;

/*convert PSC+AIH to ASC*/
data IBD_Data10c;
  set IBD_Data10b;
  if PSC=1 AND AIH=1 then ASC = 1;
  if ASC=1 then PSC=3;
  if ASC=1 then AIH=3;
  if ASC=1 then ASC_date = PSC_date;
  format ASC_date date9.;
run;

/*add a variable to denote if the patient has a history of other liver disease, EIM, or other autoimmune diseases*/
data IBD_Data10d;
  set IBD_Data10c;
  chr_liv_dis=0;
  if PSC=1 OR AIH=1 OR ASC=1 then chr_liv_dis=1;
  True_EIM=0;
  if AnkSpond = 1 OR EN = 1 OR uveitis = 1 OR L_Arthritis = 1 OR S_Arthritis = 1 OR o_ulcers = 1 OR OFG = 1 OR Psoriasis = 1 OR PG = 1 OR SI = 1 then True_EIM=1;
  Past_AutoImm = 0;
  if Asthma=1 OR Thyroid=1 OR Celiac=1 OR Diabetes=1 OR Vitiligo=1 then Past_AutoImm = 1;
run;

/*Add family history of IBD & Autoimmune disease data*/
proc sort data=DataFam;
  by StudyID;
run;

data IBD_Data11;
  set IBD_Data10d;
  FamHxIBD=(FamilyHxIBD*1);
  if FamHxIBD=. then FamHxIBD=0;
if FamHxIBD=3 then FamHxIBD=0;
run;

data Datafam1;
set Datafam;
by StudyID;

Atopy_fam=0;
AIH_fam=0;
Diabetes_fam=0;
JIA_fam=0;
RA_fam=0;
SLE_fam=0;
aloepecia_fam=0;
Arthritis_fam=0;
Celiac_fam=0;
Thyroid_fam=0;

if FamHx="Asthma" OR FamHx="Atopy" OR FamHx="Eczema" OR FamHx="allergies" OR FamHx="asthma" OR FamHx="atopy" OR FamHx="eczema" then
   Atopy_fam=1;
   if FamHx="Autoimmune hepatitis" then AIH_fam=1;
   if FamHx="Diabetes" then Diabetes_fam=1;
   if FamHx="JIA" then JIA_fam=1;
   if FamHx="Rheumatoid arthritis" then RA_fam=1;
   if FamHx="SLE" then SLE_fam=1;
   if FamHx="alopecia" then alopecia_fam=1;
   if FamHx="arthritis" then Arthritis_fam=1;
   if FamHx="celiac disease" then Celiac_fam=1;
   if FamHx="thyroid disease" then Thyroid_fam=1;

FamHx_AutoImm=0;

if FamHx="Asthma" OR FamHx="Atopy" OR FamHx="Autoimmune hepatitis" OR FamHx="Diabetes" OR FamHx="Eczema" OR FamHx="JIA" OR FamHx="Rheumatoid arthritis" OR FamHx="SLE" OR FamHx="allergies" OR FamHx="alopecia" OR FamHx="arthritis" OR FamHx="asthma" OR FamHx="atopy" OR FamHx="celiac disease" OR FamHx="eczema" OR FamHx="thyroid disease"
   then FamHx_AutoImm=1;
   drop FamHx ID;
run;

data IBD_Data11;
update IBD_Data11 Datafam1;
by StudyID;
run;

/* Add Growth z-scores at diagnosis */
/* Percentile values found at:
http://www.cdc.gov/nccdphp/dnpao/growthcharts/resources/sas.htm*/
/* The CDC provided 2 SAS programs: gc-setup-BIV & gc-calculate-BIV */
/* This code was used to generate the file to plug the growth data into the
CDC SAS programs */
data IBD_Data_Growth;
set IBD_Data1;
   AGEMOS = INT((GrowthDate - DOB)/30.4); /*convert days to months*/
   SEX = Sex;
   HEIGHT = (DxHeight*1);
WEIGHT = (DxWeight*1);
RECBUMBNT=0;
HEADCIR=0;
keep StudyID AGEMOS SEX HEIGHT WEIGHT RECBUMBNT HEADCIR;
run;
PROC EXPORT DATA= WORK.IBD_DATA_GROWTH
OUTFILE= "E:\Research_Data 2013-02-21\2012 LFT-IBD\SAS Analysis\GROWTH.xls"
DBMS=EXCELCS LABEL REPLACE;
SHEET="GROWTH";
RUN;
*/
data IBD_Data12;
set IBD_Data11;
run;
data DataGrow1;
set DataGrow;
keep StudyID WAZ WTPCT HAZ HTPCT BMI BMIZ BMIPCT; /*wt z score, wt percentile, ht z score, ht percentile, BMI z score, BMI percentile */
run;
proc sort data=DataGrow1;
by StudyID;
run;
data IBD_Data12;
update IBD_Data12 DataGrow1;
by StudyID;
run;/* Add Liver Symptoms at IBD diagnosis */
/*This data is only available for patients with abnormal liver biochemistry...not for all cases*/
data LogitudData;
set DataRed;
keep StudyID livsx_visit_type visit_date jaund_si pruritus anorexia bruising clin_hepatomegaly clin_splenomegaly clin_ascites clubbing;
if livsx_visit_type=2 or livsx_visit_type=3 then delete;
run;
proc sort data=LogitudData;
by StudyID;
run;
data IBD_Data12;
update IBD_Data12 LogitudData;
by StudyID;
run;/* Add Presumed reason associated with an event of elevated liver biochem vs diagnosis of liver disease*/
/*Liver diagnoses associated with first abnormal liver biochem abn*/
data LogitudData1;
set DataRed;
keep StudyID liv_diagnosis chronic_liv_diag chronic_dx_date chronic_liv_diag_clin chronic_liv_diag_inv acute_liv_diag acute_dx_date acute_diag_clin med_hepatotox_clin__1 med_hepatotox_clin__2 med_hepatotox_clin__3 med_hepatotox_clin__4 med_hepatotox_clin__5 med_hepatotox_clin__6 med_hepatotox_clin__7 med_hepatotox_clin__8 med_hepatotox_clin__0 inf_etiol_clin
acut_diag_inv  med_hepatotox__1  med_hepatotox__2  med_hepatotox__3  
med_hepatotox__4  med_hepatotox__5  med_hepatotox__6  med_hepatotox__7  
med_hepatotox__8  med_hepatotox__0  inf_etiol;
run;
proc sort data=LogitudData1;
   by StudyID;
run;
data IBD_Data12;
   update IBD_Data12 LogitudData1;
   by StudyID;
run;
/*/ Add Pathology labs (ANCA, ASCA) */
/*/ Just want to know if patient was ever positive*/
/*/ANCA; Negative = 0, Positive >= 20 */
data LogitudData3;
   set DataRed;
   keep StudyID anca date_lab_test;
run;
data LogitudData3a;
   set LogitudData3;
   if anca=.;
      ancaTemp=0;
   if anca="0" AND anca=.
      then ancaTemp=1;
run;
proc sort data=LogitudData3a out=LogitudData3a;
   by StudyID date_lab_test;
run;
data LogitudData3b;
   set LogitudData3a;
   by StudyID;
   if first.StudyID=1 then EverANCA = ancaTemp;
run;
/*/Using the retain statement (Cody, Learning SAS by Example: A Programmer’s
Guide, pg 517, Ch 24.8)*/
data LogitudData3c;
   set LogitudData3b;
   by StudyID;
   retain EverANCA;
   if ancaTemp = 1 then EverANCA = ancaTemp;
   else if last.ancaTemp=1 then EverANCA = ancaTemp;
   else if first.ancaTemp=1 then delete;
run;
data LogitudData3d;
   set LogitudData3c;
   by StudyID;
run;
data LogitudData3e;
   set LogitudData3d;
   drop date_lab_test anca ancaTemp;
run;
data IBD_Data12;
   update IBD_Data12 LogitudData3e;
   by StudyID;
run;
/*ASCA; Negative < 20, Positive >= 20 */
data LogitudData4;
  set DataRed;
  keep StudyID asca date_lab_test;
run;
data LogitudData4a;
  set LogitudData4;
  if asca=.
    ascaTemp=0;
  if asca>=20 then ascaTemp=1;
run;
proc sort data=LogitudData4a out=LogitudData4a;
  by StudyID date_lab_test;
run;
data LogitudData4b;
  set LogitudData4a;
  by StudyID;
  if first.StudyID=1 then EverASCA = ascaTemp;
run;
/*Using the retain statement (Cody, Learning SAS by Example: A Programmer’s Guide, pg 517, Ch 24.8)*/
data LogitudData4c;
  set LogitudData4b;
  by StudyID;
  retain EverASCA;
  if ascaTemp = 1 then EverASCA = ascaTemp;
  else if last.ascaTemp=1 then EverASCA = ascaTemp;
  else if first.ascaTemp=1 then delete;
run;
data LogitudData4d;
  set LogitudData4c;
  by StudyID;
  if first.StudyID;
run;
data LogitudData4e;
  set LogitudData4d;
  by StudyID;
  drop date_lab_test asca ascaTemp ;
run;
data IBD_Data12;
  update IBD_Data12 LogitudData4e;
  by StudyID;
run;
/* Create a dataset with CRP labs */
data LogitudData5;
  set DataRed;
  keep StudyID crp date_lab_test;
run;
data LogitudData5a;
  set LogitudData5;
  CRPdate=date_lab_test;
  format CRPdate date9.;
  drop date_lab_test;
run;
data LogitudDataCRP;
  set LogitudData5a;
  if CRP=. then delete;
run;
proc sort data=LogitudDataCRP; by StudyID; run;
proc means data=LogitudDataCRP nway chartype noprint:
class StudyID;
var crp;
output out=LogitudDataCRP1 MEAN=;
run;
data LogitudDataCRP2;
set LogitudDataCRP1;
DROP _TYPE_ _FREQ_; RUN;
data IBD_Data12;
update IBD_Data12 LogitudDataCRP2;
by StudyID;
run;
data IBD_Data_changeIBDUtoUC;
set IBD_Data12;
if InitDx=3 then InitDx=2;
run;
/********************************************************************************************/

Appendix Table 3: Dataset Manipulation – Medication data treated as Time-Varying Covariates
/*****Convert Medications from 1 column into many – according to Medication_Name, Med_Startdate, Med_Enddate*****/
/*****Add Enteral nutrition: Modulen, Tobrex, Tolerex****/
data IBD_Meds1;
set DataIBD;
keep StudyID Medication_Name Med_Startdate Med_Enddate
Enteral_feeding_name Period_used_enteral_feeding Ent_feed_Start_date
Ent_feed_End_date;
run;
data IBD_Meds1a;
set IBD_Meds1;
Med_Enddate= Med_Enddate+7;
Ent_feed_End_date= Ent_feed_End_date+7;
run;
data IBD_Meds2;
set IBD_Meds1a;
/*list all medications as not used "0", then make medication "1" if used;
Med names that were changed: 6-Mercaptopurine, Azathioprine, Budesonide (oral), Hydrocort (IV), Methotrexate, Pred (PO), Tacrolimus, Urso., Vit D*/
if Medication_Name = "6-Mercaptopurine" then MP=1;
if Medication_Name = "6-Mercaptopurine" then MP_Startdate = Med_Startdate;
if Medication_Name = "6-Mercaptopurine" then MP_Enddate = Med_Enddate;
if Medication_Name = "6-Mercaptopurine" then thio=1;
if Medication_Name = "6-Mercaptopurine" then thio_Startdate = Med_Startdate;
if Medication_Name = "6-Mercaptopurine" then thio_Enddate = Med_Enddate;
if Medication_Name = "6-Mercaptopurine" then thio_Startdate = Med_Startdate;
if Medication_Name = "6-Mercaptopurine" then thio_Enddate = Med_Startdate;
if Medication_Name = "6-Mercaptopurine" then thio_Startdate = Med_Startdate;
if Medication_Name = "Accutane" then Accutane=1;
if Medication_Name = "Accutane" then Accutane_Enddate = Med_Enddate;
if Medication_Name = "Adalimumab" then Adalimumab = 1;
if Medication_Name = "Adalimumab" then Adalimumab_Startdate = Med_Startdate;
if Medication_Name = "Adalimumab" then Adalimumab_Enddate = Med_Enddate;
if Medication_Name = "Adalimumab" then Biologics = 1;
if Medication_Name = "Adalimumab" then Biologics_Startdate = Med_Startdate;
if Medication_Name = "Adalimumab" then Biologics_Enddate = Med_Enddate;
if Medication_Name = "Azathioprine" then AZA = 1;
if Medication_Name = "Azathioprine" then AZA_Startdate = Med_Startdate;
if Medication_Name = "Azathioprine" then AZA_Enddate = Med_Enddate;
if Medication_Name = "Azathioprine" then thio = 1;
if Medication_Name = "Azathioprine" then thio_Startdate = Med_Startdate;
if Medication_Name = "Azathioprine" then thio_Enddate = Med_Enddate;
if Medication_Name = "Budesonide (oral)" then Budesonide = 1;
if Medication_Name = "Budesonide (oral)" then Budesonide_Startdate = Med_Startdate;
if Medication_Name = "Budesonide (oral)" then Budesonide_Enddate = Med_Enddate;
if Medication_Name = "Budesonide (oral)" then steroids = 1;
if Medication_Name = "Budesonide (oral)" then steroids_Startdate = Med_Startdate;
if Medication_Name = "Budesonide (oral)" then steroids_Enddate = Med_Enddate;
if Medication_Name = "Certolizumab" then Certolizumab = 1;
if Medication_Name = "Certolizumab" then Certolizumab_Startdate = Med_Startdate;
if Medication_Name = "Certolizumab" then Certolizumab_Enddate = Med_Enddate;
if Medication_Name = "Certolizumab" then Biologics = 1;
if Medication_Name = "Certolizumab" then Biologics_Startdate = Med_Startdate;
if Medication_Name = "Certolizumab" then Biologics_Enddate = Med_Enddate;
if Medication_Name = "Ciprofloxacin" then Cipro = 1;
if Medication_Name = "Ciprofloxacin" then Cipro_Startdate = Med_Startdate;
if Medication_Name = "Ciprofloxacin" then Cipro_Enddate = Med_Enddate;
if Medication_Name = "Ciprofloxacin" then Abx = 1;
if Medication_Name = "Ciprofloxacin" then Abx_Startdate = Med_Startdate;
if Medication_Name = "Ciprofloxacin" then Abx_Enddate = Med_Enddate;
if Medication_Name = "Golimumab" then Golimumab = 1;
if Medication_Name = "Golimumab" then Golimumab_Startdate = Med_Startdate;
if Medication_Name = "Golimumab" then Golimumab_Enddate = Med_Enddate;
if Medication_Name = "Golimumab" then Biologics=1;
if Medication_Name = "Golimumab" then Biologics_Startdate = Med_Startdate;
if Medication_Name = "Golimumab" then Biologics_Enddate = Med_Enddate;
if Medication_Name = "Hydrocort (IV)" then Hydrocort=1;
if Medication_Name = "Hydrocort (IV)" then Hydrocort_Startdate = Med_Startdate;
if Medication_Name = "Hydrocort (IV)" then Hydrocort_Enddate = Med_Enddate;
if Medication_Name = "Hydrocort (IV)" then steroids=1;
if Medication_Name = "Hydrocort (IV)" then steroids_Startdate = Med_Startdate;
if Medication_Name = "Hydrocort (IV)" then steroids_Enddate = Med_Enddate;
if Medication_Name = "Infliximab" then Infliximab=1;
if Medication_Name = "Infliximab" then Infliximab_Startdate = Med_Startdate;
if Medication_Name = "Infliximab" then Infliximab_Enddate = Med_Enddate;
if Medication_Name = "Infliximab" then Biologics=1;
if Medication_Name = "Infliximab" then Biologics_Startdate = Med_Startdate;
if Medication_Name = "Infliximab" then Biologics_Enddate = Med_Enddate;
if Medication_Name = "Mesalazine" then Mesalazine=1;
if Medication_Name = "Mesalazine" then Mesalazine_Startdate = Med_Startdate;
if Medication_Name = "Mesalazine" then Mesalazine_Enddate = Med_Enddate;
if Medication_Name = "Methotrexate" then MTX=1;
if Medication_Name = "Methotrexate" then MTX_Startdate = Med_Startdate;
if Medication_Name = "Methotrexate" then MTX_Enddate = Med_Enddate;
if Medication_Name = "MethylPred" then MethylPred=1;
if Medication_Name = "MethylPred" then MethylPred_Startdate = Med_Startdate;
if Medication_Name = "MethylPred" then MethylPred_Enddate = Med_Enddate;
if Medication_Name = "MethylPred" then steroids=1;
if Medication_Name = "MethylPred" then steroids_Startdate = Med_Startdate;
if Medication_Name = "MethylPred" then steroids_Enddate = Med_Enddate;
if Medication_Name = "Metronidazole" then Metronidazole=1;
if Medication_Name = "Metronidazole" then Metronidazole_Startdate = Med_Startdate;
if Medication_Name = "Metronidazole" then Metronidazole_Enddate = Med_Enddate;
if Medication_Name = "Metronidazole" then Abx=1;
if Medication_Name = "Metronidazole" then Abx_Startdate = Med_Startdate;
if Medication_Name = "Metronidazole" then Abx_Enddate = Med_Enddate;

    if Medication_Name = "Pred (PO)" then Pred=1;
if Medication_Name = "Pred (PO)" then Pred_Startdate = Med_Startdate;
    if Medication_Name = "Pred (PO)" then Pred_Enddate = Med_Enddate;
if Medication_Name = "Pred (PO)" then steroids=1;
if Medication_Name = "Pred (PO)" then steroids_Startdate = Med_Startdate;
if Medication_Name = "Pred (PO)" then steroids_Enddate = Med_Enddate;

if Medication_Name = "Sulfasalazine" then Sulfasalazine=1;
if Medication_Name = "Sulfasalazine" then Sulfasalazine_Startdate = Med_Startdate;
if Medication_Name = "Sulfasalazine" then Sulfasalazine_Enddate = Med_Enddate;

if Medication_Name = "Tacrolimus" then FK=1;
if Medication_Name = "Tacrolimus" then FK_Startdate = Med_Startdate;
if Medication_Name = "Tacrolimus" then FK_Enddate = Med_Enddate;

if Medication_Name = "Urso." then Urso=1;
if Medication_Name = "Urso." then Urso_Startdate = Med_Startdate;
if Medication_Name = "Urso." then Urso_Enddate = Med_Enddate;

if Medication_Name = "Vancomycin" then vancomycin=1;
if Medication_Name = "Vancomycin" then vancomycin_Startdate = Med_Startdate;
if Medication_Name = "Vancomycin" then vancomycin_Enddate = Med_Enddate;

if Medication_Name = "Vit D" then VitD=1;
if Medication_Name = "Vit D" then VitD_Startdate = Med_Startdate;
if Medication_Name = "Vit D" then VitD_Enddate = Med_Enddate;

if Medication_Name = "vancomycin" then vancomycin=1;
if Medication_Name = "vancomycin" then vancomycin_Startdate = Med_Startdate;
if Medication_Name = "vancomycin" then vancomycin_Enddate = Med_Enddate;

if Enteral_feeding_name = "Modulen" then EEN=1;
if Enteral_feeding_name = "Modulen" then EEN_Startdate = Ent_feed_Start_date;
if Enteral_feeding_name = "Modulen" then EEN_Enddate = Ent_feed_End_date;

if Enteral_feeding_name = "Tobrex" then EEN=1;
if Enteral_feeding_name = "Tobrex" then EEN_Startdate = Ent_feed_Start_date;
if Enteral_feeding_name = "Tobrex" then EEN_Enddate = Ent_feed_End_date;

if Enteral_feeding_name = "Tolerex" then EEN=1;
if Enteral_feeding_name = "Tolerex" then EEN_Startdate = Ent_feed_Start_date;
if Enteral_feeding_name = "Tolerex" then EEN_Enddate = Ent_feed_End_date;

DROP Medication_Name;
DROP Med_Startdate;
DROP Med_Enddate;

DROP Enteral_feeding_name;
DROP Period_used_enteral_feeding;
DROP Ent_feed_Start_date;
DROP Ent_feed_End_date;

format MP_Startdate Accutane_Startdate Adalimumab_Startdate
AZA_Startdate Budesonide_Startdate Certolizumab_Startdate Cipro_Startdate
Golimumab_Startdate
Hydrocort_Startdate Infliximab_Startdate
Mesalazine_Startdate MTX_Startdate MethylPred_Startdate
Metronidazole_Startdate Pred_Startdate
Sulfasalazine_Startdate FK_Startdate Urso_Startdate
Vancomycin_Startdate VitD_Startdate EEN_Startdate thio_Startdate
steroids_Startdate Abx_Startdate Biologics_Startdate
MP_Enddate Accutane_Enddate Adalimumab_Enddate AZA_Enddate
Budesonide_Enddate Certolizumab_Enddate Cipro_Enddate Golimumab_Enddate
Hydrocort_Enddate
Infliximab_Enddate Mesalazine_Enddate MTX_Enddate
MethyLPred_Enddate Metronidazole_Enddate Pred_Enddate Sulfasalazine_Enddate
FK_Enddate Urso_Enddate
Vancomycin_Enddate VitD_Enddate EEN_Enddate thio_Enddate
steroids_Enddate Abx_Enddate Biologics_Enddate
date9.;
run;
/*Create separate datasets for each medication*/
data IBD_Meds_6MP;
set IBD_Meds2;
keep StudyID MP MP_Startdate MP_Enddate;
if MP=. then delete;
run;
data IBD_Meds_6MP_start;
set IBD_Meds_6MP;
start = MP_Startdate;
start1 = 1;
run;
data IBD_Meds_6MP_end;
set IBD_Meds_6MP;
start = MP_Enddate;
start1 = 0;
run;
data IBD_Meds_Accutane;
set IBD_Meds2;
keep StudyID Accutane Accutane_Startdate Accutane_Enddate;
if Accutane=. then delete;
run;
data IBD_Meds_Accutane_start;
set IBD_Meds_Accutane;
start = Accutane_Startdate;
start1 = 1;
run;
data IBD_Meds_Accutane_end;
set IBD_Meds_Accutane;
start = Accutane_Enddate;
start1 = 0;
run;
data IBD_Meds_Adalimumab;
set IBD_Meds2;
keep StudyID Adalimumab Adalimumab_Startdate Adalimumab_Enddate;
if Adalimumab=. then delete;
run;
data IBD_Meds_Adalimumab_start;
set IBD_Meds_Adalimumab;
start = Adalimumab_Startdate;
start1 = 1;
run;
data IBD_Meds_Adalimumab_end;
set IBD_Meds_Adalimumab;
start = Adalimumab_Enddate;
start1 = 0;
run;
data IBD_Meds_AZA;
set IBD_Meds2;
keep StudyID AZA AZA_Startdate AZA_Enddate;
if AZA=. then delete;
run;
data IBD_Meds_AZA_start;
set IBD_Meds_AZA;
start = AZA_Startdate;
start1 = 1;
run;
data IBD_Meds_AZA_end;
set IBD_Meds_AZA;
start = AZA_Enddate;
run;
data IBD_Meds_Budesonide;
set IBD_Meds2;
keep StudyID Budesonide Budesonide_Startdate Budesonide_Enddate;
if Budesonide=. then delete;
run;
data IBD_Meds_Budesonide_start;
set IBD_Meds_Budesonide;
start = Budesonide_Startdate;
start1 = 1;
run;
data IBD_Meds_Budesonide_end;
set IBD_Meds_Budesonide;
start = Budesonide_Enddate;
start1 = 0;
run;
data IBD_Meds_Certolizumab;
set IBD_Meds2;
keep StudyID Certolizumab Certolizumab_Startdate Certolizumab_Enddate;
if Certolizumab=. then delete;
run;
data IBD_Meds_Certolizumab_start;
set IBD_Meds_Certolizumab;
start = Certolizumab_Startdate;
start1 = 1;
run;
    data IBD_Meds_Certolizumab_end;
        set IBD_Meds_Certolizumab;
        start = Certolizumab_Enddate;
        start1 = 0;
    run;

data IBD_Meds_Cipro;
    set IBD_Meds2;
    keep StudyID Cipro Cipro_Startdate Cipro_Enddate;
    if Cipro=. then delete;
run;
    data IBD_Meds_Cipro_start;
        set IBD_Meds_Cipro ;
        start = Cipro_Startdate;
        start1 = 1;
    run;
    data IBD_Meds_Cipro_end;
        set IBD_Meds_Cipro ;
        start = Cipro_Enddate;
        start1 = 0;
    run;

data IBD_Meds_Golimumab;
    set IBD_Meds2;
    keep StudyID Golimumab Golimumab_Startdate Golimumab_Enddate;
    if Golimumab=. then delete;
run;
    data IBD_Meds_Golimumab_start;
        set IBD_Meds_Golimumab;
        start = Golimumab_Startdate;
        start1 = 1;
    run;
    data IBD_Meds_Golimumab_end;
        set IBD_Meds_Golimumab;
        start = Golimumab_Enddate;
        start1 = 0;
    run;

data IBD_Meds_Hydrocort;
    set IBD_Meds2;
    keep StudyID Hydrocort Hydrocort_Startdate Hydrocort_Enddate;
    if Hydrocort=. then delete;
run;
    data IBD_Meds_Hydrocort_start;
        set IBD_Meds_Hydrocort;
        start = Hydrocort_Startdate;
        start1 = 1;
    run;
    data IBD_Meds_Hydrocort_end;
        set IBD_Meds_Hydrocort;
        start = Hydrocort_Enddate;
        start1 = 0;
    run;

data IBD_Meds_Infliximab;
    set IBD_Meds2;
    keep StudyID Infliximab Infliximab_Startdate Infliximab_Enddate;
    if Infliximab=. then delete;
run;
    data IBD_Meds_Infliximab_start;
        set IBD_Meds_Infliximab;
start = Infliximab_Startdate;
start1 = 1;
run;
data IBD_Meds_Infliximab_end;
    set IBD_Meds_Infliximab;
    start = Infliximab_Enddate;
    start1 = 0;
run;
data IBD_Meds_Mesalazine;
    set IBD_Meds2;
    keep StudyID Mesalazine Mesalazine_Startdate Mesalazine_Enddate;
    if Mesalazine= . then delete;
run;
data IBD_Meds_Mesalazine_start;
    set IBD_Meds_Mesalazine;
    start = Mesalazine_Startdate;
    start1 = 1;
run;
data IBD_Meds_Mesalazine_end;
    set IBD_Meds_Mesalazine;
    start = Mesalazine_Enddate;
    start1 = 0;
run;
data IBD_Meds_MTX;
    set IBD_Meds2;
    keep StudyID MTX MTX_Startdate MTX_Enddate;
    if MTX= . then delete;
run;
data IBD_Meds_MTX_start;
    set IBD_Meds_MTX;
    start = MTX_Startdate;
    start1 = 1;
run;
data IBD_Meds_MTX_end;
    set IBD_Meds_MTX;
    start = MTX_Enddate;
    start1 = 0;
run;
data IBD_Meds_MethylPred;
    set IBD_Meds2;
    keep StudyID MethylPred MethylPred_Startdate MethylPred_Enddate;
    if MethylPred= . then delete;
run;
data IBD_Meds_MethylPred_start;
    set IBD_Meds_MethylPred;
    start = MethylPred_Startdate;
    start1 = 1;
run;
data IBD_Meds_MethylPred_end;
    set IBD_Meds_MethylPred;
    start = MethylPred_Enddate;
    start1 = 0;
run;
data IBD_Meds_Metronidazole;
    set IBD_Meds2;
    keep StudyID Metronidazole Metronidazole_Startdate Metronidazole_Enddate;
    if Metronidazole= . then delete;
run;
  data IBD_Meds_Metronidazole_start;
    set IBD_Meds_Metronidazole;
    start = Metronidazole_Startdate;
    start1 = 1;
  run;
  data IBD_Meds_Metronidazole_end;
    set IBD_Meds_Metronidazole;
    start = Metronidazole_Enddate;
    start1 = 0;
  run;
  data IBD_Meds_Pred;
    set IBD_Meds2;
    keep StudyID Pred Pred_Startdate Pred_Enddate;
    if Pred=. then delete;
  run;
  data IBD_Meds_Pred_start;
    set IBD_Meds_Pred;
    start = Pred_Startdate;
    start1 = 1;
  run;
  data IBD_Meds_Pred_end;
    set IBD_Meds_Pred;
    start = Pred_Enddate;
    start1 = 0;
  run;
  data IBD_Meds_Sulfasalazine;
    set IBD_Meds2;
    keep StudyID Sulfasalazine Sulfasalazine_Startdate Sulfasalazine_Enddate;
    if Sulfasalazine=. then delete;
  run;
  data IBD_Meds_Sulfasalazine_start;
    set IBD_Meds_Sulfasalazine;
    start = Sulfasalazine_Startdate;
    start1 = 1;
  run;
  data IBD_Meds_Sulfasalazine_end;
    set IBD_Meds_Sulfasalazine;
    start = Sulfasalazine_Enddate;
    start1 = 0;
  run;
  data IBD_Meds_FK;
    set IBD_Meds2;
    keep StudyID FK FK_Startdate FK_Enddate;
    if FK=. then delete;
  run;
  data IBD_Meds_FK_start;
    set IBD_Meds_FK;
    start = FK_Startdate;
    start1 = 1;
  run;
  data IBD_Meds_FK_end;
    set IBD_Meds_FK;
    start = FK_Enddate;
    start1 = 0;
  run;
  data IBD_Meds_Urso;
set IBD_Meds2;
keep StudyID Urso Urso_Startdate Urso_Enddate;
if Urso=. then delete;
run;

data IBD_Meds_Urso_start;
  set IBD_Meds_Urso;
  start = Urso_Startdate;
  start1 = 1;
run;

data IBD_Meds_Urso_end;
  set IBD_Meds_Urso;
  start = Urso_Enddate;
  start1 = 0;
run;

data IBD_Meds_vancomycin;
  set IBD_Meds2;
  keep StudyID vancomycin vancomycin_Startdate vancomycin_Enddate;
if vancomycin=. then delete;
run;

data IBD_Meds_vancomycin_start;
  set IBD_Meds_vancomycin;
  start = vancomycin_Startdate;
  start1 = 1;
run;

data IBD_Meds_vancomycin_end;
  set IBD_Meds_vancomycin;
  start = vancomycin_Enddate;
  start1 = 0;
run;

data IBD_Meds_VitD;
  set IBD_Meds2;
  keep StudyID VitD VitD_Startdate VitD_Enddate;
if VitD=. then delete;
run;

data IBD_Meds_VitD_start;
  set IBD_Meds_VitD;
  start = VitD_Startdate;
  start1 = 1;
run;

data IBD_Meds_VitD_end;
  set IBD_Meds_VitD;
  start = VitD_Enddate;
  start1 = 0;
run;

data IBD_Meds_EEN;
  set IBD_Meds2;
  keep StudyID EEN EEN_Startdate EEN_Enddate;
if EEN=. then delete;
run;

data IBD_Meds_EEN_start;
  set IBD_Meds_EEN;
  start = EEN_Startdate;
  start1 = 1;
run;

data IBD_Meds_EEN_end;
  set IBD_Meds_EEN;
  start = EEN_Enddate;
  start1 = 0;
run;
data IBD_Meds_thio;
  set IBD_Meds2;
  keep StudyID thio thio_Startdate thio_Enddate;
  if thio=. then delete;
run;

data IBD_Meds_thio_start;
  set IBD_Meds_thio;
  start = thio_Startdate;
  start1 = 1;
run;

data IBD_Meds_thio_end;
  set IBD_Meds_thio;
  start = thio_Enddate;
  start1 = 0;
run;

data IBD_Meds_Biologics;
  set IBD_Meds2;
  keep StudyID Biologics Biologics_Startdate Biologics_Enddate;
  if Biologics=. then delete;
run;

data IBD_Meds_Biologics_start;
  set IBD_Meds_Biologics;
  start = Biologics_Startdate;
  start1 = 1;
run;

data IBD_Meds_Biologics_end;
  set IBD_Meds_Biologics;
  start = Biologics_Enddate;
  start1 = 0;
run;

data IBD_Meds_steroids;
  set IBD_Meds2;
  keep StudyID steroids steroids_Startdate steroids_Enddate;
  if steroids=. then delete;
run;

data IBD_Meds_steroids_start;
  set IBD_Meds_steroids;
  start = steroids_Startdate;
  start1 = 1;
run;

data IBD_Meds_steroids_end;
  set IBD_Meds_steroids;
  start = steroids_Enddate;
  start1 = 0;
run;

data IBD_Meds_Abx;
  set IBD_Meds2;
  keep StudyID Abx Abx_Startdate Abx_Enddate;
  if Abx=. then delete;
run;

data IBD_Meds_Abx_start;
  set IBD_Meds_Abx;
  start = Abx_Startdate;
  start1 = 1;
run;

data IBD_Meds_Abx_end;
  set IBD_Meds_Abx;
start = Abx_Enddate;
start1 = 0;
/*Remove repeated values of the same start date (presumably the end date would be the same as well)*/
proc sort data= IBD_Meds_6MP out=IBD_Meds_6MP_sorted NODUPKEY;
    by StudyID MP_Startdate MP_Enddate;
run;
proc sort data= IBD_Meds_Accutane out=IBD_Meds_Accutane_sorted NODUPKEY;
    by StudyID Accutane_Startdate Accutane_Enddate;
run;
proc sort data= IBD_Meds_Adalimumab out=IBD_Meds_Adalimumab_sorted NODUPKEY;
    by StudyID Adalimumab_Startdate Adalimumab_Enddate;
run;
proc sort data= IBD_Meds_AZA out=IBD_Meds_AZA_sorted NODUPKEY;
    by StudyID AZA_Startdate AZA_Enddate;
run;
proc sort data= IBD_Meds_Budesonide out=IBD_Meds_Budesonide_sorted NODUPKEY;
    by StudyID Budesonide_Startdate Budesonide_Enddate;
run;
proc sort data= IBD_Meds_Certolizumab out=IBD_Meds_Certolizumab_sorted NODUPKEY;
    by StudyID Certolizumab_Startdate Certolizumab_Enddate;
run;
proc sort data= IBD_Meds_Cipro out=IBD_Meds_Cipro_sorted NODUPKEY;
    by StudyID Cipro_Startdate Cipro_Enddate;
run;
proc sort data= IBD_Meds_Golimumab out=IBD_Meds_Golimumab_sorted NODUPKEY;
    by StudyID Golimumab_Startdate Golimumab_Enddate;
run;
proc sort data= IBD_Meds_Hydrocort out=IBD_Meds_Hydrocort_sorted NODUPKEY;
    by StudyID Hydrocort_Startdate Hydrocort_Enddate;
run;
proc sort data= IBD_Meds_Infliximab out=IBD_Meds_Infliximab_sorted NODUPKEY;
    by StudyID Infliximab_Startdate Infliximab_Enddate;
run;
proc sort data= IBD_Meds_Mesalazine out=IBD_Meds_Mesalazine_sorted NODUPKEY;
    by StudyID Mesalazine_Startdate Mesalazine_Enddate;
run;
proc sort data= IBD_Meds_MTX out=IBD_Meds_MTX_sorted NODUPKEY;
    by StudyID MTX_Startdate MTX_Enddate;
run;
proc sort data= IBD_Meds_MethylPred out=IBD_Meds_MethylPred_sorted NODUPKEY;
    by StudyID MethylPred_Startdate MethylPred_Enddate;
run;
proc sort data= IBD_Meds_Metronidazole out=IBD_Meds_Metronidazole_sorted NODUPKEY;
    by StudyID Metronidazole_Startdate Metronidazole_Enddate;
run;
proc sort data= IBD_Meds_Pred out=IBD_Meds_Pred_sorted NODUPKEY;
    by StudyID Pred_Startdate Pred_Enddate;
run;
proc sort data= IBD_Meds_Sulfasalazine out=IBD_Meds_Sulfasalazine_sorted NODUPKEY;
    by StudyID Sulfasalazine_Startdate Sulfasalazine_Enddate;
run;
proc sort data= IBD_Meds_FK out=IBD_Meds_FK_sorted NODUPKEY;
    by StudyID FK_Startdate FK_Enddate;
run;
proc sort data=IBD_Meds_Urso out=IBD_Meds_Urso_sorted NODUPKEY;
   by StudyID Urso_Startdate Urso_Enddate;
run;
proc sort data=IBD_Meds_vancomycin out=IBD_Meds_vancomycin_sorted NODUPKEY;
   by StudyID vancomycin_Startdate vancomycin_Enddate;
run;
proc sort data=IBD_Meds_VitD out=IBD_Meds_VitD_sorted NODUPKEY;
   by StudyID VitD_Startdate VitD_Enddate;
run;
proc sort data=IBD_Meds_EEN out=IBD_Meds_EEN_sorted NODUPKEY;
   by StudyID EEN_Startdate EEN_Enddate;
run;
proc sort data=IBD_Meds_thio out=IBD_Meds_thio_sorted NODUPKEY;
   by StudyID thio_Startdate thio_Enddate;
run;
proc sort data=IBD_Meds_Biologics out=IBD_Meds_Biologics_sorted NODUPKEY;
   by StudyID Biologics_Startdate Biologics_Enddate;
run;
proc sort data=IBD_Meds_steroids out=IBD_Meds_steroids_sorted NODUPKEY;
   by StudyID steroids_Startdate steroids_Enddate;
run;
proc sort data=IBD_Meds_Abx out=IBD_Meds_Abx_sorted NODUPKEY;
   by StudyID Abx_Startdate Abx_Enddate;
run;
/*Create separate datasets for each medication; one dataset for start time (start1=1), and one for end time (start1=0)*/
  data IBD_Meds_6MP_start;
    set IBD_Meds_6MP_sorted;
    start = MP_Startdate;
    start1 = 1;
run;
data IBD_Meds_6MP_end;
    set IBD_Meds_6MP_sorted;
    start = MP_Enddate;
    start1 = 0;
run;
data IBD_Meds_Accutane_start;
    set IBD_Meds_Accutane_sorted;
    start = Accutane_Startdate;
    start1 = 1;
run;
data IBD_Meds_Accutane_end;
    set IBD_Meds_Accutane_sorted;
    start = Accutane_Enddate;
    start1 = 0;
run;
data IBD_Meds_A达尔他umab_start;
    set IBD_Meds_A达尔他umab_sorted;
    start = Adal他umab_Startdate;
    start1 = 1;
run;
data IBD_Meds_A达尔他umab_end;
    set IBD_Meds_A达尔他umab_sorted;
    start = Adal他umab_Enddate;
    start1 = 0;
run;
data IBD_Meds_AZA_start;
set IBD_Meds_AZA_sorted;
start = AZA_Startdate;
start1 = 1;
run;
proc print data=IBD_Meds_AZA_start (obs=50); run;
data IBD_Meds_AZA_end;
set IBD_Meds_AZA_sorted;
start = AZA_Enddate;
start1 = 0;
run;
data IBD_Meds_Budesonide_start;
set IBD_Meds_Budesonide_sorted;
start = Budesonide_Startdate;
start1 = 1;
run;
data IBD_Meds_Budesonide_end;
set IBD_Meds_Budesonide_sorted;
start = Budesonide_Enddate;
start1 = 0;
run;
data IBD_Meds_Certolizumab_start;
set IBD_Meds_Certolizumab_sorted;
start = Certolizumab_Startdate;
start1 = 1;
run;
data IBD_Meds_Certolizumab_end;
set IBD_Meds_Certolizumab_sorted;
start = Certolizumab_Enddate;
start1 = 0;
run;
data IBD_Meds_Cipro_start;
set IBD_Meds_Cipro_sorted;
start = Cipro_Startdate;
start1 = 1;
run;
data IBD_Meds_Cipro_end;
set IBD_Meds_Cipro_sorted;
start = Cipro_Enddate;
start1 = 0;
run;
data IBD_Meds_Golimumab_start;
set IBD_Meds_Golimumab_sorted;
start = Golimumab_Startdate;
start1 = 1;
run;
data IBD_Meds_Golimumab_end;
set IBD_Meds_Golimumab_sorted;
start = Golimumab_Enddate;
start1 = 0;
run;
data IBD_Meds_Hydrocort_start;
set IBD_Meds_Hydrocort_sorted;
start = Hydrocort_Startdate;
start1 = 1;
run;
data IBD_Meds_Hydrocort_end;
set IBD_Meds_Hydrocort_sorted;
start = Hydrocort_Enddate;
start1 = 0;
run;
data IBD_Meds_Infliximab_start;
  set IBD_Meds_Infliximab_sorted;
  start = Infliximab_Startdate;
  start1 = 1;
run;
data IBD_Meds_Infliximab_end;
  set IBD_Meds_Infliximab_sorted;
  start = Infliximab_Enddate;
  start1 = 0;
run;
data IBD_Meds_Mesalazine_start;
  set IBD_Meds_Mesalazine_sorted;
  start = Mesalazine_Startdate;
  start1 = 1;
run;
data IBD_Meds_Mesalazine_end;
  set IBD_Meds_Mesalazine_sorted;
  start = Mesalazine_Enddate;
  start1 = 0;
run;
data IBD_Meds_MTX_start;
  set IBD_Meds_MTX_sorted;
  start = MTX_Startdate;
  start1 = 1;
run;
data IBD_Meds_MTX_end;
  set IBD_Meds_MTX_sorted;
  start = MTX_Enddate;
  start1 = 0;
run;
data IBD_Meds_MethylPred_start;
  set IBD_Meds_MethylPred_sorted;
  start = MethylPred_Startdate;
  start1 = 1;
run;
data IBD_Meds_MethylPred_end;
  set IBD_Meds_MethylPred_sorted;
  start = MethylPred_Enddate;
  start1 = 0;
run;
data IBD_Meds_Metronidazole_start;
  set IBD_Meds_Metronidazole_sorted;
  start = Metronidazole_Startdate;
  start1 = 1;
run;
data IBD_Meds_Metronidazole_end;
  set IBD_Meds_Metronidazole_sorted;
  start = Metronidazole_Enddate;
  start1 = 0;
run;
data IBD_Meds_Pred_start;
  set IBD_Meds_Pred_sorted;
  start = Pred_Startdate;
  start1 = 1;
run;
data IBD_Meds_Pred_end;
set IBD_Meds_Pred_sorted;
start = Pred_Enddate;
start1 = 0;
run;
data IBD_Meds_Sulfasalazine_start;
set IBD_Meds_Sulfasalazine_sorted;
start = Sulfasalazine_Startdate;
start1 = 1;
run;
data IBD_Meds_Sulfasalazine_end;
set IBD_Meds_Sulfasalazine_sorted;
start = Sulfasalazine_Enddate;
start1 = 0;
run;
data IBD_Meds_FK_start;
set IBD_Meds_FK_sorted;
start = FK_Startdate;
start1 = 1;
run;
data IBD_Meds_FK_end;
set IBD_Meds_FK_sorted;
start = FK_Enddate;
start1 = 0;
run;
data IBD_Meds_Urso_start;
set IBD_Meds_Urso_sorted;
start = Urso_Startdate;
start1 = 1;
run;
data IBD_Meds_Urso_end;
set IBD_Meds_Urso_sorted;
start = Urso_Enddate;
start1 = 0;
run;
data IBD_Meds_vancomycin_start;
set IBD_Meds_vancomycin_sorted;
start = vancomycin_Startdate;
start1 = 1;
run;
data IBD_Meds_vancomycin_end;
set IBD_Meds_vancomycin_sorted;
start = vancomycin_Enddate;
start1 = 0;
run;
data IBD_Meds_VitD_start;
set IBD_Meds_VitD_sorted;
start = VitD_Startdate;
start1 = 1;
run;
data IBD_Meds_VitD_end;
set IBD_Meds_VitD_sorted;
start = VitD_Enddate;
start1 = 0;
run;
data IBD_Meds_EEN_start;
set IBD_Meds_EEN_sorted;
start = EEN_Startdate;
start1 = 1;
run;
data IBD_Meds_EEN_end;
  set IBD_Meds_EEN_sorted;
  start = EEN_Enddate;
  start1 = 0;
run;
data IBD_Meds_thio_start;
  set IBD_Meds_thio_sorted;
  start = thio_Startdate;
  start1 = 1;
run;
data IBD_Meds_thio_end;
  set IBD_Meds_thio_sorted;
  start = thio_Enddate;
  start1 = 0;
run;
data IBD_Meds_Biologics_start;
  set IBD_Meds_Biologics_sorted;
  start = Biologics_Startdate;
  start1 = 1;
run;
data IBD_Meds_Biologics_end;
  set IBD_Meds_Biologics_sorted;
  start = Biologics_Enddate;
  start1 = 0;
run;
data IBD_Meds_steroids_start;
  set IBD_Meds_steroids_sorted;
  start = steroids_Startdate;
  start1 = 1;
run;
data IBD_Meds_steroids_end;
  set IBD_Meds_steroids_sorted;
  start = steroids_Enddate;
  start1 = 0;
run;
data IBD_Meds_Abx_start;
  set IBD_Meds_Abx_sorted;
  start = Abx_Startdate;
  start1 = 1;
run;
data IBD_Meds_Abx_end;
  set IBD_Meds_Abx_sorted;
  start = Abx_Enddate;
  start1 = 0;
run;
/*Concatenate the data*/
data IBD_Meds3;
  set IBD_Meds_6MP_start IBD_Meds_Accutane_start IBD_Meds_Adalimumab_start
  IBD_Meds_AZA_start IBD_Meds_Budesonide_start
  IBD_Meds_Certolizumab_start IBD_Meds_Cipro_start
  IBD_Meds_Golimumab_start IBD_Meds_Hydrocort_start IBD_Meds_Inflixiamb_start
  IBD_Meds_Mesalazine_start IBD_Meds_MTX_start
  IBD_Meds_MethylPred_start IBD_Meds_Metronidazole_start IBD_Meds_Pred_start
  IBD_Meds_Sulfasalazine_start
  IBD_Meds_FK_start IBD_Meds_Urso_start IBD_Meds_vancomycin_start
  IBD_Meds_VitD_start IBD_Meds_EEN_start
IBD_Meds_thio_start IBD_Meds_Biologics_start
IBD_Meds_steroids_start IBD_Meds_Abx_start
IBD_Meds_AZA_end IBD_Meds_Budesonide_end
IBD_Meds_6MP_end IBD_Meds_Certolizumab_end IBD_Meds_Cipro_end
IBD_Meds_Golimumab end IBD_Meds_Hydrocort_end IBD_Meds_Infliximab_end
IBD_Meds_Mesalazine_end IBD_Meds_MTX_end IBD_Meds_MethylPred_end
IBD_Meds_Metronidazole_end IBD_Meds_Pred_end IBD_Meds_Sulfasalazine_end
IBD_Meds_FK_end IBD_Meds_Urso_end IBD_Meds_vancomycin_end
IBD_Meds_MethylPred_end IBD_Meds_VitD_end IBD_Meds_EEN_end
IBD_Meds_thio_end IBD_Meds_Biologics_end IBD_Meds_steroids_end
IBD_Meds_Abx_end;

format start date9.;
run;

data IBD_Meds4;
  set IBD_Meds3;
    if MP=1 and MP_Enddate = . then MP_Enddate = '31JAN2014'D;
    if Accutane=1 and Accutane_Enddate = . then Accutane_Enddate = '31JAN2014'D;
    if Adalimumab =1 and Adalimumab_Enddate = . then
      Adalimumab_Enddate = '31JAN2014'D;
    if AZA = 1 and AZA_Enddate= . then AZA_Enddate = '31JAN2014'D;
    if Budesonide = 1 and Budesonide_Enddate= . then
      Budesonide_Enddate = '31JAN2014'D;
    if Certolizumab = 1 and Certolizumab_Enddate= . then
      Certolizumab_Enddate = '31JAN2014'D;
    if Cipro = 1 and Cipro_Enddate= . then Cipro_Enddate = '31JAN2014'D;
    if Golimumab = 1 and Golimumab_Enddate= . then Golimumab_Enddate = '31JAN2014'D;
    if Hydrocort = 1 and Hydrocort_Enddate= . then Hydrocort_Enddate = '31JAN2014'D;
    if Infliximab = 1 and Infliximab_Enddate= . then Infliximab_Enddate = '31JAN2014'D;
    if Mesalazine = 1 and Mesalazine_Enddate= . then Mesalazine_Enddate = '31JAN2014'D;
    if MTX = 1 and MTX_Enddate= . then MTX_Enddate = '31JAN2014'D;
    if MethylPred = 1 and MethylPred_Enddate= . then
      MethylPred_Enddate = '31JAN2014'D;
    if Metronidazole = 1 and Metronidazole_Enddate= . then
      Metronidazole_Enddate = '31JAN2014'D;
    if Pred = 1 and Pred_Enddate= . then Pred_Enddate = '31JAN2014'D;
    if Sulfasalazine = 1 and Sulfasalazine_Enddate= . then
      Sulfasalazine_Enddate = '31JAN2014'D;
    if FK = 1 and FK_Enddate= . then FK_Enddate = '31JAN2014'D;
    if Urso = 1 and Urso_Enddate= . then Urso_Enddate = '31JAN2014'D;
    if vancomycin = 1 and vancomycin_Enddate= . then
      vancomycin_Enddate = '31JAN2014'D;
    if EEN = 1 and EEN_Enddate= . then EEN_Enddate = '31JAN2014'D;
    if thio = 1 and thio_Enddate= . then thio_Enddate = '31JAN2014'D;
    if Biologics = 1 and Biologics_Enddate= . then Biologics_Enddate = '31JAN2014'D;
    if steroids = 1 and steroids_Enddate= . then steroids_Enddate = '31JAN2014'D;
    if Abx = 1 and Abx_Enddate= . then Abx_Enddate = '31JAN2014'D;
    if start = . then start = '31JAN2014'D;
run;
/*This code creates a dataset with the times at which meds were changed.*/
/*For each patient, you then need to use these times to construct intervals
(a,b): the first interval will be (missing,27Nov2001), the next
(27Nov2001,28Dec2001).*/
/*Once you’ve done that, you can add columns for each med to indicate whether
or not the patient was on the med at that time.*/
proc means data=IBD_Meds5  NOPRINT;
  by StudyID start start1;
  var MP Accutane Adalimumab AZA Budesonide Certolizumab Cipro Golimumab
        Hydrocort Infliximab Mesalazine MTX MethylPred Metronidazole Pred
        Sulfasalazine FK Urso vancomycin VitD EEN thio Biologics steroids Abx;
  output out=medtimes;
run;
data medtimes;
  set medtimes;
  where _STAT_="MAX";
run;
data medtimes;
  set medtimes;
  where start<>.;
run;
data intervals1;
  set medtimes;
  keep StudyID start;
  rename start=time;
run;
proc print data=intervals1 (obs=50); Title "Dataset with the times at which
meds were changed"; run;
PROC EXPORT
DATA=intervals1
DBMS=EXCELCS REPLACE
LABEL
OUTFILE="D:\Research Data 2013-02-21\2012 LFT-IBD\SAS
Analysis\Intervals1.xls"
REPLACE;
PROC EXPORT
DATA=IBD_Meds5
DBMS=EXCELCS REPLACE
LABEL
OUTFILE="D:\Research Data 2013-02-21\2012 LFT-IBD\SAS Analysis\IBD_Meds5.xls"
REPLACE;
/*previous commands took too much memory*/
proc datasets library=work kill nolist;
quit;
run;
/*Start over the data commands here*/
PROC IMPORT OUT= WORK.intervals1
  DATAFILE="D:\Research Data 2013-02-21\2012 LFT-IBD\SAS
Analysis\Intervals1.xls"
  DBMS=EXCELCS REPLACE;
  RANGE="INTERVALS1$";
  SCANTEXT=YES;
  USEDATE=YES;
SCANTIME=YES;
RUN;
/*Convert the time column to start_time and end_time*/
proc sort data=intervals1 out=intervals2 NODUPKEY;
   by StudyID descending time;
run;
data intervals3;
   set intervals2;
   by StudyID;
   retain temp_time;
   format temp_time date9.;
   if first.StudyID=1 then temp_time = time; /*Set the date for the first abnormal result*/
   format end_time date9.;
   if first.StudyID~=`1` then do;
      start_time=time;
      end_time=temp_time;
      temp_time=time;
      format start_time end_time date9.;
      output;
   end;
run;
proc sort data=intervals3 out=intervals4;
   by StudyID start_time;
run;
data intervals;
   set intervals4;
   drop time temp_time;
run;
PROC EXPORT
DATA=intervals
DBMS=EXCELCS REPLACE
LABEL
OUTFILE="D:\Research Data 2013-02-21\2012 LFT-IBD\SAS Analysis\intervals.xls"
REPLACE;
/*previous commands took too much memory*/
proc datasets library=work kill nolist;
quit;
run;
PROC IMPORT OUT= WORK.intervals
   DATAFILE = "D:\Research Data 2013-02-21\2012 LFT-IBD\SAS Analysis\intervals.xls"
   DBMS=EXCELCS REPLACE;
   RANGE="INTERVALS$";
   SCANTEXT=YES;
   USEDATE=YES;
   SCANTIME=YES;
RUN;
PROC IMPORT OUT= WORK.IBD_Meds5
   DATAFILE = "D:\Research Data 2013-02-21\2012 LFT-IBD\SAS Analysis\IBD_Meds5.xls"
   DBMS=EXCELCS REPLACE;
   RANGE="IBD_MEDS5$";
   SCANTEXT=YES;
   USEDATE=YES;
   SCANTIME=YES;
RUN;
data IBD_Meds6a;
  set intervals;
run;
proc sort data=IBD_Meds6a out=IBD_Meds6;
  by StudyID start_time;
run;
/*MP*/
proc sql;
  create table IBD_Meds_MP as
  select intervals.StudyID as temp_StudyID,
    start_time, end_time,
    IBD_Meds5.StudyID as Data_StudyID,
    MP, MP_Startdate, MP_Enddate, start, start1
  from intervals,
    IBD_Meds5;
quit;
data IBD_Meds_MP1;
  set IBD_Meds_MP;
  if temp_StudyID ~= Data_StudyID then delete;
  if MP = . then delete;
run;
data IBD_Meds_MP2;
  set IBD_Meds_MP1;
    if start_time < MP_Startdate then delete;
    if end_time > MP_Enddate then delete;
    StudyID = temp_StudyID;
  drop temp_StudyID Data_StudyID start start1;
run;
proc sort data=IBD_Meds_MP2 out=IBD_Meds_MP3 NODUPKEY;
  by StudyID start_time;
run;
data IBD_Meds6;
  update IBD_Meds6 IBD_Meds_MP3;
  by StudyID start_time;
run;
proc datasets library=work nolist;
delete IBD_Meds_MP IBD_Meds_MP1 IBD_Meds_MP2 IBD_Meds_MP3;
quit;
run;
/*Accutane*/
proc sql;
  create table IBD_Meds_Accutane as
  select intervals.StudyID as temp_StudyID,
    start_time, end_time,
    IBD_Meds5.StudyID as Data_StudyID,
    Accutane, Accutane_Startdate, Accutane_Enddate, start, start1
  from intervals,
    IBD_Meds5;
quit;
data IBD_Meds_Accutane1;
  set IBD_Meds_Accutane;
    if temp_StudyID ~= Data_StudyID then delete;
    if Accutane = . then delete;
run;
data IBD_Meds_Accutane2;
set IBD_Meds_Accutane1;

if start_time < Accutane_Startdate then delete;
if end_time > Accutane_Enddate then delete;
StudyID = temp_StudyID;
drop temp_StudyID Data_StudyID start start1;
run;
proc sort data=IBD_Meds_Accutane2 out=IBD_Meds_Accutane3 NODUPKEY;
   by StudyID start_time;
run;
data IBD_Meds6;
   update IBD_Meds6 IBD_Meds_Accutane3;
   by StudyID start_time;
run;
proc datasets library=work nolist;
delete IBD_Meds_ACCUTANE IBD_Meds_ACCUTANE1 IBD_Meds_ACCUTANE2
  IBD_Meds_ACCUTANE3;
quit;
run;
/*Adalimumab*/
proc sql;
   create table IBD_Meds_Adalimumab as
      select intervals.StudyID as temp_StudyID,
            start_time, end_time,
            IBD_Meds5.StudyID as Data_StudyID,
            Adalimumab, Adalimumab_Startdate, Adalimumab_Enddate,
      start, start1
         from intervals,
            IBD_Meds5;
   quit;
data IBD_Meds_Adalimumab1;
   set IBD_Meds_Adalimumab;
   if temp_StudyID ~= Data_StudyID then delete;
   if Adalimumab = . then delete;
run;
data IBD_Meds_Adalimumab2;
   set IBD_Meds_Adalimumab1;

   if start_time < Adalimumab_Startdate then delete;
   if end_time > Adalimumab_Enddate then delete;
   StudyID = temp_StudyID;
   drop temp_StudyID Data_StudyID start start1;
run;
proc sort data=IBD_Meds_Adalimumab2 out=IBD_Meds_Adalimumab3 NODUPKEY;
   by StudyID start_time;
run;
data IBD_Meds6;
   update IBD_Meds6 IBD_Meds_Adalimumab3;
   by StudyID start_time;
run;
proc datasets library=work nolist;
delete IBD_Meds_ADALIMUMAB IBD_Meds_ADALIMUMAB1 IBD_Meds_ADALIMUMAB2
  IBD_Meds_ADALIMUMAB3;
quit;
run;
/*AZA*/
proc sql;
   create table IBD_Meds_AZA as
select intervals.StudyID as temp_StudyID, 
    start_time, end_time, 
    IBD_Meds5.StudyID as Data_StudyID, 
    AZA, AZA_Startdate, AZA_Enddate, start, start1 
from intervals, 
    IBD_Meds5;
quit;
data IBD_Meds_AZA1;
    set IBD_Meds_AZA;
    if temp_StudyID ^= Data_StudyID then delete;
    if AZA = . then delete;
run;
data IBD_Meds_AZA2;
    set IBD_Meds_AZA1;
    if start_time < AZA_Startdate then delete;
    if end_time > AZA_Enddate then delete;
    StudyID = temp_StudyID;
    drop temp_StudyID Data_StudyID start start1;
run;
proc sort data=IBD_Meds_AZA2 out=IBD_Meds_AZA3 NODUPKEY;
    by StudyID start_time;
run;
data IBD_Meds6;
    update IBD_Meds6 IBD_Meds_AZA3;
    by StudyID start_time;
run;
proc datasets library=work nolist;
delete IBD_Meds_AZA IBD_Meds_AZA1 IBD_Meds_AZA2 IBD_Meds_AZA3;
quit;
run;
/*Budesonide*/
proc sql;
    create table intervals.StudyID as temp_StudyID, 
    start_time, end_time, 
    IBD_Meds5.StudyID as Data_StudyID, 
    Budesonide, Budesonide_Startdate, Budesonide_Enddate, 
start, start1 
from intervals, 
    IBD_Meds5;
quit;
data IBD_Meds_Budesonide1;
    set IBD_Meds_Budesonide;
    if temp_StudyID ^= Data_StudyID then delete;
    if Budesonide = . then delete;
run;
data IBD_Meds_Budesonide2;
    set IBD_Meds_Budesonide1;
    if start_time < Budesonide_Startdate then delete;
    if end_time > Budesonide_Enddate then delete;
    StudyID = temp_StudyID;
    drop temp_StudyID Data_StudyID start start1;
run;
proc sort data=IBD_Meds_Budesonide2 out=IBD_Meds_Budesonide3 NODUPKEY;
    by StudyID start_time;
run;
data IBD_Meds6;
    update IBD_Meds6 IBD_Meds_Budesonide3;
    by StudyID start_time;
run;
proc datasets library=work nolist;
delete IBD_Meds_BUDESONIDE IBD_Meds_BUDESONIDE1 IBD_Meds_BUDESONIDE2
     IBD_Meds_BUDESONIDE3;
quit;
run;
/*Certolizumab*/
proc sql;
    create table IBD_Meds_Certolizumab as
    select intervals.StudyID as temp_StudyID,
          start_time, end_time,
          IBD_Meds5.StudyID as Data_StudyID,
          Certolizumab, Certolizumab_Startdate, Certolizumab_Enddate,
          start, start1
    from intervals,
         IBD_Meds5;
quit;
data IBD_Meds_Certolizumab;
    set IBD_Meds_Certolizumab;
    if temp_StudyID ^= Data_StudyID then delete;
    if Certolizumab = . then delete;
run;
data IBD_Meds_Certolizumab2;
    set IBD_Meds_Certolizumab;
    if start_time < Certolizumab_Startdate then delete;
    if end_time > Certolizumab_Enddate then delete;
    StudyID = temp_StudyID;
    drop temp_StudyID Data_StudyID start start1;
run;
proc sort data=IBD_Meds_Certolizumab2 out=IBD_Meds_Certolizumab3 nodupkey;
    by StudyID start_time;
run;
data IBD_Meds6;
    update IBD_Meds6 IBD_Meds_Certolizumab3;
    by StudyID start_time;
run;
proc datasets library=work nolist;
delete IBD_Meds_CERTOLIZUMAB IBD_Meds_CERTOLIZUMAB1 IBD_Meds_CERTOLIZUMAB2
     IBD_Meds_CERTOLIZUMAB3;
quit;
run;
/*Cipro*/
proc sql;
    create table IBD_Meds_Cipro as
    select intervals.StudyID as temp_StudyID,
          start_time, end_time,
          IBD_Meds5.StudyID as Data_StudyID,
          Cipro, Cipro_Startdate, Cipro_Enddate, start, start1
    from intervals,
         IBD_Meds5;
quit;
data IBD_Meds_Cipro;
    set IBD_Meds_Cipro;
    if temp_StudyID ^= Data_StudyID then delete;
if Cipro = . then delete;
run;
data IBD_Meds_Cipro2;
  set IBD_Meds_Cipro1;
    if start_time < Cipro_Startdate then delete;
    if end_time > Cipro_Enddate then delete;
    StudyID = temp StudyID;
  drop temp StudyID Data_StudyID start start1;
run;
proc sort data=IBD_Meds_Cipro2 out=IBD_Meds_Cipro3 NODUPKEY;
  by StudyID start_time;
run;
data IBD_Meds6;
  update IBD_Meds6 IBD_Meds_Cipro3;
  by StudyID start_time;
run;
data IBD_Meds6;
  update IBD_Meds6 IBD_Meds_Cipro3;
  by StudyID start_time;
run;
proc datasets library=work nolist;
delete IBD_Meds_CIPRO IBD_Meds_CIPRO1 IBD_Meds_CIPRO2 IBD_Meds_CIPRO3;
quit;
run; /*Golimumab*/
proc sql;
  create table IBD_Meds_Golimumab as
  select intervals.StudyID as temp_StudyID,
    start_time, end_time,
    IBD_Meds5.StudyID as Data_StudyID,
    Golimumab, Golimumab_Startdate, Golimumab_Enddate, start,
    start1
    from intervals,
    IBD_Meds5;
quit;
data IBD_Meds_Golimumab1;
  set IBD_Meds_Golimumab;
  if temp_StudyID ~= Data_StudyID then delete;
  if Golimumab = . then delete;
run;
data IBD_Meds_Golimumab2;
  set IBD_Meds_Golimumab1;
    if start_time < Golimumab_Startdate then delete;
    if end_time > Golimumab_Enddate then delete;
    StudyID = temp StudyID;
  drop temp StudyID Data_StudyID start start1;
run;
proc sort data=IBD_Meds_Golimumab2 out=IBD_Meds_Golimumab3 NODUPKEY;
  by StudyID start_time;
run;
data IBD_Meds6;
  update IBD_Meds6 IBD_Meds_Golimumab3;
  by StudyID start_time;
run;
proc datasets library=work nolist;
/*Hydrocort*/
proc sql;
cREATE TABLE IBD_Meds_Hydrocort AS
SELECT intervals.StudyID AS temp_StudyID,
    start_time, end_time,
    IBD_Meds5.StudyID AS Data_StudyID,
    Hydrocort, Hydrocort_Startdate, Hydrocort_Enddate, start,
    start1
FROM intervals,
    IBD_Meds5;
QUIT;
DATA IBD_Meds_Hydrocort1;
SET IBD_Meds_Hydrocort;
IF temp_StudyID ^= Data_StudyID THEN DELETE;
IF Hydrocort = . THEN DELETE;
RUN;
DATA IBD_Meds_Hydrocort2;
SET IBD_Meds_Hydrocort1;
    IF start_time < Hydrocort_Startdate THEN DELETE;
    IF end_time > Hydrocort_Enddate THEN DELETE;
    StudyID = temp_StudyID;
    DROP temp_StudyID Data_StudyID start start1;
RUN;
PROC SORT DATA=IBD_Meds_Hydrocort2 OUT=IBD_Meds_Hydrocort3 NODUPKEY;
    BY StudyID start_time;
RUN;
DATA IBD_Meds6;
    UPDATE IBD_Meds6 IBD_Meds_Hydrocort3;
    BY StudyID start_time;
RUN;
PROC DATASETS LIBRARY=work Nolist;
DELETE IBD_Meds_HYDROCORT IBD_Meds_HYDROCORT1 IBD_Meds_HYDROCORT2
    IBD_Meds_HYDROCORT3;
QUIT;
RUN;
/*Infliximab*/
proc sql;
cREATE TABLE IBD_Meds_Infliximab AS
SELECT intervals.StudyID AS temp_StudyID,
    start_time, end_time,
    IBD_Meds5.StudyID AS Data_StudyID,
    Infliximab, Infliximab_Startdate, Infliximab_Enddate, start,
    start1
FROM intervals,
    IBD_Meds5;
QUIT;
DATA IBD_Meds_Infliximab1;
SET IBD_Meds_Infliximab;
IF temp_StudyID ^= Data_StudyID THEN DELETE;
IF Infliximab = . THEN DELETE;
RUN;
DATA IBD_Meds_Infliximab2;
SET IBD_Meds_Infliximab1;
```sql
if start_time < Infliximab_Startdate then delete;
if end_time > Infliximab_Enddate then delete;
StudyID = temp_SudyID;
drop temp_SudyID Data_SudyID start start1;
run;
proc sort data=IBD_Meds_Infliximab2 out=IBD_Meds_Infliximab3 NODUPKEY;
by StudyID start_time;
run;
data IBD_Meds6;
update IBD_Meds6 IBD_Meds_Infliximab3;
by StudyID start_time;
run;
proc datasets library=work nolist;
delete IBD_Meds_INFLIXIMAB IBD_Meds_INFLIXIMAB1 IBD_Meds_INFLIXIMAB2
IBD_Meds_INFLIXIMAB3;
quit;
run;
/*Mesalazine*/
proc sql;
create table IBD_Meds_Mesalazine as
select intervals.StudyID as temp_SudyID,
start_time, end_time,
IBD_Meds5.StudyID as Data_SudyID,
Mesalazine, Mesalazine_Startdate, Mesalazine_Enddate,
start, start1
from intervals,
IBD_Meds5;
quit;
data IBD_Meds_Mesalazine1;
set IBD_Meds_Mesalazine;
if temp_SudyID ~= Data_SudyID then delete;
if Mesalazine = . then delete;
run;
data IBD_Meds_Mesalazine2;
set IBD_Meds_Mesalazine1;
if start_time < Mesalazine_Startdate then delete;
if end_time > Mesalazine_Enddate then delete;
StudyID = temp_SudyID;
drop temp_SudyID Data_SudyID start start1;
run;
proc sort data=IBD_Meds_Mesalazine2 out=IBD_Meds_Mesalazine3 NODUPKEY;
by StudyID start_time;
run;
data IBD_Meds6;
update IBD_Meds6 IBD_Meds_Mesalazine3;
by StudyID start_time;
run;
proc datasets library=work nolist;
delete IBD_Meds_MESALAZINE IBD_Meds_MESALAZINE1 IBD_Meds_MESALAZINE2
IBD_Meds_MESALAZINE3;
quit;
run;
/*MTX*/
proc sql;
create table IBD_Meds_MTX as
select intervals.StudyID as temp_SudyID,
```
start_time, end_time,
   IBD_Meds5.StudyID as Data_StudyID,
   MTX, MTX_Startdate, MTX_Enddate, start, start1
   from intervals,
   IBD_Meds5;
quit;

data IBD_Meds_MTX1;
   set IBD_Meds_MTX;
   if temp_StudyID ~= Data_StudyID then delete;
   if MTX = . then delete;
run;

data IBD_Meds_MTX2;
   set IBD_Meds_MTX1;
   if start_time < MTX_Startdate then delete;
   if end_time > MTX_Enddate then delete;
   StudyID = temp_StudyID;
   drop temp_StudyID Data_StudyID start start1;
run;

proc sort data=IBD_Meds_MTX2 out=IBD_Meds_MTX3 NODUPKEY;
   by StudyID start_time;
run;

data IBD_Meds6;
   update IBD_Meds6 IBD_Meds_MTX3;
   by StudyID start_time;
run;

proc datasets library=work nolist;
   delete IBD_Meds_MTX IBD_Meds_MTX1 IBD_Meds_MTX2 IBD_Meds_MTX3;
quit;
run;

/*MethylPred*/
proc sql;
   create table IBD_Meds_MethylPred as
   select intervals.StudyID as temp_StudyID,
       start_time, end_time,
       IBD_Meds5.StudyID as Data_StudyID,
       MethylPred, MethylPred_Startdate, MethylPred_Enddate,
       start, start1
       from intervals,
       IBD_Meds5;
quit;

data IBD_Meds_MethylPred1;
   set IBD_Meds_MethylPred;
   if temp_StudyID ~= Data_StudyID then delete;
   if MethylPred = . then delete;
run;

data IBD_Meds_MethylPred2;
   set IBD_Meds_MethylPred1;
   if start_time < MethylPred_Startdate then delete;
   if end_time > MethylPred_Enddate then delete;
   StudyID = temp_StudyID;
   drop temp_StudyID Data_StudyID start start1;
run;
proc sort data=IBD_Meds_MethylPred2 out=IBD_Meds_MethylPred3 NODUPKEY;
   by StudyID start_time;
run;

data IBD_Meds6;
update IBD_Meds6 IBD_Meds_Methylpred3;
by StudyID start_time;
run;
proc datasets library=work nolist;
delete IBD_Meds_METHYLPRED IBD_Meds_METHYLPRED1 IBD_Meds_METHYLPRED2 IBD_Meds_METHYLPRED3;
quit;
run;
/*Metronidazole*/
proc sql;
create table IBD_Meds_Metronidazole as
select intervals.StudyID as temp_StudyID,
       start_time, end_time,
       IBD_Meds5.StudyID as Data_StudyID,
       Metronidazole, Metronidazole_Startdate, Metronidazole_Enddate, start, start1
from intervals,
     IBD_Meds5;
quit;
data IBD_Meds_Metronidazole1;
set IBD_Meds_Metronidazole;
if temp_StudyID ~= Data_StudyID then delete;
if Metronidazole = . then delete;
run;
data IBD_Meds_Metronidazole2;
set IBD_Meds_Metronidazole1;
if start_time < Metronidazole_Startdate then delete;
if end_time > Metronidazole_Enddate then delete;
StudyID = temp_StudyID;
drop temp_StudyID Data_StudyID start start1;
run;
proc sort data=IBD_Meds_Metronidazole2 out=IBD_Meds_Metronidazole3 NODUPKEY;
by StudyID start_time;
run;
data IBD_Meds6;
update IBD_Meds6 IBD_Meds_Metronidazole3;
by StudyID start_time;
run;
proc datasets library=work nolist;
delete IBD_Meds_METRONIDAZOLE IBD_Meds_METRONIDAZOLE1 IBD_Meds_METRONIDAZOLE2 IBD_Meds_METRONIDAZOLE3;
quit;
run;
/*Pred*/
proc sql;
create table IBD_Meds_Pred as
select intervals.StudyID as temp_StudyID,
       start_time, end_time,
       IBD_Meds5.StudyID as Data_StudyID,
       Pred, Pred_Startdate, Pred_Enddate, start, start1
from intervals,
     IBD_Meds5;
quit;
data IBD_Meds_Pred1;
set IBD_Meds_Pred;
if temp_StudyID ~= Data_StudyID then delete;
if Pred = . then delete;
run;
data IBD_Meds_Pred2;
set IBD_Meds_Pred1;

if start_time < Pred_Startdate then delete;
if end_time > Pred_Enddate then delete;
StudyID = temp_StudyID;
drop temp_StudyID Data_StudyID start start1;
run;
proc sort data=IBD_Meds_Pred2 out=IBD_Meds_Pred3 NODUPKEY;
by StudyID start_time;
run;
data IBD_Meds6;
update IBD_Meds6 IBD_Meds_Pred3;
by StudyID start_time;
run;
proc datasets library=work nolist;
delete IBD_Meds_PRED IBD_Meds_PRED1 IBD_Meds_PRED2 IBD_Meds_PRED3;
quit;
run;
/*Sulfasalazine*/
proc sql;
create table IBD_Meds_Sulfasalazine as
select intervals.StudyID as temp_StudyID,
start_time, end_time,
IBD_Meds5.StudyID as Data_StudyID,
Sulfasalazine, Sulfasalazine_Startdate,
Sulfasalazine_Enddate, start, start1
from intervals,
IBD_Meds5;
quit;
data IBD_Meds_Sulfasalazine1;
set IBD_Meds_Sulfasalazine;
if temp_StudyID ~= Data_StudyID then delete;
if Sulfasalazine = . then delete;
run;
data IBD_Meds_Sulfasalazine2;
set IBD_Meds_Sulfasalazine1;

if start_time < Sulfasalazine_Startdate then delete;
if end_time > Sulfasalazine_Enddate then delete;
StudyID = temp_StudyID;
drop temp_StudyID Data_StudyID start start1;
run;
proc sort data=IBD_Meds_Sulfasalazine2 out=IBD_Meds_Sulfasalazine3 NODUPKEY;
by StudyID start_time;
run;
data IBD_Meds6;
update IBD_Meds6 IBD_Meds_Sulfasalazine3;
by StudyID start_time;
run;
proc datasets library=work nolist;
delete IBD_Meds_SULFASALAZINE IBD_Meds_SULFASALAZINE1 IBD_Meds_SULFASALAZINE2 IBD_Meds_SULFASALAZINE3;
quit;
run;
/*FK*/
proc sql;
create table IBD_Meds_FK as
select intervals.StudyID as temp_StudyID,
    start_time, end_time,
    IBD_Meds5.StudyID as Data_StudyID,
    FK, FK_Startdate, FK_Enddate, start, start1
from intervals,
    IBD_Meds5;
quit;
data IBD_Meds_FK1;
set IBD_Meds_FK;
if temp_StudyID =~ Data_StudyID then delete;
if FK = . then delete;
run;
data IBD_Meds_FK2;
set IBD_Meds_FK1;
if start_time < FK_Startdate then delete;
if end_time > FK_Enddate then delete;
StudyID = temp_StudyID;
drop temp_StudyID Data_StudyID start start1;
run;
proc sort data=IBD_Meds_FK out=IBD_Meds_FK3 NODUPKEY;
    by StudyID start_time;
run;
data IBD_Meds6;
    update IBD_Meds6 IBD_Meds_FK3;
    by StudyID start_time;
run;
proc datasets library=work nolist;
delete IBD_Meds_FK IBD_Meds_FK1 IBD_Meds_FK2 IBD_Meds_FK3;
quit;
run;
/*vancomycin*/
proc sql;
create table IBD_Meds_vancomycin as
select intervals.StudyID as temp_StudyID,
    start_time, end_time,
    IBD_Meds5.StudyID as Data_StudyID,
    vancomycin, vancomycin_Startdate, vancomycin_Enddate,
start, start1
from intervals,
    IBD_Meds5;
quit;
data IBD_Meds_vancomycin1;
set IBD_Meds_vancomycin;
if temp_StudyID =~ Data_StudyID then delete;
if vancomycin = . then delete;
run;
data IBD_Meds_vancomycin2;
set IBD_Meds_vancomycin1;
if start_time < vancomycin_Startdate then delete;
if end_time > vancomycin_Enddate then delete;
StudyID = temp_StudyID;
drop temp_StudyID Data_StudyID start start1;
run;
proc sort data=IBD_Meds_vancomycin2 out=IBD_Meds_vancomycin3 NODUPKEY;
by StudyID start_time;
run;
data IBD_Meds6;
update IBD_Meds6 IBD_Meds_vancomycin3;
by StudyID start_time;
run;
proc datasets library=work nolist;
delete IBD_Meds_VANCOMYCIN IBD_Meds_VANCOMYCIN1 IBD_Meds_VANCOMYCIN2
IBD_Meds_VANCOMYCIN3;
quit;
run;
/*EEN*/
proc sql;
create table IBD_Meds_EEN as
select intervals.StudyID as temp_StudyID,
      start_time, end_time,
      IBD_Meds5.StudyID as Data_StudyID,
      EEN, EEN_Startdate, EEN_Enddate, start, start1
from intervals,
     IBD_Meds5;
quit;
data IBD_Meds_EEN1;
set IBD_Meds_EEN;
if temp_StudyID ~= Data_StudyID then delete;
if EEN = . then delete;
run;
data IBD_Meds_EEN2;
set IBD_Meds_EEN1;
if start_time < EEN_Startdate then delete;
if end_time > EEN_Enddate then delete;
StudyID = temp_StudyID Data_StudyID start start1;
drop temp_StudyID Data_StudyID start start1;
run;
proc sort data=IBD_Meds_EEN2 out=IBD_Meds_EEN3 NODUPKEY;
by StudyID start_time;
run;
data IBD_Meds6;
update IBD_Meds6 IBD_Meds_EEN3;
by StudyID start_time;
run;
proc datasets library=work nolist;
delete IBD_Meds_EEN IBD_Meds_EEN1 IBD_Meds_EEN2 IBD_Meds_EEN3;
quit;
run;
/*thio*/
proc sql;
create table IBD_Meds_thio as
select intervals.StudyID as temp_StudyID,
      start_time, end_time,
      IBD_Meds5.StudyID as Data_StudyID,
      thio, thio_Startdate, thio_Enddate, start, start1
from intervals,
     IBD_Meds5;
quit;
data IBD_Meds_thio1;
set IBD_Meds_thio;
if temp_StudyID ~= Data_StudyID then delete;
if thio = . then delete;
run;
data IBD_Meds_thio2;
  set IBD_Meds_thio1;
  if start_time < thio_Startdate then delete;
  if end_time > thio_Enddate then delete;
  StudyID = temp_StudyID;
  drop temp_StudyID Data_StudyID start start1;
run;
proc sort data=IBD_Meds_thio2 out=IBD_Meds_thio3 NODUPKEY;
  by StudyID start_time;
run;
data IBD_Meds6;
  update IBD_Meds6 IBD_Meds_thio3;
  by StudyID start_time;
run;
proc datasets library=work nolist;
  delete IBD_Meds_THIO IBD_Meds_THIO1 IBD_Meds_THIO2 IBD_Meds_THIO3;
quit;
run;
/*Biologics*/
proc sql;
  create table IBD_Meds_Biologics as
  select intervals.StudyID as temp_StudyID, start_time, end_time,
    IBD_Meds5.StudyID as Data_StudyID,
    Biologics, Biologics_Startdate, Biologics_Enddate, start,
    start1
  from intervals,
    IBD_Meds5;
quit;
data IBD_Meds_Biologics1;
  set IBD_Meds_Biologics;
  if temp_StudyID ^= Data_StudyID then delete;
  if Biologics = . then delete;
run;
data IBD_Meds_Biologics2;
  set IBD_Meds_Biologics1;
  if start_time < Biologics_Startdate then delete;
  if end_time > Biologics_Enddate then delete;
  StudyID = temp_StudyID;
  drop temp_StudyID Data_StudyID start start1;
run;
proc sort data=IBD_Meds_Biologics2 out=IBD_Meds_Biologics3 NODUPKEY;
  by StudyID start_time;
run;
data IBD_Meds6;
  update IBD_Meds6 IBD_Meds_Biologics3;
  by StudyID start_time;
run;
proc datasets library=work nolist;
  delete IBD_Meds_BIOLOGICS IBD_Meds_BIOLOGICS1 IBD_Meds_BIOLOGICS2
  IBD_Meds_BIOLOGICS3;
quit;
run;
/*steroids*/
proc sql;
  create table IBD_Meds_steroids as
  select intervals.StudyID as temp_StudyID,
         start_time, end_time,
         IBD_Meds5.StudyID as Data_StudyID,
         steroids, steroids_Startdate, steroids_Enddate, start,
         start1
  from intervals,
       IBD_Meds5;
QUIT;
DATA IBD_Meds_steroids1;
  set IBD_Meds_steroids;
  if temp_StudyID ~= Data_StudyID then delete;
  if steroids = . then delete;
RUN;
DATA IBD_Meds_steroids2;
  set IBD_Meds_steroids1;
  if start_time < steroids_Startdate then delete;
  if end_time > steroids_Enddate then delete;
  StudyID = temp_StudyID;
  drop temp_StudyID Data_StudyID start start1;
RUN;
PROC SORT DATA=IBD_Meds_steroids2 OUT=IBD_Meds_steroids3 NODUPKEY;
  BY StudyID start_time;
RUN;
DATA IBD_Meds6;
  UPDATE IBD_Meds6 IBD_Meds_Steroids3;
  BY StudyID start_time;
RUN;
PROC DATASETS LIBRARY=work Nolist;
DELETE IBD_Meds_STEROIDS IBD_Meds_STEROIDS1 IBD_Meds_STEROIDS2
       IBD_Meds_STEROIDS3;
QUIT;
RUN;
/*Abx*/
PROC SQL;
  CREATE TABLE IBD_Meds_Abx AS
  SELECT intervals.StudyID as temp_StudyID,
         start_time, end_time,
         IBD_Meds5.StudyID as Data_StudyID,
         Abx, Abx_Startdate, Abx_Enddate, start, start1
  FROM intervals,
       IBD_Meds5;
QUIT;
DATA IBD_Meds_Abx1;
  set IBD_Meds_Abx;
  if temp_StudyID ~= Data_StudyID then delete;
  if Abx = . then delete;
RUN;
DATA IBD_Meds_Abx2;
  set IBD_Meds_Abx1;
  if start_time < Abx_Startdate then delete;
  if end_time > Abx_Enddate then delete;
  StudyID = temp_StudyID;
  drop temp_StudyID Data_StudyID start start1;
RUN;
*Appendix Table 4: Dataset Manipulation – Liver enzymes data merged with clinical IBD data*

data IBD_Data13;
    set IBD_Data_changeIBDUtoUC;
run;
data IBD_Data14;
    merge IBD_Data13 all_ptn_time_to_event_FU;
    by StudyID;
run;
/* Add Pathology labs (ANCA, ANA, ASMA, ALKA) */
/* "LabPath" 1= positive tests between the time frame of the IBD diagnosis and the abnormal liver biochem (+30 days), 0=all negative*/
data LogitudData2;
    set DataRed;
    keep StudyID date_lab_test anca ana asm alkm asca;
run;
proc sort data=LogitudData2;
    by StudyID;
run;
data all_ptn_time_to_event_FU_1;
    set all_ptn_time_to_event_FU;
    keep StudyID Date_Collected; /* date of Abn result vs the last date of liver biochem in patients with normal results*/
run;
data LogitudData2;
    merge LogitudData2 all_ptn_time_to_event_FU_1;
    by StudyID;
run;
data LogitudData2a;
    set LogitudData2;
    if (anca~=. OR ana~=. OR asm~=. OR alkm~.);
LabPathTEMP=0;
    if (date_lab_test <= (Date_Collected+30)) AND ((anca~="0" AND anca~=.)
      OR (ana~="0" AND ana~=.) OR (asm~="0" AND asm~=.) OR (alkm~="0" AND alkm~=.)
    ) then LabPathTEMP=1;
run;
proc sort data=LogitudData2a out=LogitudData2a;
    by StudyID date_lab_test;
run;
/*Using the retain statement (Cody, Learning SAS by Example: A Programmer’s
Guide, pg 517, Ch 24.8)*/
data LogitudData2b;
    set LogitudData2a;
    by StudyID;
    if first.StudyID=1 then LabPath = LabPathTEMP;
    if first.StudyID=1 then PathDate = date_lab_test; /*Set the
date for the path lab result*/
    format PathDate date9.;
run;
data LogitudData2c;
    set LogitudData2b;
    by StudyID;
    retain LabPath PathDate;
    if LabPathTEMP = 1 then do;
        LabPath = LabPathTEMP;
        PathDate = date_lab_test;
        format PathDate date9.;
        output;
    end;
    else do;
        if last.LabPathTEMP=1 then LabPath = LabPathTEMP;
        if last.LabPathTEMP=1 then PathDate = date_lab_test;
        else if first.LabPathTEMP=1 then delete;
        output;
    end;
run;
data LogitudData2d;
    set LogitudData2c;
    by StudyID;
    if first.StudyID;
run;
data LogitudData2e;
    set LogitudData2d;
    drop Date_Collected date_lab_test anca ana asm alkm asca LabPathTEMP;
run;
data IBD_Data15;
    merge IBD_Data14 LogitudData2e;
    by StudyID;
run;
/*****Add variables for medication analysis: time varying covariates*******/
/*****Add medAbn 0=no, 1=yes****/
proc sort data=IBD_Meds out=IBD_Meds;
    by StudyID start_time;
run;
data IBD_Meds1;
merge IBD_Meds all_ptn_time_to_event_FU;
by StudyID;
run;

data IBD_Meds2;
set IBD_Meds1;
if StudyID = . then delete;

if orig_coll_date < datedx then Date_Collected = datedx;
if end_time<datedx then delete;
if start_time>Date_Collected then delete;
if start_time<datedx OR start_time= . then do; start_time=datedx; end;
if end_time>Date_Collected OR end_time= . then do;
end_time=Date_Collected; end;
run;
data IBD_Meds_ADD_Dx_Interval1;
set IBD_Meds2;
by StudyID;
if First.StudyID=1 AND start_time~datedx;
run;
data IBD_Meds_ADD_Dx_Interval2;
set IBD_Meds_ADD_Dx_Interval1;
end_time=start_time;
start_time=DateDx;
run;
/*Concatenate to dataset*/
data IBD_Meds_ADD_Dx_Interval3;
set IBD_Meds2 IBD_Meds_ADD_Dx_Interval2;
run;
proc sort data=IBD_Meds ADD_Dx_Interval3 out=IBD_Meds2a;
by StudyID start_time;
run;
data IBD_Meds3;
set IBD_Meds2a;
/*Meds used in each time interval: start date before start_time, start
date exists, end date after end_time of interval; *Note: all end dates exist,
and '.' values set to Date_Collected*/
if MP_Startdate <= start_time AND MP_Enddate >= end_time
then MP_b4abn=1;
if Adalimumab_Startdate <= start_time AND Adalimumab_Enddate >= end_time
then Adalimumab_b4abn=1;
if AZA_Startdate <= start_time AND AZA_Enddate >= end_time
then AZA_b4abn=1;
if Budesonide_Startdate <= start_time AND Budesonide_Enddate >= end_time
then Budesonide_b4abn=1;
if Certolizumab_Startdate <= start_time AND Certolizumab_Enddate >= end_time
then Certolizumab_b4abn=1;
if Cipro_Startdate <= start_time AND Cipro_Enddate >= end_time
then Cipro_b4abn=1;
if Golimumab_Startdate <= start_time AND Golimumab_Enddate >= end_time
then Golimumab_b4abn=1;
if Hydrocort_Startdate <= start_time AND Hydrocort_Startdate=. AND Hydrocort_Enddate >= end_time then Hydrocort_b4abn=1;
if Infliximab_Startdate <= start_time AND Infliximab_Startdate=. AND Infliximab_Enddate >= end_time then Infliximab_b4abn=1;
if Mesalazine_Startdate <= start_time AND Mesalazine_Startdate=. AND Mesalazine_Enddate >= end_time then Mesalazine_b4abn=1;
if MTX_Startdate <= start_time AND MTX_Startdate=. AND MTX_Enddate >= end_time then MTX_b4abn=1;
if MethylPred_Startdate <= start_time AND MethylPred_Startdate=. AND MethylPred_Enddate >= end_time then MethylPred_b4abn=1;
if Metronidazole_Startdate <= start_time AND Metronidazole_Startdate=. AND Metronidazole_Enddate >= end_time then Metronidazole_b4abn=1;
if Pred_Startdate <= start_time AND Pred_Startdate=. AND Pred_Enddate >= end_time then Pred_b4abn=1;
if Sulfasalazine_Startdate <= start_time AND Sulfasalazine_Startdate=. AND Sulfasalazine_Enddate >= end_time then Sulfasalazine_b4abn=1;
if FK_Startdate <= start_time AND FK_Startdate=. AND FK_Enddate >= end_time then FK_b4abn=1;
if Urso_Startdate <= start_time AND Urso_Startdate=. AND Urso_Enddate >= end_time then Urso_b4abn=1;
if vancomycin_Startdate <= start_time AND vancomycin_Startdate=. AND vancomycin_Enddate >= end_time then vancomycin_b4abn=1;
if EEN_Startdate <= start_time AND EEN_Startdate=. AND EEN_Enddate >= end_time then EEN_b4abn=1;
if thio_Startdate <= start_time AND thio_Startdate=. AND thio_Enddate >= end_time then thio_b4abn=1;
if steroids_Startdate <= start_time AND steroids_Startdate=. AND steroids_Enddate >= end_time then steroids_b4abn=1;
if Abx_Startdate <= start_time AND Abx_Startdate=. AND Abx_Enddate >= end_time then Abx_b4abn=1;
run;
data IBD_Meds4;
merge IBD_Meds3 IBD_Data15;
by StudyID;
if orig_coll_date < datedx then Date_Collected = DateDx;
if start_time=. then do; start_time=DateDx; end;
if end_time>Date_Collected OR end_time=. then do;
end_time=Date_Collected; end;
run;
data IBD_Meds5;
set IBD_Meds4;
if MP_b4abn=. then MP_b4abn=0;
if Adalimumab_b4abn=. then Adalimumab_b4abn=0;
if AZA_b4abn=. then AZA_b4abn=0;
if Budesonide_b4abn=. then Budesonide_b4abn=0;
if Certolizumab_b4abn=. then Certolizumab_b4abn=0;
if Cipro_b4abn=. then Cipro_b4abn=0;
if Golimumab_b4abn=. then Golimumab_b4abn=0;
if Hydrocort_b4abn=. then Hydrocort_b4abn=0;
if Infliximab_b4abn=. then Infliximab_b4abn=0;
if Mesalazine_b4abn=. then Mesalazine_b4abn=0;
if MTX_b4abn=. then MTX_b4abn=0;
if MethylPred_b4abn=. then MethylPred_b4abn=0;
if Metronidazole_b4abn=. then Metronidazole_b4abn=0;
if Pred_b4abn=. then Pred_b4abn=0;
if Sulfasalazine_b4abn=. then Sulfasalazine_b4abn=0;
if FK_b4abn=. then FK_b4abn=0;
if Urso_b4abn=. then Urso_b4abn=0;
if vancomycin_b4abn=. then vancomycin_b4abn=0;
if EEN_b4abn=. then EEN_b4abn=0;
if thio_b4abn=. then thio_b4abn=0;
if Biologics_b4abn=. then Biologics_b4abn=0;
if steroids_b4abn=. then steroids_b4abn=0;
if Abx_b4abn=. then Abx_b4abn=0;
run;
data IBD_Meds7;
  set IBD_Meds5;
  format start_time_h datetime20. end_time_h datetime20. datex $9. datey $9. ;
  datex = put(start_time, date9.); /* convert the date to a character string;  
  ddate = put(end_time, date9.); /* convert the date to a character string;  
  start_time_h = input(datex||"00:00:00",datetime20.); /* this input function reads a character value;  
  * the put function displays the SAS date into a readable date, but remember the date is still retained as a date value;  
  end_time_h = input(datey||"00:00:00",datetime20.); 
  put start_time_h / start_time_h datetime20.;  
  put end_time_h / end_time_h datetime20.;
run;
data IBD_Meds8;
set IBD_Meds7;
  if start_time_h = end_time_h then end_time_h=end_time_h+43200; /*12 hours later; 
  drop datex datey;
run;
data IBD_Meds9;
set IBD_Meds8;
  by StudyID;
  retain temp_start_time_h temp_end_time_h;
  format temp_start_time_h datetime20. temp_end_time_h datetime20.;
  if first.StudyID=1 then temp_start_time_h = start_time_h;
  if first.StudyID=1 then temp_end_time_h = end_time_h;
  if first.StudyID<>1 then do;
    if start_time_h=temp_start_time_h then 
    start_time_h=temp_end_time_h;
end;
    temp_start_time_h = start_time_h;
    temp_end_time_h = end_time_h;
run;
data IBD_Meds10;
    set IBD_Meds9;
    start_time=start_time_h;
    end_time=end_time_h;
    drop start_time_h end_time_h temp_start_time_h temp_end_time_h;
    format start_time end_time datetime20.;
    if start_time=end_time then end_time = end_time+3600;
run;
proc sort data=IBD_Meds10 out=IBD_Meds11;
    by StudyID start_time;
run;
/*proc print data=IBD_Meds11 (obs=50); Title "IBD & meds"; run;*/
data IBD_Meds12;
    set IBD_Meds11;
    if StudyID="101" AND Mesalazine_Startdate='04APR2002'd THEN Mesalazine_b4abn=0 AND steroids_b4abn=0;
    if StudyID="102" AND Abx_Startdate='07MAY2003'd THEN Abx_b4abn=0;
    if StudyID="103" AND Sulfasalazine_Startdate='12SEP2006'd THEN Sulfasalazine_b4abn=0;
    if StudyID="104" AND steroids_Startdate='20DEC2006'd THEN steroids_b4abn=0;
    if StudyID="105" AND Abx_Startdate='31OCT2007'd THEN Abx_b4abn=0;
    if StudyID="106" AND steroids_Startdate='20JUN2011'd THEN steroids_b4abn=0;
    if StudyID="107" AND Sulfasalazine_Startdate='29JUL2009'd THEN Sulfasalazine_b4abn=0;
    if StudyID="108" AND steroids_Startdate='04OCT2010'd THEN steroids_b4abn=0;
    if StudyID="109" AND Sulfasalazine_Startdate='22NOV2005'd THEN Sulfasalazine_b4abn=0;
    if StudyID="110" AND Sulfasalazine_Startdate='09NOV2010'd THEN Sulfasalazine_b4abn=0;
    if StudyID="111" AND steroids_Startdate='18SEP2009'd THEN steroids_b4abn=0;
    if StudyID="112" AND Sulfasalazine_Startdate='22APR2002'd THEN Sulfasalazine_b4abn=0;
    if StudyID="113" AND steroids_Startdate='03APR2006'd THEN steroids_b4abn=0;
    if StudyID="114" AND Sulfasalazine_Startdate='17JAN2003'd THEN Sulfasalazine_b4abn=0;
    if StudyID="115" AND Abx_Startdate='08NOV2004'd THEN Abx_b4abn=0;
    if StudyID="116" AND Abx_Startdate='06NOV2007'd THEN Abx_b4abn=0;
    if StudyID="117" AND steroids_Startdate='28JUN2010'd THEN steroids_b4abn=0;
    if StudyID="118" AND Sulfasalazine_Startdate='27FEB2009'd THEN Sulfasalazine_b4abn=0;
    if StudyID="119" AND Sulfasalazine_Startdate='15JUL2010'd THEN Sulfasalazine_b4abn=0;
if StudyID="1" AND Abx_Startdate='10NOV2009'd THEN
Abx_b4abn=0;
if StudyID="11" AND Sulfasalazine_Startdate='11MAY2010'd THEN
Sulfasalazine_b4abn=0;
if StudyID="12" AND steroids_Startdate='08OCT2010'd THEN
steroids_b4abn=0 AND Abx_b4abn=0;
if orig_coll_date<DateDx THEN
MP_b4abn=0;
if orig_coll_date<DateDx THEN
Adalimumab_b4abn=0;
if orig_coll_date<DateDx THEN
AZA_b4abn=0;
if orig_coll_date<DateDx THEN
Budesonide_b4abn=0;
if orig_coll_date<DateDx THEN
Certolizumab_b4abn=0;
if orig_coll_date<DateDx THEN
Cipro_b4abn=0;
if orig_coll_date<DateDx THEN
Golimumab_b4abn=0;
if orig_coll_date<DateDx THEN
Hydrocort_b4abn=0;
if orig_coll_date<DateDx THEN
Infliximab_b4abn=0;
if orig_coll_date<DateDx THEN
Mesalazine_b4abn=0;
if orig_coll_date<DateDx THEN
MTX_b4abn=0;
if orig_coll_date<DateDx THEN
MethylPred_b4abn=0;
if orig_coll_date<DateDx THEN
Metronidazole_b4abn=0;
if orig_coll_date<DateDx THEN
Pred_b4abn=0;
if orig_coll_date<DateDx THEN
Sulfasalazine_b4abn=0;
if orig_coll_date<DateDx THEN
FK_b4abn=0;
if orig_coll_date<DateDx THEN
vancomycin_b4abn=0;
if orig_coll_date<DateDx THEN
EEN_b4abn=0;
if orig_coll_date<DateDx THEN
steroids_b4abn=0;
if orig_coll_date<DateDx THEN
Abx_b4abn=0;
if orig_coll_date<DateDx THEN
Biologics_b4abn=0;
run;
/*proc print data=IBD_Meds_ADD_Dx_Interval2; Title "Missing interval between
diagnosis and 1st med"; run;*/
data IBD_Data;
  set IBD_Meds12;
  By StudyID;
  if last.StudyID=0 then Abn=50;
run;
/*********************************************************************
Appendix Table 5: Statistical Analysis – Description of cohort
/* *** Number of biochem measurements per patient *** */
data Num_Bioch_Meas1;
  set All_Liv_Biochem;
run;
proc sort data=Num_Bioch_Meas1 out=Num_Bioch_Meas2 NODUPKEY;
  by StudyID Date_Collected;
run;
data Num_Bioch_Meas3;
  set Num_Bioch_Meas2;
  count + 1;
  by StudyID;
  if first.StudyID then count = 1;
run;
data Num_Bioch_Meas4;
  set Num_Bioch_Meas3;
  keep count;
run
proc freq data=Num_Bioch_Meas4; Title "Number of biochemistry measurements/patient"; run;

data Num_Bioch_Meas5;
  set Num_Bioch_Meas3;
  by HSCNumber;
  if last.HSCNumber=0 then delete; /*the last biochemistry measurement should be last.HSCNumber=1*/
run;

/* Average # of biochemistry measurements*/
/* Test for normality: Shapiro-Wilk test b/c < 2000 variables*/
title "Test of normality: # Biochemistry measurements";
proc univariate data=Num_Bioch_Meas5 normal plots;
  var count;
  histogram;
  qqplot;
run;

/* # Biochemistry measurements is NOT normally distributed, Shapiro-Wilk test < 0.001; Median 11; 25% Q1 5, 75% Q3 21*/

/*STATS: Table 1, description of demographic data according to whether patients developed abnormal biochemistry or not */
/* TABLE 1: AGE*/
/* Test for normality: Shapiro-Wilk test b/c < 2000 variables*/
title "Test of normality: Age at diagnosis (yrs)";
proc univariate data=all_ptn_time_to_event_FU normal plots;
  var AgeDxYrs;
  histogram;
  qqplot;
run;

/* Age at diagnosis is NOT normally distributed, Shapiro-Wilk test < 0.001; Median 4442.000; 25% Q1 3449.0, 75% Q3 5227.0 */
/* Calculation of median age by abnormal/normal liver biochemistry */
/* AGE median p value: Wilcoxon-Rank Sum test performed for continuous, non-normally distributed variables*/
/* TABLE 1: SEX*/
/* Calculation of # Females by abnormal/normal liver biochemistry */
title "Proportion of FEMALE patients by abnormal/normal liver biochemistry";
proc freq data=all_ptn_time_to_event_FU;
  tables abn * sex;
run;
/* TABLE 1: IBD TYPE*/
/* Calculation of IBD types by abnormal/normal liver biochemistry */
title "Proportion of IBD types by sex ";
proc freq data=IBD_Data;
  tables sex * InitDx /chisq fisher;
run;
/* InitDx 1=CD, 2=UC 3=IBD-U*/
/* TABLE 1: CD location by Paris classification*/
/* L1: Distal 1/3 ileum ± limited cecal disease*/
/* L2: Colonic*/
/* L3: Ileocolonic*/
/* L4a: Upper disease, proximal to ligament of Treitz*/
/* L4b: Upper disease, proximal to ligament of Treitz + proximal to distal 1/3 ileum*/
data IBD_Data_class;
  set IBD_Data;
  if L1=1 AND (L4a="." AND L4b="." ) then class=1;
  if L1=1 AND L4a=1 then class=2;
if L1=1 AND L4b=1 then class=3;
if L2=1 AND (L4a = "." AND L4b = ".") then class=4;
if L2=1 AND L4a=1 then class=5;
if (L1= 1 AND L4b=1) then class=6;
if L3=1 AND (L4a = "." AND L4b = ".") then class=7;
if L3=1 AND L4a=1 then class=8;
if L3=1 AND L4b=1 then class=9;
if (L1=. AND L2=. AND L3=.) AND L4a=1 then class=10;
if (L1=. AND L2=. AND L3=.) AND L4b=1 then class=11;

run;
title "CD location by Paris classification";
proc freq data=IBD_Data_class;
        tables class * InitDx /chisq fisher;
run;
/* TABLE 1: UC location by Paris classification*/
/* E1: Ulcerative proctitis*/
/* E2: Left sided disease*/
/* E3: Extensive disease (proximal to the splenic flexure, but distal to
the hepatic flexure)*/
/* E4: Pancolitis*/
data IBD_Data_class_UC;
        set IBD_Data;
       if E1=1 then class=1;
if E2=1 then class=2;
if E3=1 then class=3;
if E4=1 then class=4;
if U_E1=1 then class=1;
if U_E2=1 then class=2;
if U_E3=1 then class=3;
if U_E4=1 then class=4;
if InitDx=2 AND (UpperGI="macro" OR Oral="macro" OR Esophagus="macro"
OR Stomach="macro" OR Duodenum="macro") then class_Upper=5;
run;
title "UC location by Paris classification";
proc freq data=IBD_Data_class_UC;
        tables class * InitDx /chisq fisher;
run;
title "UC location by Paris classification";
proc freq data=IBD_Data_class_UC;
        tables class_Upper * InitDx /chisq fisher;
run;
/* TABLE 1: PCDAI / PUCAI at Diagnosis*/
/* Test for normality: Shapiro-Wilk test b/c < 2000 variables*/
title "Test of normality: PCDAI for CD at diagnosis";
data IBD_Data_CD;
        set IBD_Data;
       if InitDx = 1;
run;
proc univariate data=IBD_Data_CD normal plots;
        var PCDAI;
        histogram;
        qqplot;
run;
/* PCDAI for CD at diagnosis is normally distributed, Shapiro-Wilk test
0.0896; Mean 36.4840764 Std Deviation 15.945663 */
title "Test of normality: mathPCDAI for CD at diagnosis";
proc univariate data=IBD_Data_CD normal plots;
        var mathPCDAI;
/* mathPCDAI for CD at diagnosis is normally distributed, Shapiro-Wilk test 0.1164; Mean 58.9 Std Deviation 26.0 */
title "Test of normality: PUCAI for UC at diagnosis"
data IBD_Data_UC;
  set IBD_Data;
  if InitDx = 2;
run;
proc univariate data=IBD_Data_UC normal plots;
  var PUCAI;
  histogram;
  qqplot;
run;
/* PUCAI for UC at diagnosis is NOT normally distributed, Shapiro-Wilk test 0.0018; Median 55.0, IQR 25:35, 75: 65 */
/* TABLE 1: Growth z-scores at diagnosis */
/*Percentile values found at: http://www.cdc.gov/nccdphp/dnpao/growthcharts/resources/sas.htm*/
/* Test for normality: Shapiro-Wilk test b/c < 2000 variables*/
title "Test of normality: Weight z-scores at diagnosis"
proc sort data=IBD_Data;
  by InitDx;
run;
proc univariate data=IBD_Data normal plots;
  var WAZ; /*Wt z score*/
  by InitDx;
  histogram;
  qqplot;
run;
/* Wt z score at diagnosis is normally distributed for CD & UC, Shapiro-Wilk test 0.4219 for CD, 0.0915 for UC */
/* Test for normality: Shapiro-Wilk test b/c < 2000 variables*/
title "Test of normality: Height z-scores at diagnosis"
proc univariate data=IBD_Data normal plots;
  var HAZ; /*Ht z score*/
  by InitDx;
  histogram;
  qqplot;
run;
/* Ht z score at diagnosis is NOT normally distributed for CD & UC, Shapiro-Wilk test 0.0271 for CD, 0.0246 for UC */
/* Test for normality: Shapiro-Wilk test b/c < 2000 variables*/
title "Test of normality: BMI z-scores at diagnosis"
proc univariate data=IBD_Data_OneRowPerPtn normal plots;
  var BMIZ; /*BMI z score*/
  by InitDx;
  histogram;
  qqplot;
run;
/* BMI z score at diagnosis is NOT normally distributed for CD & UC, Shapiro-Wilk test 0.0435 for CD, 0.0016 for UC */
/* TABLE 1: AVERAGE CRP */
/* Test for normality: Shapiro-Wilk test b/c < 2000 variables*/
title "Test of normality: CRP"
proc sort data=IBD_Data;
By InitDx;
run;
proc univariate data=IBD_Data normal plots;
   var CRP;
   by InitDx;
   histogram;
   qqplot;
run;
/* CRP is normally NOT distributed at all, Shapiro-Wilk test < 0.05 */
/* TABLE 1: EIM list */
Title "EIM frequencies"
run;
proc freq data=IBD_Data;
   tables AIH * InitDx / chisq fisher;
run;
proc freq data=IBD_Data;
   tables PSC * InitDx / chisq fisher;
run;
proc freq data=IBD_Data;
   tables ASC * InitDx / chisq fisher;
run;
proc freq data=IBD_Data;
   tables AnkSpond * InitDx / chisq fisher;
run;
proc freq data=IBD_Data;
   tables Arthralgia * InitDx / chisq fisher;
run;
proc freq data=IBD_Data;
   tables Asthma * InitDx / chisq fisher;
run;
proc freq data=IBD_Data;
   tables Thyroid * InitDx / chisq fisher;
run;
proc freq data=IBD_Data;
   tables Celiac * InitDx / chisq fisher;
run;
proc freq data=IBD_Data;
   tables EN * InitDx / chisq fisher;
run;
proc freq data=IBD_Data;
   tables Diabetes * InitDx / chisq fisher;
run;
proc freq data=IBD_Data;
   tables uveitis * InitDx / chisq fisher;
run;
proc freq data=IBD_Data;
   tables L_Arthritis * InitDx / chisq fisher;
run;
proc freq data=IBD_Data;
   tables S_Arthritis * InitDx / chisq fisher;
run;
proc freq data=IBD_Data;
   tables o_ulcers * InitDx / chisq fisher;
run;
proc freq data=IBD_Data;
   tables OFG * InitDx / chisq fisher;
run;
proc freq data=IBD_Data;
   tables Psoriasis * InitDx / chisq fisher;
run;
proc freq data=IBD_Data;
  tables PG * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
  tables SI * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
  tables Vitiligo * InitDx /chisq fisher;
run;
/* TABLE 1: Autoimmune Ab list*/
Title "Pathology Labs - present or absent in cohort";run;
proc freq data=IBD_Data;
  tables LabPath * InitDx /chisq fisher;
run;
Title "ANCA - present or absent in cohort";run;
proc freq data=IBD_Data;
  tables EverANCA * InitDx /chisq fisher;
run;
Title "ASCA - present or absent in cohort";run;
proc freq data=IBD_Data;
  tables EverASCA * InitDx /chisq fisher;
run;
/* TABLE 1: Family Hx of IBD 1st degree vs 2nd degree */
Title "Family Hx of IBD 1st degree vs 2nd degree";run;
proc freq data=IBD_Data;
  tables FamHxIBD * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
  tables FirstDegree * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
  tables SecondDegree * InitDx /chisq fisher;
run;
/* TABLE 1: Family Hx of autoimmune disease */
title "Family Hx of autoimmune disease";
proc freq data=IBD_Data;
  tables Atopy_fam * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
  tables AIH_fam * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
  tables Diabetes_fam * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
  tables JIA_fam * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
  tables SLE_fam * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
  tables alopecia_fam * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
  tables Arthritis_fam * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
tables Celiac_fam * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
tables Thyroid_fam * InitDx /chisq fisher;
run; /* TABLE 1: Medication use by diagnosis */
title "Medication use by diagnosis ";
proc freq data=IBD_Data;
tables MP * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
tables Adalimumab * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
tables AZA * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
tables Budesonide * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
tables Certolizumab * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
tables Golimumab * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
tables Hydrocort * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
tables Infliximab * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
tables Mesalazine * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
tables MTX * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
tables MethylPred * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
tables Metronidazole * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
tables Cipro * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
tables Pred * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
tables Sulfasalazine * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
tables FK * InitDx /chisq fisher;
run;
proc freq data=IBD_Data;
tables Urso * InitDx /chisq fisher;
run;
Appendix Table 6: Statistical Analysis – Description of liver enzymes abnormalities

data whichLFTabn1;
   set All_Liv_Biochem;
   if test = "Bilirubin-Conjugated,blood" then delete;
   if test = "AST,blood" then delete;
run;
data whichLFTabn2;
   set First_Abn_liver_test;
   if test = "Bilirubin-Conjugated,blood" then delete;
   if test = "AST,blood" then delete;
   if ULN = "." then delete;
   keep StudyID Date_Collected;
run;
data whichLFTabn3;
   set whichLFTabn2;
   Date_1st_abn = Date_Collected;
   format Date_1st_abn date9.;
run;
proc sort data = whichLFTabn1;
   by StudyID;
run;
proc sort data = whichLFTabn3;
   by StudyID;
run;
data whichLFTabn4;
   merge whichLFTabn1 whichLFTabn3;
   by StudyID;
run;
data whichLFTabn5;
   set whichLFTabn4;
   if Date_Collected = Date_1st_abn;
run;
data whichLFTabn6;
   set whichLFTabn5;
   if Result > ULN;
run;
proc freq data = whichLFTabn6;
run;
/*Test */
/*Test - Frequency - Percent - Cumulative Frequency - Cumulative Percent*/
/*ALT,blood 119 72.12 119 72.12 */
/*GGT,blood 46 27.88 165 100.00 */
*******************************************************************************/

Appendix Table 7: Statistical Analysis – Time to Event analysis
/* Survival Analysis – Kaplan-Meier estimate of time to abnormal biochemistry */
/*first abn result*/
data IBD_Data_FORM; /*format time variable to months*/
set IBD_Data;
Event_time = abn_censor_duration/365*12;
if abn = 50 then delete;
run;
ods graphics on;
Title "Time to event analysis for abnormal liver biochemistry";
proc lifetest data=IBD_Data_FORM nelson plots=(survival(atrisk) logsurv);
/*   BY StudyID;*/
  time Event_time*Abn(0);
run;
ods graphics off;
ods rtf close;
/*Stratify by type of IBD*/
ods graphics on;
Title "Time to event analysis for abnormal liver biochemistry according to IBD diagnosis";
proc lifetest data=IBD_Data_FORM nelson plots=(survival(atrisk) logsurv);
  time Event_time*Abn(0);
  strata InitDx;
run;
ods graphics off;
ods rtf close;
/*Time to event - all patients EXCEPT those with chronic liver disease*/
/*first abn result*/
data IBD_Data_FORM_NoLiver; /*format time variable to months*/
  set IBD_DataNoLiver;
  Event_time = abn_censor_duration/365*12;
run;
ods graphics on;
Title "Time to event analysis for abnormal liver biochemistry";
proc lifetest data=IBD_Data_FORM_NoLiver nelson plots=(survival(atrisk) logsurv);
  time Event_time*Abn(0);
run;
ods graphics off;
ods rtf close;
/*time to 2x ULN*/
data IBD_Data_2ULN_FORM; /*format time variable to months*/
  set IBD_Data_2ULN ;
  Event_time = abn_censor_duration_2/365*12;
if Abn_2ULN=50 then delete;
run;
ods graphics on;
Title "Time to event analysis for 2x ULN abnormal liver biochemistry";
proc lifetest data=IBD_Data_2ULN_FORM nelson plots=(survival(atrisk) logsurv);
  time Event_time*Abn_2ULN(0);
run;
ods graphics off;
ods rtf close;
/*Stratify by type of IBD*/
ods graphics on;
Title "Time to event analysis for 2xULN abnormal liver biochemistry according to IBD diagnosis";
proc lifetest data=IBD_Data_2ULN_FORM nelson plots=(survival(atrisk) logsurv);
  time Event_time*Abn_2ULN(0);
  strata InitDx;
/*time to persistently elevated abn biochem 30-90 days - 1-2x ULN*/
data IBD_Data_30_90_FORM; /*format time variable to months*/
set IBD_Data_30_90 ;
   Event_time = abn_censor_duration_pers/365*12;
run;
ods graphics on;
Title "Time to event analysis for Persistence of abnormal liver biochemistry for 30-90 days";
proc lifetest data=IBD_Data_30_90_FORM nelson plots=(survival(atrisk) logsurv);
   time Event_time*Abn_pers90(0);
run;
ods graphics close;
ods rtf close;
/*Stratify by type of IBD*/
data IBD_Data2_30_90_FORM; /*format time variable to months*/
set IBD_Data2_30_90 ;
   Event_time = abn_censor_duration_pers/365*12;
run;
ods graphics on;
Title "Time to event analysis for Persistence of abnormal liver biochemistry for 30-90 days according to IBD diagnosis";
proc lifetest data=IBD_Data2_30_90_FORM nelson plots=(survival(atrisk) logsurv);
   time Event_time*Abn_pers90(0);
   strata InitDx;
run;
ods graphics close;
ods rtf close;
/*time to persistently elevated abn biochem 30-90 days - >=2x ULN*/
/*Stratify by type of IBD*/
/*--- Cox-Proportional Hazards Model ---*/
/*--- UNIVARIATE ANALYSIS -- ANY ABN BIOCHEMISTRY > ULN---*/

Appendix Table 8: Statistical Analysis – Cox Proportional Hazards modeling
/*----- Cox-Proportional Hazards Model --------------------------*/
/*----- UNIVARIATE ANALYSIS -- ANY ABN BIOCHEMISTRY > ULN-----*/
data IBD_Data_CD;
set IBD_Data;
if InitDx = 1;
if PCDAI >= 30 then PCDAI_corr = 1;
if PCDAI < 30 then PCDAI_corr = 0;
run;
data IBD_Data_UC;
set IBD_Data;
if InitDx = 2;
if PUCAI >= 30 then PUCAI_corr = 1;
if PUCAI < 30 then PUCAI_corr = 0;
run;
data IBD_Data_IBDU;
set IBD_Data;
if InitDx = 3;
if PCDAI >= 30 then PCDAI_corr = 1;
if PCDAI < 30 then PCDAI_corr = 0;
if PUCAI >= 30 then PUCAI_corr = 1;
if PUCAI < 30 then PUCAI_corr = 0;
run;
Title "Cox-Proportional Hazards Model - UNIVARIATE: InitDx against the development of abn biochem - with dataset treated time varying covariate";
proc phreg data = IBD_Data covsandwich(aggregate);
    ID StudyID;
    class InitDx (ref="1");
    model (start_time, end_time)*Abn(0, 50) = InitDx / RL ties=efron;
run;
Title "Cox-Proportional Hazards Model - UNIVARIATE: AgeDxYrs against the development of abn biochem - with dataset treated time varying covariate";
proc phreg data = IBD_Data covsandwich(aggregate);
    ID StudyID;
    model (start_time, end_time)*Abn(0, 50) = AgeDxYrs / RL ties=efron;
run;
Title "Cox-Proportional Hazards Model - UNIVARIATE: Sex against the development of abn biochem - with dataset treated time varying covariate";
proc phreg data = IBD_Data covsandwich(aggregate);
    ID StudyID;
    class Sex (ref="1");
    model (start_time, end_time)*Abn(0, 50) = Sex / RL ties=efron;
run;
Title "Cox-Proportional Hazards Model - UNIVARIATE: PCDAI against the development of abn biochem - with dataset treated time varying covariate";
proc phreg data = IBD_Data_CD covsandwich(aggregate);
    ID StudyID;
    model (start_time, end_time)*Abn(0, 50) = PCDAI / RL ties=efron;
run;
Title "Cox-Proportional Hazards Model - UNIVARIATE: PUCAI against the development of abn biochem - with dataset treated time varying covariate";
proc phreg data = IBD_Data_UC covsandwich(aggregate);
    ID StudyID;
    model (start_time, end_time)*Abn(0, 50) = PUCAI / RL ties=efron;
run;
Title "Cox-Proportional Hazards Model - UNIVARIATE: PCDAI against the development of abn biochem - with dataset treated time varying covariate";
proc phreg data = IBD_Data_IBDU covsandwich(aggregate);
    ID StudyID;
    model (start_time, end_time)*Abn(0, 50) = PCDAI / RL ties=efron;
run;
Title "Cox-Proportional Hazards Model - UNIVARIATE: PUCAI against the development of abn biochem - with dataset treated time varying covariate"

proc phreg data = IBD_Data_IBDU covsandwich(aggregate);
   ID StudyID;
   model (start_time, end_time)*Abn(0, 50) = PUCAI / RL ties=efron;
run;

Title "Cox-Proportional Hazards Model - UNIVARIATE: True_EIM against the development of abn biochem - with dataset treated time varying covariate"

proc phreg data = IBD_Data covsandwich(aggregate);
   ID StudyID;
   class True_EIM (ref="0");
   model (start_time, end_time)*Abn(0, 50) = True_EIM / RL ties=efron;
run;

Title "Cox-Proportional Hazards Model - UNIVARIATE: Past_AutoImm against the development of abn biochem - with dataset treated time varying covariate"

proc phreg data = IBD_Data covsandwich(aggregate);
   ID StudyID;
   class Past_AutoImm (ref="0");
   model (start_time, end_time)*Abn(0, 50) = Past_AutoImm / RL ties=efron;
run;

Title "Cox-Proportional Hazards Model - UNIVARIATE: FamHxIBD against the development of abn biochem - with dataset treated time varying covariate"

proc phreg data = IBD_Data covsandwich(aggregate);
   ID StudyID;
   class FamHxIBD (ref="0");
   model (start_time, end_time)*Abn(0, 50) = FamHxIBD / RL ties=efron;
run;

Title "Cox-Proportional Hazards Model - UNIVARIATE: FamHx_AutoImm against the development of abn biochem - with dataset treated time varying covariate"

proc phreg data = IBD_Data covsandwich(aggregate);
   ID StudyID;
   class FamHx_AutoImm (ref="0");
   model (start_time, end_time)*Abn(0, 50) = FamHx_AutoImm / RL ties=efron;
run;

Title "Cox-Proportional Hazards Model - UNIVARIATE: WAZ against the development of abn biochem - with dataset treated time varying covariate"

proc phreg data = IBD_Data covsandwich(aggregate);
   ID StudyID;
   model (start_time, end_time)*Abn(0, 50) = WAZ / RL ties=efron;
run;

Title "Cox-Proportional Hazards Model - UNIVARIATE: HAZ against the development of abn biochem - with dataset treated time varying covariate"

proc phreg data = IBD_Data covsandwich(aggregate);
   ID StudyID;
   model (start_time, end_time)*Abn(0, 50) = HAZ / RL ties=efron;
run;

Title "Cox-Proportional Hazards Model - UNIVARIATE: BMI_Z against the development of abn biochem - with dataset treated time varying covariate"

proc phreg data = IBD_Data covsandwich(aggregate);
   ID StudyID;
   model (start_time, end_time)*Abn(0, 50) = BMI_Z / RL ties=efron;
run;

Title "Cox-Proportional Hazards Model - UNIVARIATE: LabPATH against the development of abn biochem - with dataset treated time varying covariate"

proc phreg data = IBD_Data covsandwich(aggregate);
   ID StudyID;
   class LabPATH (ref="0");
model (start_time, end_time)*Abn(0, 50) = LabPATH / RL ties=efron;
run;

Title "Cox-Proportional Hazards Model - UNIVARIATE: EverANCA against the development of abn biochem - with dataset treated time varying covariate";
proc phreg data = IBD_Data covsandwich(aggregate);
   ID StudyID;
   class EverANCA (ref="0");
   model (start_time, end_time)*Abn(0, 50) = EverANCA / RL ties=efron;
run;

Title "Cox-Proportional Hazards Model - UNIVARIATE: EverASCA against the development of abn biochem - with dataset treated time varying covariate";
proc phreg data = IBD_Data covsandwich(aggregate);
   ID StudyID;
   class EverASCA (ref="0");
   model (start_time, end_time)*Abn(0, 50) = EverASCA / RL ties=efron;
run;

Title "Cox-Proportional Hazards Model - UNIVARIATE: Biologics_b4abn against the development of abn biochem - with dataset treated time varying covariate";
proc phreg data = IBD_Data covsandwich(aggregate);
   ID StudyID;
   class Biologics_b4abn (ref="0");
   model (start_time, end_time)*Abn(0, 50) = Biologics_b4abn / RL ties=efron;
run;

Title "Cox-Proportional Hazards Model - UNIVARIATE: Infliximab_b4abn against the development of abn biochem - with dataset treated time varying covariate";
proc phreg data = IBD_Data covsandwich(aggregate);
   ID StudyID;
   class Infliximab_b4abn (ref="0");
   model (start_time, end_time)*Abn(0, 50) = Infliximab_b4abn / RL ties=efron;
run;

Title "Cox-Proportional Hazards Model - UNIVARIATE: Adalimumab_b4abn against the development of abn biochem - with dataset treated time varying covariate";
proc phreg data = IBD_Data covsandwich(aggregate);
   ID StudyID;
   class Adalimumab_b4abn (ref="0");
   model (start_time, end_time)*Abn(0, 50) = Adalimumab_b4abn / RL ties=efron;
run;

Title "Cox-Proportional Hazards Model - UNIVARIATE: MTX_b4abn against the development of abn biochem - with dataset treated time varying covariate";
proc phreg data = IBD_Data covsandwich(aggregate);
   ID StudyID;
   class MTX_b4abn (ref="0");
   model (start_time, end_time)*Abn(0, 50) = MTX_b4abn / RL ties=efron;
run;

Title "Cox-Proportional Hazards Model - UNIVARIATE: Thio_b4abn against the development of abn biochem - with dataset treated time varying covariate";
proc phreg data = IBD_Data covsandwich(aggregate);
   ID StudyID;
   class Thio_b4abn (ref="0");
   model (start_time, end_time)*Abn(0, 50) = Thio_b4abn / RL ties=efron;
run;
Title "Cox-Proportional Hazards Model - UNIVARIATE: Steroids_b4abn against the development of abn biochem - with dataset treated time varying covariate";
proc phreg data = IBD_Data covsandwich(aggregate);
  ID StudyID;
  class Steroids_b4abn (ref="0");
  model (start_time, end_time)*Abn(0, 50) = Steroids_b4abn / RL ties=efron;
run;
Title "Cox-Proportional Hazards Model - UNIVARIATE: Abx_b4abn against the development of abn biochem - with dataset treated time varying covariate";
proc phreg data = IBD_Data covsandwich(aggregate);
  ID StudyID;
  class Abx_b4abn (ref="0");
  model (start_time, end_time)*Abn(0, 50) = Abx_b4abn / RL ties=efron;
run;
Title "Cox-Proportional Hazards Model - UNIVARIATE: Sulfasalazine_b4abn against the development of abn biochem - with dataset treated time varying covariate";
proc phreg data = IBD_Data covsandwich(aggregate);
  ID StudyID;
  class Sulfasalazine_b4abn (ref="0");
  model (start_time, end_time)*Abn(0, 50) = Sulfasalazine_b4abn / RL ties=efron;
run;
Title "Cox-Proportional Hazards Model - UNIVARIATE: Mesalazine_b4abn against the development of abn biochem - with dataset treated time varying covariate";
proc phreg data = IBD_Data covsandwich(aggregate);
  ID StudyID;
  class Mesalazine_b4abn (ref="0");
  model (start_time, end_time)*Abn(0, 50) = Mesalazine_b4abn / RL ties=efron;
run;
Title "Cox-Proportional Hazards Model - UNIVARIATE: EEN_b4abn against the development of abn biochem - with dataset treated time varying covariate";
proc phreg data = IBD_Data covsandwich(aggregate);
  ID StudyID;
  class EEN_b4abn (ref="0");
  model (start_time, end_time)*Abn(0, 50) = EEN_b4abn / RL ties=efron;
run;
Title "Cox-Proportional Hazards Model - Multivariable model - with dataset treated time varying covariate";
proc phreg data = IBD_Data covsandwich(aggregate);
  ID StudyID;
  class InitDx (ref="1");
  model (start_time, end_time)*Abn(0, 50) = InitDx AgeDxYrs Sex Thio_b4abn Steroids_b4abn Abx_b4abn Sulfasalazine_b4abn Mesalazine_b4abn EEN_b4abn/ RL ties=efron;
run;
/*Excluded: FamHxIBD due to high p values > 1.0*/
/*Testing assumptions: Model NOT overspecified -> 9 df in a subset of 300 patients with 119 events*/
/*Testing assumptions: Verifying multicollinearity */
Title "Cox-Proportional Hazards Model: testing assumptions, Verifying multicollinearity"

```plaintext
proc reg data = IBD_Data;
   model Abn = InitDx AgeDxYrs Sex Thio_b4abn Steroids_b4abn Abx_b4abn Sulfasalazine_b4abn Mesalazine_b4abn EEN_b4abn / TOL collin vif;
run;

/* Variance Inflation (VIF) < 10 (and mostly <1.2) AND condition index < 12 (mostly < 2.9), TOLERANCE VALUES ARE > 0.25 (AND MOSTLY > 0.8)...therefore variables are not collinear */
```

| Variable        | DF | Parameter Estimate | Standard Error | t Value | Pr > |t| Tolerance | Variance Inflation |
|-----------------|----|--------------------|----------------|---------|-------|-----------|-------------------|
| Intercept       | 1  | -4.052             | 0.97514        | -4.16   | <.0001| 0.79717   | 1.25443           |
| InitDx          | 1  | -0.6008            | 0.13505        | -4.45   | <.0001| 0.82623   | 1.21032           |
| AgeDxYrs        | 1  | 0.32622            | 0.87948        | 0.37    | 0.7107| 0.98001   | 1.0204            |
| Sex             | 1  | 1.41832            | 1.26128        | 1.12    | 0.261 | 0.96951   | 1.03145           |
| Thio_b4abn      | 1  | 10.8081            | 0.89404        | 12.09   | <.0001| 0.9235    | 1.08284           |
| Steroids_b4abn  | 1  | 2.83754            | 0.99388        | 2.86    | 0.0044| 0.93323   | 1.07155           |
| Abx_b4abn       | 1  | -0.1049            | 1.14166        | -0.09   | 0.9268| 0.83129   | 1.20295           |
| Sulfasalazine_b4abn | 1 | -0.1298            | 1.19728        | -0.11   | 0.9137| 0.91635   | 1.09129           |
| Mesalazine_b4abn| 1  | -1.1458            | 2.04124        | -0.56   | 0.5747| 0.94655   | 1.05646           |
| EEN_b4abn       | 1  | -4.052             | 0.97514        | -4.16   | <.0001| 0.79717   | 1.25443           |

Collinearity Diagnostics

<table>
<thead>
<tr>
<th>#</th>
<th>Eigen value</th>
<th>Condition Index</th>
<th>Inte</th>
<th>Init Dx</th>
<th>Age Dx Yrs</th>
<th>Sex</th>
<th>Fam Hx IBD</th>
<th>steroids_b4abn</th>
<th>Abx_b4abn</th>
<th>Sulfasalazine_b4abn</th>
<th>Mesalazine_b4abn</th>
<th>EEN_b4abn</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.27363</td>
<td>1</td>
<td>0.00</td>
<td>0.00</td>
<td>0.0024</td>
<td>0.003</td>
<td>0.00539</td>
<td>0.00989</td>
<td>0.0084</td>
<td>0.00587</td>
<td>0.0057</td>
<td>0.0017</td>
</tr>
<tr>
<td>2</td>
<td>1.03111</td>
<td>2.26153</td>
<td>9.2E</td>
<td>0.00</td>
<td>0.0007</td>
<td>1.4E</td>
<td>0.07303</td>
<td>0.00509</td>
<td>0.0151</td>
<td>0.04374</td>
<td>0.15134</td>
<td>0.52088</td>
</tr>
<tr>
<td>3</td>
<td>0.96671</td>
<td>2.33564</td>
<td>4.43</td>
<td>0.00</td>
<td>0.0006</td>
<td>3.1E</td>
<td>0.10999</td>
<td>0.01026</td>
<td>0.0195</td>
<td>0.2796</td>
<td>0.24566</td>
<td>0.18233</td>
</tr>
<tr>
<td>4</td>
<td>0.86609</td>
<td>2.46759</td>
<td>8E</td>
<td>0.00</td>
<td>0.0001</td>
<td>0.000</td>
<td>0.71991</td>
<td>0.00118</td>
<td>0.0710</td>
<td>0.1144</td>
<td>0.02104</td>
<td>0.12186</td>
</tr>
<tr>
<td>5</td>
<td>0.74261</td>
<td>2.66487</td>
<td>6.9E</td>
<td>0.00</td>
<td>0.0000</td>
<td>0.000</td>
<td>0.03657</td>
<td>0.02573</td>
<td>0.4387</td>
<td>0.14642</td>
<td>0.22497</td>
<td>0.03916</td>
</tr>
<tr>
<td>6</td>
<td>0.5398</td>
<td>3.12563</td>
<td>0.00</td>
<td>0.00</td>
<td>0.0004</td>
<td>0.003</td>
<td>0.1258</td>
<td>0.14494</td>
<td>0.3734</td>
<td>0.283</td>
<td>0.18602</td>
<td>7.9E-06</td>
</tr>
</tbody>
</table>
### Collinearity Diagnostics

<table>
<thead>
<tr>
<th>#</th>
<th>Eigenvalue</th>
<th>Condition Index</th>
<th>Proportion of Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Intercept</td>
<td>InitDx</td>
</tr>
<tr>
<td>0.36991</td>
<td>3.77578</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.08652</td>
<td>7.80738</td>
<td>0.00</td>
<td>0.53</td>
</tr>
<tr>
<td>0.02098</td>
<td>15.8558</td>
<td>0.98</td>
<td>0.42</td>
</tr>
</tbody>
</table>

/* ----- MULTIVARIABLE COX MODEL - ABN BIOCHEMISTRY > 2x ULN ---------*/
Title "Cox-Proportional Hazards Model - UNIVARIATE: InitDx against the development of Abn_2ULN biochem - with dataset treated time varying covariate";

```plaintext
proc phreg data = IBD_Data_2ULN covsandwich(aggregate);
  ID StudyID;
  class InitDx (ref="1");
  model (start_time, end_time)*Abn_2ULN(0, 50) = InitDx True_EIM BMIZ Steroids_b4abn Abx_b4abn Sulfasalazine_b4abn Mesalazine_b4abn EEN_b4abn/
    RL ties=efron;
run;
/*Excluded: Past_AutoImm removed because p value too high > 1.0*/

/*Testing assumptions: Model NOT overspecified -> 8 predictors in a subset of 300 patients with 58 events */

/*Testing assumptions: Verifying multicollinearity */
Title "Cox-Proportional Hazards Model: testing assumptions #5, Verifying multicollinearity";

```plaintext
proc reg data = IBD_Data_2ULN;
  model Abn_2ULN = InitDx True_EIM BMIZ Steroids_b4abn Abx_b4abn Sulfasalazine_b4abn Mesalazine_b4abn EEN_b4abn /
    TOL collin vif;
run;
/* Variance Inflation (VIF) < 10 (and mostly < 1.2) AND condition index < 12 (mostly < 9), TOLERANCE VALUES ARE > 0.25 (AND MOSTLY > 0.77)...therefore variables are not collinear */

### Parameter Estimates

<table>
<thead>
<tr>
<th>Variable</th>
<th>DF</th>
<th>Parameter Estimate</th>
<th>Standard Error</th>
<th>t Value</th>
<th>Pr &gt;</th>
<th>Tolerance</th>
<th>Variance Inflation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td>0.08096</td>
<td>0.0269</td>
<td>3.01</td>
<td>0.0026</td>
<td></td>
<td></td>
</tr>
<tr>
<td>InitDx</td>
<td>1</td>
<td>0.04041</td>
<td>0.01678</td>
<td>2.41</td>
<td>0.0161</td>
<td>0.77386</td>
<td>1.29222</td>
</tr>
<tr>
<td>True_EIM</td>
<td>1</td>
<td>-0.1056</td>
<td>0.02176</td>
<td>-4.85</td>
<td>&lt;0.0001</td>
<td>0.91735</td>
<td>1.09009</td>
</tr>
<tr>
<td>BMIZ</td>
<td>1</td>
<td>-0.0314</td>
<td>0.00499</td>
<td>-6.29</td>
<td>&lt;0.0001</td>
<td>0.96591</td>
<td>1.0353</td>
</tr>
</tbody>
</table>
Parameter Estimates

<table>
<thead>
<tr>
<th>Variable</th>
<th>DF</th>
<th>Parameter Estimate</th>
<th>Standard Error</th>
<th>t Value</th>
<th>Pr &gt;</th>
<th>Tolerance</th>
<th>Variance Inflation</th>
</tr>
</thead>
<tbody>
<tr>
<td>steroids_b4abn</td>
<td>1</td>
<td>-0.0107</td>
<td>0.01509</td>
<td>-0.71</td>
<td>0.4771</td>
<td>0.92215</td>
<td>1.08442</td>
</tr>
<tr>
<td>Abx_b4abn</td>
<td>1</td>
<td>0.07584</td>
<td>0.01607</td>
<td>4.72</td>
<td>&lt;.0001</td>
<td>0.945</td>
<td>1.0582</td>
</tr>
<tr>
<td>Sulfasalazine_b4abn</td>
<td>1</td>
<td>-0.095</td>
<td>0.0187</td>
<td>-5.08</td>
<td>&lt;.0001</td>
<td>0.89205</td>
<td>1.12101</td>
</tr>
<tr>
<td>EEN_b4abn</td>
<td>1</td>
<td>0.00274</td>
<td>0.03235</td>
<td>0.08</td>
<td>0.9326</td>
<td>0.93812</td>
<td>1.06596</td>
</tr>
</tbody>
</table>

Collinearity Diagnostics

<table>
<thead>
<tr>
<th>#</th>
<th>Eigenvalue</th>
<th>Condition Index</th>
<th>Intercept</th>
<th>InitDx</th>
<th>Age Dx Yrs</th>
<th>True_EIM</th>
<th>Steroids_b4abn</th>
<th>Abx_b4abn</th>
<th>Mesalazine_b4abn</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.6985</td>
<td>3</td>
<td>0.00482</td>
<td>0.00527</td>
<td>0.01026</td>
<td>0.01009</td>
<td>0.02131</td>
<td>0.01918</td>
<td>0.01529</td>
</tr>
<tr>
<td>2</td>
<td>1.0189</td>
<td>1</td>
<td>3.2E-05</td>
<td>0.00053</td>
<td>0.00936</td>
<td>0.15377</td>
<td>0.00456</td>
<td>2E-05</td>
<td>0.03953</td>
</tr>
<tr>
<td>3</td>
<td>0.9862</td>
<td>7</td>
<td>7.4E-05</td>
<td>1.26E-09</td>
<td>0.10277</td>
<td>0.15377</td>
<td>0.00456</td>
<td>2E-05</td>
<td>0.15488</td>
</tr>
<tr>
<td>4</td>
<td>0.9002</td>
<td>2</td>
<td>2.02696</td>
<td>2.7E-05</td>
<td>0.26215</td>
<td>0.44244</td>
<td>0.0124</td>
<td>0.00025</td>
<td>0.01075</td>
</tr>
<tr>
<td>5</td>
<td>0.8351</td>
<td>5</td>
<td>2.10441</td>
<td>8.2E-05</td>
<td>0.00282</td>
<td>0.44828</td>
<td>0.22121</td>
<td>7.3E-05</td>
<td>0.23589</td>
</tr>
<tr>
<td>6</td>
<td>0.7251</td>
<td>3</td>
<td>2.25843</td>
<td>2.4E-05</td>
<td>0.00119</td>
<td>0.00784</td>
<td>0.05976</td>
<td>0.01208</td>
<td>0.56523</td>
</tr>
<tr>
<td>7</td>
<td>0.5052</td>
<td>2</td>
<td>2.70566</td>
<td>0.00186</td>
<td>0.00542</td>
<td>0.03248</td>
<td>0.08014</td>
<td>0.35296</td>
<td>0.27931</td>
</tr>
<tr>
<td>8</td>
<td>0.2893</td>
<td>9</td>
<td>3.57499</td>
<td>0.05075</td>
<td>0.07083</td>
<td>0.00713</td>
<td>0.00113</td>
<td>0.58776</td>
<td>0.0238</td>
</tr>
<tr>
<td>9</td>
<td>0.0411</td>
<td>9</td>
<td>9.47552</td>
<td>0.94227</td>
<td>0.91391</td>
<td>0.11973</td>
<td>0.03147</td>
<td>0.00097</td>
<td>0.09595</td>
</tr>
</tbody>
</table>

/*----- MULTIVARIABLE COX MODEL - PERSISTENTLY ABN BIOCHEM 30-90 DAYS 1-2x ULN------*/
Title "Cox-Proportional Hazards Model - MULTIVARIABLE MODEL: PERSISTENTLY ABN BIOCHEM 30-90 DAYS 1-2x ULN - with dataset treated time varying covariate";
proc phreg data = IBD_Data_30_90 covarianc(aggregate);
   ID StudyID;
   class InitDx (ref=1);
   model (start_time, end_time)*Abn_pers90(0, 50) = InitDx AgeDxYrs
         Biologics_b4abn Steroids_b4abn Abx_b4abn Mesalazine_b4abn/ RL ties=efron;
run;
/*ANCA & PUCAI: removed because too many missing variables*/
/* Testing assumptions: Model overspecified by two VAR -> 6 predictors in a subset of 300 patients with 37 events*/
/*Testing assumptions: Verifying multicollinearity */
Title "Cox-Proportional Hazards Model: testing assumptions, Verifying multicollinearity";
proc reg data = IBD_Data_30_90;
   model Abn_pers90 = InitDx AgeDxYrs Biologics_b4abn Steroids_b4abn Abx_b4abn Mesalazine_b4abn / TOL collin vif;
run;
/* Variance Inflation (VIF) < 10 (and mostly <1.1) AND condition index < 12, TOLERANCE VALUES ARE > 0.25 (AND MOSTLY > 0.87)...therefore variables are not collinear */

Parameter Estimates

| Variable              | DF | Parameter Estimate | Standard Error | t Value | Pr > |t| Tolerance | Variance Inflation |
|-----------------------|----|--------------------|----------------|---------|------|----------|-------------------|
| Intercept             | 1  | 47.718             | 1.78663        | 26.71   | <.0001|         | 0                 |
| InitDx                | 1  | -2.7397            | 0.75499        | -3.63   | 0.0003| 0.87226  | 1.14644           |
| AgeDxYrs              | 1  | -0.458             | 0.10512        | -4.36   | <.0001| 0.94205  | 1.06151           |
| Biologics_b4abn       | 1  | -4.0356            | 1.00231        | -4.03   | <.0001| 0.97255  | 1.02823           |
| Steroids_b4abn        | 1  | 8.25337            | 0.71019        | 11.62   | <.0001| 0.9596   | 1.04607           |
| Abx_b4abn             | 1  | 3.78564            | 0.77387        | 4.89    | <.0001| 0.95428  | 1.04791           |
| Mesalazine_b4abn      | 1  | -1.4243            | 0.93901        | -1.52   | 0.1294| 0.95917  | 1.04257           |

Collinearity Diagnostics

<table>
<thead>
<tr>
<th>#</th>
<th>Eigenvalue</th>
<th>Condition Index</th>
<th>Intercept</th>
<th>InitDx</th>
<th>Age DxYrs</th>
<th>Biologics_b4abn</th>
<th>Steroids_b4abn</th>
<th>Abx_b4abn</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.15942</td>
<td>1</td>
<td>0.00205</td>
<td>0.00476</td>
<td>0.00426</td>
<td>0.00896</td>
<td>0.01717</td>
<td>0.01581</td>
</tr>
<tr>
<td>2</td>
<td>0.95535</td>
<td>2.08658</td>
<td>2.55E-07</td>
<td>0.00035</td>
<td>9.05E-08</td>
<td>0.48603</td>
<td>9.9E-05</td>
<td>0.00286</td>
</tr>
<tr>
<td>3</td>
<td>0.74932</td>
<td>2.35604</td>
<td>4.4E-06</td>
<td>4.2E-05</td>
<td>3.6E-05</td>
<td>0.36419</td>
<td>0.04471</td>
<td>0.22213</td>
</tr>
<tr>
<td>4</td>
<td>0.62531</td>
<td>2.57911</td>
<td>0.00102</td>
<td>0.00849</td>
<td>0.00119</td>
<td>0.05037</td>
<td>0.11746</td>
<td>0.67178</td>
</tr>
<tr>
<td>5</td>
<td>0.37353</td>
<td>3.33699</td>
<td>0.00939</td>
<td>0.01909</td>
<td>0.02446</td>
<td>0.06724</td>
<td>0.80311</td>
<td>0.01119</td>
</tr>
<tr>
<td>6</td>
<td>0.11169</td>
<td>6.10265</td>
<td>0.00034</td>
<td>0.43196</td>
<td>0.34958</td>
<td>0.00752</td>
<td>0.01728</td>
<td>0.05416</td>
</tr>
<tr>
<td>7</td>
<td>0.02539</td>
<td>12.8004</td>
<td>0.98719</td>
<td>0.53531</td>
<td>0.62047</td>
<td>0.01569</td>
<td>0.00018</td>
<td>0.02207</td>
</tr>
</tbody>
</table>

/*----- MULTIVARIABLE COX MODEL – PERSISTENTLY ABN BIOCHEM 30-90 DAYS >2x ULN--------*/
Title "Cox-Proportional Hazards Model – MULTIVARIABLE MODEL: PERSISTENTLY ABN BIOCHEM 30-90 DAYS >2x ULN – with dataset treated time varying covariate";
proc phreg data = IBD_Data2_30_90 covsandwich(aggregate);
   ID StudyID;
   class InitDx (ref="1");
   model (start_time, end_time)*Abn_pers90(0, 50) = InitDx FamHxIBD Steroids_b4abn Abx_b4abn/ RL ties=efron;
run;
/*EverANCA & autimm ab & PUCAI : removed – too many missing variables */
/* Testing assumptions: Model overspecified by two VAR -> 4 predictors in a subset of 300 patients with 21 events */

/* Testing assumptions: Verifying multicollinearity */
Title "Cox-Proportional Hazards Model: testing assumptions #5, Verifying multicollinearity";
proc reg data = IBD_Data2_30_90;
  model Abn_pers90 = InitDx FamHxIBD Steroids_b4abn Abx_b4abn/ TOL collin vif;
run;
/* Variance Inflation (VIF) < 10 (and mostly <1.1) AND condition index < 8 (mostly < 8), TOLERANCE VALUES ARE > 0.25 (AND MOSTLY > 0.92s)...therefore variables are not collinear */

<table>
<thead>
<tr>
<th>Parameter Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>Intercept</td>
</tr>
<tr>
<td>InitDx</td>
</tr>
<tr>
<td>FamHxIBD</td>
</tr>
<tr>
<td>steroids_b4abn</td>
</tr>
<tr>
<td>Abx_b4abn</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Collinearity Diagnostics</th>
</tr>
</thead>
<tbody>
<tr>
<td>#</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
</tbody>
</table>
Copyright Acknowledgements