Investigating Genetic and Pro-fibrogenic Immune Markers of Hepatitis C Outcomes in HIV-Hepatitis C (HCV) Co-infected Individuals

Nasheed Moqueet

Department of Epidemiology, Biostatistics and Occupational Health
McGill University
Montreal, Quebec, Canada

April 2016

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

TABLE OF CONTENTS

ABSTRACT

Background: In HIV-Hepatitis C (HCV) co-infected individuals, impaired HCV-specific immune response and high inflammation leads to adverse outcomes like lower HCV spontaneous clearance and faster liver fibrosis progression. Liver fibrosis can be a precursor to advanced, possibly irreversible liver damage and is therefore an important intervention point, especially in the early stages. While HCV cure is possible, high costs of effective direct-acting antiviral agents (DAAs), low treatment uptake, HCV re-infection, and other hepatotoxic exposures remain problems in the co-infected population. Characterizing the genetic and immune markers of the underlying immunological mechanisms triggered by HCV persistence in co-infected persons can help in understanding disease etiology and improve treatment decision-making by identifying higher-risk individuals. This is especially important because HCV viral cure is less likely when fibrosis has progressed to advanced cirrhosis.

Objectives: Several host genetic and immune factors have been studied in other populations and reported as markers of HCV pathogenesis. We wanted to examine their roles in the Canadian HIV-HCV co-infected population, which has a unique genetic mix due to an overrepresentation of Aboriginal peoples. Our objectives in this dissertation are the following:

1. Test the association of HCV spontaneous clearance and three single nucleotide polymorphisms (SNPs) near the *Interferon Lambda 3 (IFNL3)* gene (rs12979860, rs8099917, functional variant rs8103142) and compare the SNP frequencies between Canadian whites and Aboriginal peoples
2. Test the association of IFNL SNPs with significant liver fibrosis after HCV clearance fails

3. Assess whether pro-fibrogenic immune and genetic markers improve ability to predict three-year risk of significant liver fibrosis over clinical risk factors alone

Methods and Results: All study samples were derived from eligible subpopulations of the Canadian Co-infection Cohort (CCC). The Canadian Co-infection Cohort Study (n=1,423), established in 2003, is an open prospective cohort of HIV-HCV co-infected individuals recruited from 19 centres across Canada, representing approximately 23% of the co-infected population under care. Data and samples are collected at visits every six months from participants who had serologic evidence of HIV and HCV at entry. We conducted survival analyses using Cox proportional hazards models for each of the three objectives.

For all three SNPs, the homozygous wild-type genotype was of interest, consisting of 2 copies of the major alleles (e.g. rs12979860 CC vs. non-CC genotypes; rs8099917 TT vs. non-TT; and rs8103142 TT vs. non-TT). Haplotype analysis, adjusted for ethnicity, tested the effect of TCT, the haplotype containing the major allele at all 3 SNPs (T at rs8103142, C at rs12979860 and T at rs8099917) on spontaneous clearance (Objective 1) and liver fibrosis (Objective 2). Significant liver fibrosis was measured by APRI (Aspartate aminotransferase-to-platelet ratio index) score of 1.5 or higher.

IFNL genotypes and spontaneous clearance (Objective 1): Individuals who had HCV treatment or less than two HCV RNA tests were excluded. Spontaneous clearance was defined as two consecutive HCV RNA-negative tests, at least six months apart. Self-reported ethnicity was used to account for population stratification. Results showed that, as in other populations, the IFNL

genotypes of interest were linked with clearance rates at least three times higher than in those lacking the genotypes, after adjusting for sex and ethnicity. The major "beneficial" alleles, genotypes and haplotypes were all more frequent among Aboriginal peoples than whites, but this only partially explained why Aboriginal individuals had higher clearance rates.

IFNL genotypes and significant liver fibrosis (Objective 2): Viral persistence was verified by positive HCV RNA test at first available visit after excluding those with prevalent liver disease or on HCV treatment. Other known risk factors were included in the final Cox model. Each IFNL genotype, associated with pro-inflammatory responses and higher clearance, was linked with a higher risk of significant liver fibrosis. The relationship with rs8099917 TT was strongest, indicating a 79% increase in fibrosis risk. Haplotype analysis also supported the link with higher risk of liver fibrosis.

Predicting 3-year risk of liver fibrosis with immune markers (Objective 3): Individuals were excluded if they had cleared HCV, were on HCV treatment or had liver disease at baseline. A case cohort design was used as an economical way to measure markers not collected during regular follow-up and as a more practical way to estimate 3-year risk of significant liver fibrosis. Specific immune markers were measured from first available plasma or serum in the randomly selected subcohort and fibrosis cases only. Cox proportional hazards models with robust variance and Barlow weights was used. Prediction metrics (discrimination, calibration and risk classification) were compared between Model 1 (selected clinical predictors only) and Model 2 (clinical predictors from Model 1 plus selected markers at IFNL rs8099917 and 5 immune markers: IL-8, sICAM-1, RANTES, hsCRP, and sCD14). Both models were well-calibrated. The

improvement in discrimination with model 2 was small, but the model with the markers fit and classified risk better.

Conclusion: Specific IFNL genotypes indicated a higher likelihood of spontaneous HCV clearance in co-infected Canadians. They were far more common in Aboriginal peoples, who cleared more often. Other mechanisms likely also contribute as IFNL genotypes did not fully account for their higher clearance rates. Once clearance fails, the same IFNL polymorphisms, reflecting a pro-inflammatory response, also are associated with a higher risk of developing significant liver fibrosis. Other markers of the heightened inflammation that accelerate liver disease in HIV-HCV co-infection can improve ability to predict 3-year risk of significant liver fibrosis, but require further cost-benefit analyses and external validation in other populations.

RÉSUMÉ

Contexte : Chez les individus co-infectés avec le VIH et l'hépatite C, l'altération de la réponse immunitaire contre l'hépatite C et un niveau d'inflammation élevé mènent à des effets indésirables tels qu'une réduction du taux de résolution spontanée de l'hépatite C et une progression accélérée de la fibrose du foie. La fibrose du foie peut être un précurseur de dommages avancés au foie pouvant être irréversibles. Il s'agit donc d'un point d'intervention important, particulièrement lors des stades précoces. Bien qu'il est possible de guérir l'hépatite C, le cout élevé des antiviraux à action directe, le faible taux de traitements amorcés, la réinfection par l'hépatite C, ainsi que l'exposition a'dautres éléments hépatotoxiques demeurent un problème pour la population co-infectée. La caractérisation des marqueurs génétiques et immunitaires sous-jacent aux mécanismes immunologiques déclenchés par la persistance de l'hépatite C chez les gens co-infectés pourrait améliorer notre compréhension de l'étiologie de la maladie ainsi que la prise de décision quant au traitement en identifiant les individus présentant un risque élevé. Ceci est particulièrement important puisque la probabilité de guérison de l'hépatite C diminue lorsque la fibrose a progressé en une cirrhose avancée.

Objectifs : Plusieurs facteurs immunitaires et génétiques de l'hôte ont été étudiés et identifiés comme marqueurs de pathogénicité de l'hépatite C dans d'autres populations. Nous avons voulu examiner leur rôle dans la population canadienne co-infectée par le VIH et l'hépatite C qui est caractérisée par un mélange génétique unique en raison de la surreprésentation des peuples autochtones. Nos objectifs pour cette dissertation sont les suivants :

1. Tester l'association entre la résolution spontanée de l'hépatite C et trois polymorphismes mononucléotidiques (SNPs) près du gène *Interferon Lambda 3 (IFNL3)* (rs12979860,

rs8099917, variant à effet dommageable rs8103142) et comparer la fréquence des SNPs entre les Canadiens blancs et autochtones.

2. Tester l'association entre les SNPs IFNL et la fibrose du foie suite à un échec de résolution de l'hépatite C.
3. Évaluer la capacité des marqueurs immunitaires et génétiques pro-fibrogènes à améliorer les prédictions du risque sur trois ans de fibrose du foie par rapport aux facteurs de risque cliniques seuls.

Méthodes et résultats : Tous les échantillons à l'étude ont été tirés de sous-populations éligibles de la Cohorte canadienne de co-infection (CCC). La Cohorte canadienne de co-infection (n=1423), établie en 2003, est une cohorte prospective ouverte d'individus co-infectés par le VIH et l'hépatite C, recrutés dans 19 centres à travers le Canada et représentant environ 23% de la population co-infectée recevant des soins. Les données et échantillons sont collectés lors de visites aux six mois chez les participants présentant des évidences sérologiques d'infections au VIH et à l'hépatite C lors de leur entrée dans la cohorte. Nous avons performé des analyses de survie avec modèles à risques proportionnels de Cox pour chacun des trois objectifs.

Pour chacun des trois SNPs, le génotype d'intérêt était l'homozygote sauvage, composé de deux copies de l'allèle majeur (p. ex. génotypes rs12979860 CC vs. non-CC; rs8099917 TT vs. non-TT; et rs8103142 TT vs. non-TT). Une analyse d'haplotype ajustée pour l'ethnicité a été employée pour tester l'effet de TCT, l'haplotype comprenant l'allèle majeur sur les trois SNPs (T sur rs8103142, C sur rs12979860 et T sur rs8099917) sur la résolution spontanée (Objectif 1)

et la fibrose du foie (Objectif 2). La fibrose du foie a été mesurée par un score APRI (*Aspartate aminotransferase-to-platelet ratio index*) de 1,5 ou plus.

Génotypes IFNL et résolution spontanée (Objectif 1) : Les individus ayant été traités pour l'hépatite C ou ayant moins de deux tests d'ARN pour l'hépatite C ont été exclus. La résolution spontanée a été définie comme deux tests d'ARN de l'hépatite C négatifs consécutifs, séparés d'au moins six mois. L'ethnicité auto-rapportée a été utilisée pour stratifier la population. Les résultats obtenus démontrent que, comme chez les autres populations, les génotypes INFL d'intérêt étaient associés à des taux de résolution au moins trois fois plus importants que chez ceux ne possédant pas ce génotype, et ce, après ajustement pour le sexe et l'ethnicité. Les principaux allèles, génotypes et haplotypes « bénéfiques » étaient tous plus fréquents chez les autochtones que chez les blancs, ce qui n'explique qu'en partie le fait que les autochtones présentaient un taux de résolution plus élevé.

Génotypes IFNL et fibrose du foie (Objectif 2) : La persistance virale a été confirmée par un test d'ARN de l'hépatite C positif lors de la première visite disponible, excluant les participants avec une maladie du foie ou traités pour l'hépatite C. D'autres facteurs de risque connus ont été inclus dans le modèle de Cox final. Chaque génotype IFNL associé à une réponse pro-inflammatoire et un taux accru de résolution a été lié à un risque accru de fibrose du foie. La relation avec rs8099917 TT était la plus forte, indiquant une augmentation du risque de fibrose du foie de 79%. Ce lien avec un risque accru de fibrose du foie était aussi soutenu par l'analyse d'haplotype.

Prédire le risque sur trois ans de fibrose du foie à l'aide de marqueurs immunitaires (Objectif 3) : Nous avons exclus les individus dont l'infection à l'hépatite C avait été résolue, sous traitement pour l'hépatite C, ou ayant une maladie du foie lors de la première visite. Un design cas-cohorte a été employé afin de mesurer de façon économique les marqueurs non collectés lors des visites de suivi régulier et estimer le risque sur trois ans de fibrose du foie. Les marqueurs immunitaires sélectionnés ont uniquement été mesurés dans le premier échantillon de plasma ou de sérum disponible dans la sous-cohorte sélectionnée de manière aléatoire ainsi que chez les cas de fibrose. Des modèles à risques proportionnels de Cox avec variance robuste et poids de Barlow ont été utilisés. Des mesures de prédiction (discrimination, calibration et classification de risque) ont été comparés entre le modèle 1 (prédicteurs cliniques sélectionnés uniquement) et le modèle 2 (prédicteurs cliniques du modèle 1 en plus d'une sélection de marqueurs à IFNL rs8099917 et de cinq marqueurs immunitaires : IL-8, sICAM-1, RANTES, hsCRP et sCD14). Les deux modèles étaient bien calibrés. Le modèle 2 a conféré une modeste amélioration de la discrimination, mais un degré d'ajustement et une classification du risque supérieurs.

Conclusions: Certains génotypes IFNL ont indiqué une probabilité accrue de résolution spontanée de l'hépatite C chez les Canadiens co-infectés. Ces génotypes étaient beaucoup plus communs chez les peuples autochtones, chez qui la résolution était plus fréquente. Il est probable que d'autres mécanismes y contribuent puisque les génotypes IFNL ne pouvaient expliquer en totalité leur plus haut taux de résolution spontanée. Lors d'échec de résolution, les mêmes polymorphismes IFNL reflétant une réponse pro-inflammatoire étaient aussi associés à un risque accru de fibrose du foie. Bien que d'autres marqueurs d'inflammation élevée pouvant accélérer la progression des maladies du foie chez les gens co-infectés avec le VIH et l'hépatite C peuvent

améliorer la capacité de prédiction du risque sur trois ans de fibrose du foie, de plus amples études de cout-efficacité et de validation externes chez d'autres populations sont nécessaires.

ACKNOWLEDGMENTS

Thank you to my outstanding supervisors, Marina Klein and Robert Platt, for their guidance and support. I am especially grateful for their clear thinking and insightful feedback which have made me a better researcher, epidemiologist, and writer. I also thank them for the opportunities they have given me and the ones they encouraged me to pursue.

A special thank you to my thesis committee and co-authors, especially Claire Infante-Rivard and Jim Young, for their expertise and assistance. From the Klein research team, thanks to Rhyan Pineda, Kathleen Rollet-Kurhajec, and Sahar Saeed for making the biological and statistical analyses possible and to Chantal Burelle, for her administrative wizardry, especially with manuscript submission. From McGill, I would like to acknowledge all my professors and the wonderful staff, especially Andrew Griffin and Andre Yves Gagnon. From CanHepC, special thanks to Norma Choucha for her exceptional efficiency.

I am deeply appreciative of the friends I have made at McGill, for Elise, Genevieve, Leah, Layla, and Laurence for their humor and camaraderie, and especially to Laurence for the French translation of the abstract in this dissertation. And to my sisters from another mother-- Courtney, Naveen, Saunia, and Almo---thanks for the right balance of cheerleading, tough love and sidesplitting laughter.

A special thank you to to my family, but especially to my husband for his unconditional love and faith in me, and to my parents, who raised me to believe that the pursuit of educational excellence is empowering and who always pushed me to do my very best. I am also grateful to

my in-laws for their help, and to Shaheen and Ruby, who loved and civilized my wild daughter. Finally, I thank my two-year-old daughter, Maya, for not getting sick too often so I could focus on my work and for letting me grow her curly hair so she can look like the mad/genius scientist I hope she grows up to be.

STATEMENT OF FINANCIAL SUPPORT

I am incredibly grateful for the funding support I received during my doctoral training: a *Frederick Banting and Charles Best Canada Graduate Scholarship Doctoral Award* from the Canadian Institutes of Health Research (CIHR DRA; 201010MDR), as well as a doctoral award from the Canadian Network on Hepatitis C (CanHepC, formerly National CIHR Research Training Program in Hepatitis C or NCRTP-Hep C), both of which also covered travel costs for workshops and conferences. I also received funding from my supervisor, Marina Klein, as well as the Department of Epidemiology, Biostatistics and Occupational Health at McGill University.

CONTRIBUTION OF AUTHORS

Manuscript 1. Distribution of *IFNL3* Genotypes in Canadian Aboriginal Peoples and Association with HCV Spontaneous Clearance

Authors: Nasheed Moqueet, Claire Infante-Rivard, Robert W. Platt, Jim Young, Curtis Cooper, Mark Hull, Sharon Walmsley, Marina B. Klein and the Canadian Co-Infection Study Investigators. *Int J Mol Sci*. 2015.

Nasheed Moqueet and Marina B. Klein designed the study. Nasheed Moqueet drafted the manuscript and conducted all the analyses with help from Claire Infante-Rivard, Robert W. Platt, Jim Young and Marina B. Klein. All authors have commented on the manuscript.

Manuscript 2. IFNL Genotypes and Significant Liver Fibrosis

Authors: Nasheed Moqueet, Curtis Cooper, John Gill, Mark Hull, Marina B. Klein, Canadian Co-infection Cohort. *JID.* 2016

Nasheed Moqueet and Marina Klein designed the study. Nasheed Moqueet drafted the manuscript and conducted all the analyses with help from Marina Klein. All of the authors have commented on the manuscript.

Manuscript 3. Developing a Prognostic Model for Significant Liver Fibrosis using Immune and Genetic Markers

Authors: Nasheed Moqueet, Cynthia Kanagaratham, Danuta Radzioch, Sahar Saeed, Robert W. Platt, Marina B. Klein.

Nasheed Moqueet, Marina Klein, and Robert W. Platt designed the study. Nasheed Moqueet drafted the manuscript and conducted all the analyses with help from Robert W. Platt. Cynthia Kanagaratham conducted all the Luminex experiments and analysis in the laboratory of Danuta Radzioch. Sahar Saeed coordinated the sample selection, collection, and analysis.

STATEMENT OF ORIGINALITY

This dissertation makes an original contribution to the field of HCV etiology and the role of immune and genetic markers in the context of HIV-HCV co-infection. All three manuscripts present novel findings from the Canadian co-infected population, which has a unique genetic makeup but also shares characteristics of co-infected populations in other high-income countries. Our findings have implications for understanding HCV pathogenesis and clinical decision-making.

Manuscript 1 was the first published report on the distribution and effect of IFNL genotypes in a Canadian co-infected population, which has an overrepresentation of Aboriginal peoples. Higher frequency of the beneficial markers in Aboriginal individuals partially explained why they cleared more often, but even after accounting for IFNL genotypes, clearance rates in Aboriginal peoples were higher than Canadian whites. This is one of the first studies that suggest an alternative mechanism of HCV resolution unique to Aboriginal peoples, as the frequency and effect of IFNL genotypes in Canadian whites was very similar to those reported in other populations of European descent.

Manuscript 2 built on the results of Manuscript 1 and indicated that if HCV clearance fails, the same IFNL genotypes that were linked to higher clearance could also be markers for a higher risk of significant liver fibrosis. Our study was one of the few conducted in an HIV-HCV co-infected population, where hepatotoxic exposures are common and liver fibrosis progression is accelerated. The findings support those from a larger study in patients without HIV infection[1] which found that the IFNL genotypes were associated with both higher necroinflammation and

higher risk of liver disease, regardless of underlying cause. This could have implications for clinical decision-making and follow-up care, especially if hepatotoxic behaviours persist or re-infection occurs.

Manuscript 3 reported the novel application of an infrequently-used study design-- a case cohort-- to examine if genetic and immune markers improved ability to predict 3-year future risk of significant liver fibrosis beyond traditional clinical predictors alone. While many others have examined the ability of noninvasive markers to improve diagnosis of liver disease, ours is one of the first that specifically examines prognosis (i.e. estimating future risk) of liver fibrosis, which is especially timely in an era of expensive HCV treatment.

While I acknowledge the contribution of Laurence Brunet, who did the French translation of my abstract and the guidance and assistance from my supervisors and co-authors on methodology, statistics and subject matter, the studies presented in this thesis represent my original work.

LIST OF ABBREVIATIONS

AIC	Akaike Information Criterion
APRI	Aspartate aminotransferase to platelet ratio index
AST	Aspartate aminotransferase
AUROC	Area under the Receiver Operating Characteristic
C-statistic	Concordance statistic
cART	Combination antiretroviral therapy
CI	Confidence interval
CXCL11	Chemokine (C-X-C motif) Ligand 11
CXCL9	Chemokine (C-X-C motif) Ligand 9
DAA	Direct-acting antivirals
ECM	Extracellular matrix
EIA	Enzyme immunoassay
ELISA	Enzyme-linked immunosorbent assay
ESLD	End-stage Liver Disease
GB Test	Gronnesby and Borgan Test
GGT	Gamma-glutamyltransferase
GWAS	Genome-wide association study
HCC	Hepatocellular carcinoma
HCV	Hepatitis C virus
HR	Hazard ratio
HSC	Hepatic stellate cells
hsCRP	High-sensitivity C-reactive protein
HWE	Hardy-Weinberg Equilibrium
IFNL3	Interferon Lambda 3
IL-6	Interleukin-6
IL-8	Interleukin-8
IL28B	Interleukin 28B
IPW	Inverse probability weights
LD	Linkage disequilibrium
LPS	Lipopolysaccharide
MCP-1	Monocyte Chemotactic Protein-1
MI	Multiple imputation
MICE	Multivariate Imputation using Chained Equations
MIP1α	Macrophage Inflammatory Protein 1 alpha
MIP1β	Macrophage Inflammatory Protein 1 beta
MSM	Men who have sex with men
OR	Odds ratio
PBMC	Peripheral blood mononuclear cells
PWID	People who inject drugs
RANTES	Regulated upon Activation, Normal T cell Expressed and Secreted protein
RIBA II	Recombinant immunoblot assay II
RNA	Ribonucleic acid
sCD14	Soluble CD14
sICAM-1	Soluble Intercellular Adhesion Molecule 1
SNP	Single nucleotide polymorphism
sVCAM-1	Soluble Vascular Cell Adhesion Molecule 1
TGF-β	Transforming Growth Factor Beta
TNF-α	Tumor Necrosis Factor alpha

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1. Introduction

HIV and Hepatitis C (HCV) co-infection affects over 2.3 million individuals worldwide.[2] Globally, 6.2% of HIV-infected individuals also have HCV co-infection.[2] In this population, partially due to HIV-related immune dysregulation and inflammation, HCV is more likely to persist, leading to liver damage which can progress quickly to irreversible and even fatal outcomes.[3,4] HCV cure is most likely and beneficial in the early stages of liver fibrosis development[5], which is thus a key intervention point. However, in Canada, as in many other countries, the high cost of effective HCV direct-acting antiviral agents (DAAs) has meant that reimbursement has been restricted to those with advanced fibrosis.[6] Other obstacles such as HCV re-infection and prevalence of other hepatotoxic exposures also remain.[7] Therefore, understanding disease etiology and identifying higher-risk individuals is especially important to improve clinical decision-making and to target treatment in co-infected persons.

Several host genetic and immune factors have been studied in other populations and described as markers of HCV pathogenesis. Genetic markers such as single nucleotide polymorphims (SNPs) near the *Interferon Lambda 3 (IFNL3,* also referred to as *IFNL4* in the literature*)* gene have been linked with higher spontaneous and treatment-induced HCV clearance rates. The distribution of these markers also vary by ethnicity but have not been characterized in the Canadian co-infected population, which has a disproportionate number of Canadian Aboriginal peoples, who have a distinct and complex genetic history and face many challenges accessing care. Some published studies also indicated that Canadian Aboriginal peoples have differences in immunological responses to HCV, though none had characterized distribution or function of IFNL SNPs.[8-10]

When spontaneous HCV clearance fails, some studies[1,11,12] found that IFNL SNPs were also linked to liver disease. While this may seem counterintuitive, it is biologically plausible. The mechanism behind IFNL is still under investigation but is considered pro-inflammatory, which is why it could be associated with liver disease, which is caused by chronic inflammation rather than HCV replication. Understanding the IFNL-liver fibrosis link in the context of co-infection could yield insight into the mechanisms of liver disease progression in a heightened inflammatory environment.

Direct markers of the underlying inflammation that drive liver disease could also help identify higher-risk individuals, thus helping to determine treatment urgency and optimize treatment efficacy. Several pro-fibrogenic markers have been characterized and studied in the literature, but mainly for diagnostic, not prognostic, purposes. Developing a prognostic model in a population that is representative of the co-infected individuals in clinical care is important as features such as marker and disease prevalence or marker correlations with other known risk factors can all affect the estimates of the discriminatory capacity of a marker.[13]

The overarching goal of this dissertation then, is to examine the link of genetic and immune markers with various HCV outcomes as well as assess their prognostic ability to predict liver fibrosis in the context of HIV-HCV co-infection.

1.1 RESEARCH OBJECTIVES

Specifically, our objectives are the following:

1. Test the association of HCV spontaneous clearance and three single nucleotide polymorphisms (SNPs) near the *Interferon Lambda 3 (IFNL3)* gene (rs12979860, rs8099917, and the functional variant rs8103142) and compare the SNP frequencies between Canadian whites and Aboriginal peoples
2. Test the association of IFNL SNPs with significant liver fibrosis after HCV clearance fails
3. Assess whether pro-fibrogenic immune and genetic markers improve ability to predict three-year risk of significant liver fibrosis over clinical risk factors alone

1.2 FORMAT OF THE THESIS

This dissertation is based on 3 manuscripts, each corresponding to the research objectives above and contained in Chapters 4-6. Chapter 2 consists of an extensive literature review. Chapter 3 provides additional details on topics common to the manuscripts, such as the source population (the Canadian Co-infection Cohort), statistical analysis (Cox proportional hazards models, which was used in all 3 manuscripts) and variables (IFNL genotype as independent variables for Manuscripts 1 and 2 and significant liver fibrosis as outcome for Manuscripts 2 and 3).

Each manuscript is presented in its own chapter, which also contains a preface with information specific to each research objective followed by an appendix with supplementary figures and results of sensitivity analyses. Chapter 4 includes Manuscript 1, which examines the relationship between IFNL genotypes and spontaneous HCV clearance and compares IFNL frequency between Aboriginal peoples and Canadian-born whites. Chapter 5 contains Manuscript 2, which examines the link between IFNL genotypes and significant liver fibrosis after HCV clearance

fails. Chapter 6 contains Manuscript 3, which examines the ability of genetic and immune markers to improve clinical prognosis of liver fibrosis beyond clinical predictors. Chapter 7 presents a cohesive discussion of all 3 studies as well as implications of the findings on clinical care and research, followed by concluding remarks.

CHAPTER 2. Literature Review

2.1 EPIDEMIOLOGY OF HIV

Globally, the World Health Organization (WHO) and UNAIDS estimates that almost 36.9 million (95% Confidence Interval: 34.3, 41.4 million) people were living with HIV in 2014, mainly in sub-Saharan Africa.[14] Because of the advances in combination antiretroviral therapy (cART), almost half of the global HIV-positive population (~41%) are on antiretroviral treatment. As a result, people are living longer, contributing to an increase in HIV prevalence, even as HIV incidence continues to fall, mostly due to reductions in heterosexual transmission.[15] In addition to sexual contact, HIV can also be spread by contaminated needles and blood as well as from mother to child during pregnancy, birth or breastfeeding.

In Canada, according to national estimates from 2011, approximately 71,300 people were living with HIV, for a prevalence of 208 per 100,000 (Range: 171–245.1 per 100,000).[16] In the same time period (i.e. in 2011), there were 3,175 new infections.[16] The biggest risk group is still men who have sex with men (MSM), who make up 46.6% of new infections,[16] though this varies regionally and ethnically. For example, in Saskatchewan, HIV is most concentrated among people who inject drugs (PWID) which is also the case among Canadian Aboriginal peoples nationally.[16,17] Saskatchewan also remains the province with the highest HIV diagnosis rate, with 11.4 diagnosed cases per 100,000, which is double the national diagnosis rate in 2013.[18] This high burden, which also disproportionately affects the Aboriginal population in Saskatchewan, underscores the importance of research specific to this community.

2.2 EPIDEMIOLOGY OF HEPATITIS C (HCV)

Worldwide, the number of people living with HCV (130–150 million, ~2% prevalence)[19] is almost four times the number living with HIV. Most of the HCV-infected live in Africa, Eastern Europe and Asia (prevalence ranging from 3.2% in China[20,21] to 20% in Egypt[22]), where risk factors and HCV genotypes differ from Western Europe, North America and Australia (prevalence <2%).[19] There are six major HCV genotypes and numerous subtypes.[23,24] The most common and efficient route of HCV transmission is percutaneously, through injection drug use or during medical procedures (such as the schistosomiasis vaccination campaign in Egypt), accidental needle injuries, or transfusion of blood or blood products. HCV can also spread vertically, though rarely, from mother-to-child or via sexual contact.[24,25]

In Canada, it is estimated that 242,500 people were infected with HCV, with almost 11,357 new infections being reported in 2009 for an incidence rate of 33.7 per 100,000.[26,27] Like HIV, HCV incidence is also higher in Saskatchewan than other provinces, mainly driven by injection drug use.[28] Model projections estimate that the numbers for 2013 could be similar with viremic HCV cases at 251,990 (Uncertainty interval 177,890 to 314,800) and the highest number of cases in the 40- to 54-year age group.[29] These projections also suggest that while HCV prevalence might have peaked in 2003, at current treatment uptake, the proportion of patients with advanced HCV-related liver disease will increase between 2013 and 2035, reflecting aging of the infected population and slow rates of fibrosis progression.[29]

2.3 EPIDEMIOLOGY OF HIV-HCV CO-INFECTION

HIV and Hepatitis C co-infection affects over 2.3 million individuals (Interquartile Range 1.27-4.42 million) worldwide.[2] In high-income countries, about 25-30% of HIV-positive individuals overall also have HCV, though this varies by subpopulations. For example, 82.4% (IQR 55.2–88.5%) of HIV-positive injection drug users are co-infected with HCV, while prevalence estimates are lower for those who acquired HIV through sexual transmission--6.4% (IQR 3.2–10%) among men who have sex with men (MSM) and 4% (IQR 1.2–8.4%) of heterosexuals.[2] This also reflects the efficiency of percutaneous transmission of HCV which is 10 times more infectious than HIV.[4]

In Canada, it was estimated that 13-15,000 individuals were co-infected in 2003.[30] The main route of HCV transmission has been via injection drug use, with studies estimating that after 5 years of injecting, 50-90% of users are infected with HCV.[31] In Saskatchewan, for example, where the burden of HIV is among the highest in Canada, 70%-90% of those diagnosed with HIV are also co-infected with hepatitis C.[28] Thus, while co-transmission of both HIV and HCV can occur, co-infected individuals with percutaneous exposure are usually infected first-and for much longer- with HCV.[32]

2.4 CANADIAN ABORIGINAL PEOPLES AND HIV, HCV, AND HIV-HCV CO-INFECTION

Canadian Aboriginal peoples (Metis, Inuits, and First Nations) are over-represented in the HIV, HCV and HIV-HCV co-infected populations.

2.4.1 Aboriginal peoples and HIV monoinfection

Despite representing only 3.8% of the Canadian population,[33] Aboriginal peoples represented about 8% of all people living with HIV and AIDS, and about 12.5% of new HIV and AIDS cases diagnosed in Canada in 2008.[17] This problem is especially notable among injection drug users, as the number of HIV infections among Aboriginal PWID was five-fold greater than for the rest of the population.[30] In fact, injection drug use is the main mode of HIV exposure in the Aboriginal community, unlike the general Canadian population, where MSM account for the largest proportion of new HIV infections.[17]

2.4.2 Aboriginal peoples and HCV monoinfection

Aboriginal peoples have also disproportionately been affected by HCV. Between 1999-2004, people of Aboriginal descent accounted for 15.2% of new HCV infections,[34,35] and similarly, from 2004-2008,[26] the rates of acute HCV infection were 5.5 times higher in the Aboriginal population than in non-Aboriginal people.[17] Higher HCV incidence and prevalence in Canadian Aboriginal peoples was also confirmed by a meta-analysis that compared trends globally from various drug injecting populations from 1989-2006.[36] However, despite higher HCV acquisition, less than 5% of HCV-infected Aboriginal peoples had detectable HCV RNA compared with 75% of non-Aboriginal Canadians in one study,[37] possibly reflecting a distinct clinical and immunological response to the virus.[8,38,39]

2.4.3 Aboriginal peoples and HIV-HCV co-infection

Due to injection drug use, rates of HIV-HCV co-infection are higher in Aboriginal peoples. In studies from North Alberta,[40] Ontario,[41] and British Columbia,[42]Aboriginal peoples were over-represented in the HIV-HCV co-infected populations. In the Canadian Coinfection Cohort (CCC), which includes approximately 23% of the co-infected population under care, a disproportionate amount (18%) of the participants report Aboriginal ancestry.[43]

Despite being overrepresented among the HCV mono and co-infected populations, Canadian Aboriginal peoples also display higher HCV spontaneous clearance rates[8] and HCV treatment responses.[44] We therefore wanted to characterize host genetic factors responsible for such distinct immunological profiles in Manuscript 1. Our results could be useful in understanding disease history and tailoring treatment for this marginalized population.

2.5 HIV PATHOGENESIS

Effective cART can prolong life by reducing AIDS-related conditions, but it cannot fully eradicate the virus, which persists in reservoirs. HIV infects and destroys CD4 T cells, thus making the individual susceptible to opportunistic infections. With effective cART, viral levels drop and CD4 T cells can be partially restored, though increased immune activation and residual inflammation still persist.[15] As a result, half of the deaths in developed countries are no longer due to AIDS but are due to non-AIDS-defining cancers and cardiovascular, kidney, liver and neurological diseases.[45,46]

In HIV-HCV co-infected individuals, HCV infection may weaken the immune restorative effect of cART,[47] though in general, it does not greatly influence the course of HIV or time to AIDS.[48] In contrast, HIV has a particularly negative effect on progression of liver disease in those with HIV-HCV co-infection. This is detailed in the section "**2.8 Accelerated Development of Liver Disease in Co-infected individuals**" and Figure 2.2.

2.6 NATURAL HISTORY OF HCV

Acute HCV refers to the first 6 months after infection, during which the majority of patients (~70-80%) are asymptomatic.[19,49] Symptoms, when displayed, are non-specific: fever, fatigue, decreased appetite, nausea, vomiting, abdominal pain, dark urine, grey-coloured faeces, joint pain and jaundice (yellowing of skin and the whites of the eyes).[19,50] After acute HCV infection, multiple pro-inflammatory mediators contribute to recruitment of immune cells to the liver in an attempt to restrict viral replication.[3] In 20-45% of cases, the response is successful in clearing the virus without treatment, i.e. spontaneous HCV clearance occurs.[51,52] More commonly, and especially in co-infected individuals, infection becomes chronic, leading to persistent inflammation and liver regeneration and fibrosis, which can worsen to cirrhosis, hepatocellular carcinoma or end-stage liver disease (ESLD), which can be fatal.[3,53] See Figure 2.2.

More details are also given in the section "**2.8 Accelerated Development of Liver Disease in Co-infected individuals.**" Thus, liver fibrosis progression, which is a dynamic process, is a critical link in the pathogenesis of HCV-related disease, and a potential intervention point.

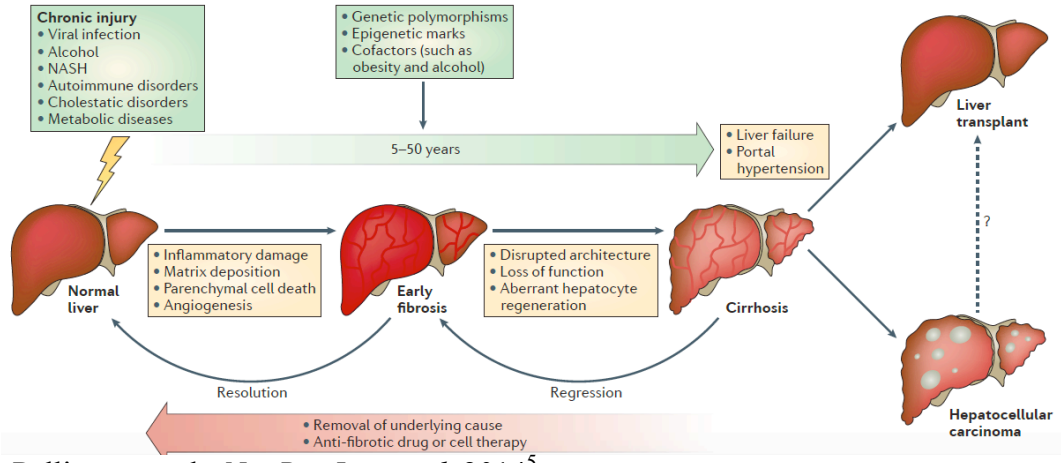
2.7 LIVER FIBROSIS BIOLOGY

The development of fibrosis is summarized in Figure 2.1. A complex interplay of cells and signals create an inflammatory environment that lead to liver fibrosis, which is triggered by repeated liver injury such as occurs in chronic HCV infection. Other factors that have been associated with fibrosis progression include: age at HCV infection, alcohol intake (>50 g/day), male sex, host genetic factors, hepatitis B co-infection, excess weight, liver steatosis, presence of metabolic syndrome and/or type II diabetes, immunodeficiency related to HIV or the use of immunosuppressant drugs such as those used after liver transplantation, and HIV therapy interruption.[54,55]

Liver fibrosis is associated with major changes in both the quantity and the quality of extracellular matrix (ECM), the tissue scaffolding necessary for both structural and functional support.[53] In advanced stages of fibrosis, the liver contains up to 6 times more ECM than normal due to both accelerated synthesis and decreased degradation.[53,56,57] The sequence is roughly as follows: HCV infects liver cells (hepatocytes), setting off a chain of events that leads to activated liver macrophages (Kupffer cells).[53] These cells release molecules that stimulate the hepatic stellate cells (HSCs), the main producers of ECM in the liver. The HSCs, stimulated both by the inflammatory milieu as well as by HCV proteins,[58] differentiate into myofibroblast-like cells, acquiring contractile, fibrogenic and proinflammatory properties. [53,56,59,60] That is, the activated HSCs secrete inflammatory cytokines, express cell adhesion molecules, and control the activation of lymphocytes.[53] The result then, is a vicious cycle where inflammatory and fibrogenic cells stimulate each other, distorting the hepatic architecture, eventually leading to fibrosis.

While a fibrotic liver can worsen to cirrhosis and ESLD, it can also be repaired and even revert to its near-normal liver architecture, especially if the causative agent is removed (i.e. viral cure occurs).[5,61] Resolution (that is, complete reversal to near-normal liver architecture) of early hepatic fibrosis can occur, though with cirrhosis, regression (improvement but not full reversal) of fibrosis is more likely.[61] Cirrhosis regression can improve some clinical outcomes, though the risk of HCC and portal hypertension, which can lead to life-threatening complications such as gastrointestinal bleeding, encephalopathy, ascites, spontaneous bacterial peritonitis, and hepatorenal syndrome remain.[62-65] In fact, advanced cirrhosis is irreversible, especially if accompanied by angio-architectural changes such as vascularized septa[65] or thickening of old scars with collagen cross-linking.[66] At this stage, even if viral clearance occurs with effective HCV treatment, the liver damage has reached a point of no return.[5] Monitoring liver damage is therefore vital, not only to track disease progression and determine treatment urgency, but also to assess treatment efficacy, as probability of HCV cure declines with very advanced liver disease.

Figure 2.1 Overall Mechanism of Liver Fibrosis Development



Source: Pellicoro et. al., *Nat Rev Immunol*, 2014[5]

2.8 ACCELERATED DEVELOPMENT OF LIVER DISEASE IN CO-INFECTED INDIVIDUALS

Fibrosis and liver disease are accelerated in co-infected individuals, compared to HCV monoinfected individuals (Figure 2.2). For example, one study from the pre-cART era found that individuals infected only with HCV progress to cirrhosis in 23.2 years compared to 6.9 years in those with both HIV and HCV.[67] Another study found a smaller difference, but still found co-infected individuals progressing to cirrhosis almost a decade before monoinfected individuals (38 years in monoinfected vs. 26 years in co-infected).[68] This effect is most pronounced in those not on HIV treatment (RR=2.5, 95% CI: 1.8–3.4), though even with HIV treatment, co-infected individuals have a 70% higher risk of cirrhosis (RR=1.7 (1.1–2.8)).[69] For example, a 2013 study found that HIV-HCV co-infection was associated with liver fibrosis as advanced as those without HIV who were 10 years older, despite some cART use.[70] Other studies have also found higher risk[71] and faster progression of fibrosis[72,73] in co-infected individuals, despite being on HIV treatment or exhibiting HIV control. These studies reported that over a period of 3 years, almost 25% of co-infected individuals progressed 2 or more Ishak stages of fibrosis compared to 10% of HCV monoinfected persons.[55]A meta-analysis further confirmed that, compared to HCV monoinfection, HIV co-infection was associated with a ~2-fold increased relative risk of cirrhosis.[74] And once cirrhosis develops, there is a 6-fold acceleration to decompensation and death[74] and a higher risk of hepatocellular cancer (HCC).[56]

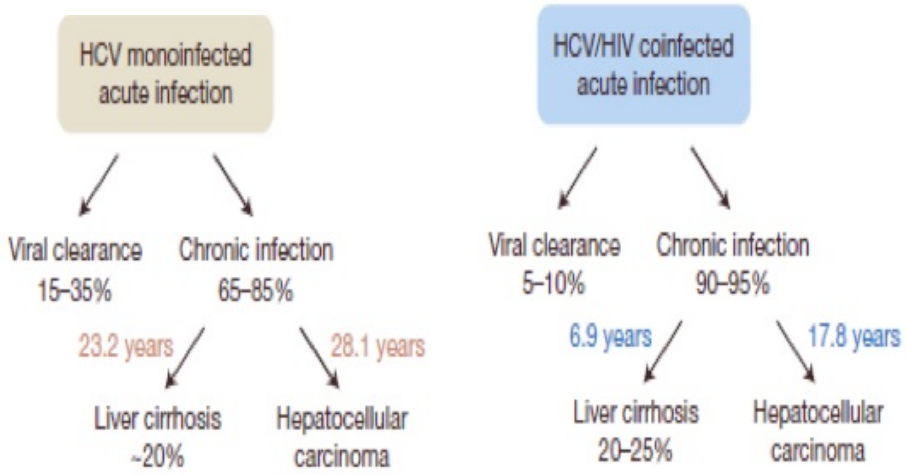
Reasons for this acceleration in co-infected individuals are biological and possibly HIV-therapy related. HIV suppresses the immune response to HCV, leading to weaker, narrower HCV-specific T-cell response, higher HCV replication and persistence, and increased inflammation.[3,74]

Evidence supporting this includes findings that lower CD4 counts (indicating poorer HIV control) are associated with faster fibrosis progression[75,76] and that HIV envelope protein directly promotes liver cell death, leading to release of TGF- β, one of the most important mediating cytokines in human fibrogenesis.[53,77-79] More details on the inflammatory process are outlined in section "**2.14 Inflammation and Fibrosis In Co-Infected Individuals**" and Figure 2.2 (b).

Thus, while HIV control might control inflammation and fibrosis, lack of adherence and access to antiretrovirals may counteract these benefits. Poor adherence is an important consideration among active injection drug users,[80] who make up the majority of Canada's co-infected population and who often experience numerous competing priorities. Treatment interruptions can be quite harmful, leading to HIV viral rebound, increased inflammation,[81] and HIV resistance. Data from several randomized clinical trials which used CD4 count to guide HIV treatment indicate that treatment interruptions not only hamper immune reconstitution,[82] but also lead to adverse liver outcomes.[81,83,84] Among co-infected individuals, HIV treatment interruption is especially damaging, and has been associated with faster liver fibrosis progression[85] and a higher risk of death than in monoinfected individuals.[86] Finally, though some studies have suggested that HIV antiretrovirals themselves may be hepatotoxic,[87] the benefits of HIV control seem to outweigh the effect on fibrosis progression in co-infected individuals.[88]

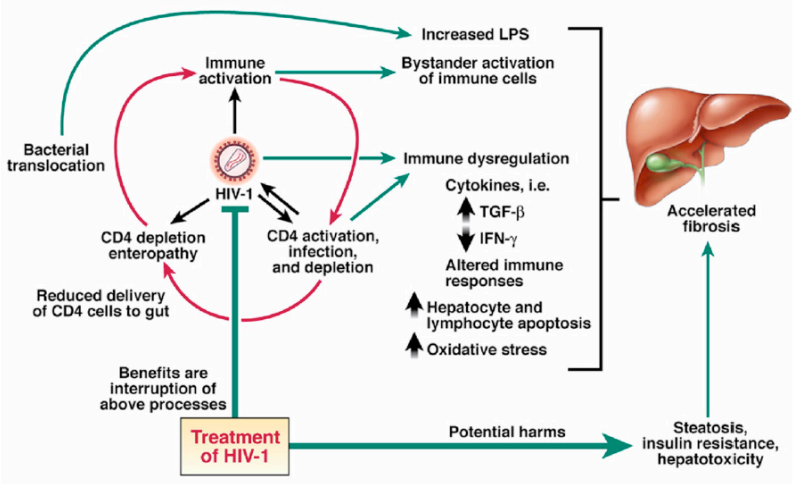
Figure 2.2 Accelerated Fibrogenesis in HIV-HCV Co-infection

a. Pathogenesis in HCV mono-infection vs. HIV-HCV Co-infection



Source: Roe et. al., *Expert Rev Mol Med,* 2008[3]

b. Putative Mechanisms of Accelerated Fibrosis Progression in HIV-HCV Co-infection



Source: Kim & Chung, *Gastroenterology,* 2009[78]

2.9 MEASURES OF LIVER FIBROSIS

Liver biopsy is considered the gold standard for staging liver disease.[89] Several grading systems exist which classify severity of liver damage based on histologic information from biopsies. The most widely used scales include the 4-stage Batts-Ludwig, the 4-stage METAVIR systems and the 6-stage ISHAK score.[90] Significant liver fibrosis corresponds to stages F2 and above in the Batts-Ludwig and METAVIR grades and F3 and above in the ISHAK system.

Despite being the gold standard for staging liver disease, liver biopsies have several drawbacks. They are invasive, costly, risky and subject to selection bias as well as sampling error due to location and size of the liver sample being analysed.[53,89] Furthermore, in injection drug users, who make up the majority of the co-infected population in developed countries, healthcare access may be poor and biopsies are even more difficult to obtain, limiting sample size.[91] For these reasons, liver biopsies are not practical for longitudinal research purposes and other markers have been used.

Noninvasive markers, especially from routine laboratory tests, are suitable for repeated measures, and have been developed for research as well as diagnosing and managing HCV co-infection. One such marker, the aspartate aminotransferase (AST)-to-platelet ratio index (APRI), is an indirect serum measure that reflects alterations in hepatic function[92] and a cutoff of 1.5 has been validated in our study and others to detect significant liver fibrosis (METAVIR stage F2 or higher).[93] This cutoff has higher specificity than sensitivity for diagnosing significant liver fibrosis. In a meta-analysis, the Area under the Receiver Operating Characteristic (AUROC) in

co-infected populations was reported as 0.75.[92] The sensitivity was 0.37 (95% CI: 0.35, 0.39), while the specificity was high at 0.93 (95% CI: 0.91, 0.94).[92]

Despite the poor sensitivity, APRI cutoffs of 1.5 and 2 have been shown to be associated with cirrhosis, other adverse liver and clinical outcomes, and death in our study cohort,[94] as well as in others.[95,96] (More details are given in Chapter 3, section **"3.3 Measuring Significant Liver Fibrosis**"). Both AST and platelets are routinely measured in clinical settings to monitor liver damage and in the CCC, are measured at study visits every six months. This availability, widespread use in other studies, and correlation with clinical outcomes makes APRI a practical option for measuring significant liver fibrosis in our study.

Another widely used alternative, especially in HIV-HCV co-infected populations, is the FIB-4 index, which uses AST, platelet count, age and ALT.[97] A FIB-4 index cutoff at 3.25 corresponds to a METAVIR fibrosis stage F3 (severe or advanced fibrosis) and higher. In HIV-HCV co-infected individuals, this cutoff had an AUROC of 0.765 and specificity of 97%,[97] while in an HCV monoinfected population, it had an even higher estimated AUROC 0.85 (95% CI 0.82-0.89), with a positive predictive value (PPV) of 82%, and specificity of 0.98.[98] However, as FIB-4 $\geq$ 3.25 indicates more advanced fibrosis than an APRI $\geq$ 1.5, for our analyses, it often resulted in fewer events, resulting in lower power. More importantly, for the research question in Manuscript 2, the study by Eslam et. al. found that IFNL genotypes were most important in the earlier fibrosis transitions (F0-F1 and F1-F2) rather than the later ones (F2-F3 and F3-F4).[1] APRI $\geq$1.5 allowed us to potentially capture the biologically significant transition to F2 and higher, which a FIB-4 $\geq$ 3.25 might have missed.

Other measures in use include hyaluronic acid and transient elastography (Fibroscan). Hyaluronic acid is a product of ECM breakdown with other inflammatory properties[99] and has high sensitivity and specificity in co-infected patients, though it is not routinely collected as part of clinical care.[100] Fibroscan, an ultrasound-based measure of hepatic stiffness, has greater accuracy for detecting advanced liver disease rather than earlier fibrotic stages.[101,102] and is fast replacing liver biopsy for clinical decision making. However, these measures are not available longitudinally for the majority of CCC participants and are thus not viable alternatives for classifying disease outcome.

The need for noninvasive measures of liver disease still remains as disease staging can help with many aspects of clinical care, such as discriminating treatment need and predicting treatment efficacy, as well as earlier detection of disease progression. Below is a summary of the diagnostic performance of several biomarkers in use for chronic liver disease from several etiologies (Table 2.1).

Table 2.1 Diagnostic Performance of Serum Biomarkers for Significant Fibrosis (F≥2) and Cirrhosis (F4) in Patients with Chronic Liver Disease of Various Etiologies

Biomarkers	Etiologies	Year	Patients (n)	F≥2 (%)	F4 (%)	Cut-offs	AUROC	Se (%)	Sp (%)	CC (%)
FibroTest® [21]	HCV	2001	339	80		>0.48	0.87	75	85	46
Forns Index [22]	HCV	2002	476	26		<4.2 >6.9	0.81	30-94	51-95	45
APRI [23]	HCV	2003	270	50		≤0.5 >1.5	0.80	41-91	47-95	44
					17	<1.0 ≥2.0	0.89	57-89	75-93	72
FibroSpectII® [24]	HCV	2004	696	52		>0.36	0.83	77	73	75
MP3 [25]	HCV	2004	194	45		<0.3 >0.4	0.82	35-65	85-96	n.a.
FPI [26]	HCV	2005	302	48		≤0.2 ≥0.8	0.77	42-85	48-98	40-49
Hepascore® [27]	HCV	2005	211	57		≥0.5	0.82	63	89	92
					16	>0.84	0.89	71	89	n.a.
Lok index [28]	HCV	2005	1141		38	<0.2 ≥0.5	0.81	40-98	53-99	52
GUCI [29]	HCV	2005	179		12	>0.1	0.85	80	70	n.a.
ViraHep-C [30]	HCV	2006	398	37		≤0.22 >0.55	0.83	51-90	54-90	52
Fibroindex [31]	HCV	2007	360	50		≤1.25 ≥2.25	0.83	30-40	97-97	35
FIB-4 [32]	HCV	2007	830		17*	<1.45 >3.25	0.85	38-74	81-98	68
HALT-C model [33]	HCV	2008	512		38	<0.2 ≥0.5	0.81	47-88	45-92	48
Hui Score [36]	HBV	2005	235	25		≤0.15 >0.5	0.79	37-88	50-88	49
Zeng score [37]	HBV	2005	372	58		<3.0 >8.7	0.77	40-98	28-90	35
SHASTA [38]	HIV-HCV	2005	95	27		<0.3 >0.8	0.87	15-88	72-100	42
FIB-4 [39]	HIV-HCV	2006	832		22*	<1.45 >3.25	0.76	70	97	62
ELF® [34]	Mixed	2004	1021/496**	40		0.102	0.78	87	51	n.a.
		2005			12	n.a.	0.89	n.a.	n.a.	n.a.
Fibrometer® [35]	Mixed	2007	598/503**	56		n.a.	0.89	80	84	82
NFS [40]	NAFLD	2008	733		27*	<-1.455 >0.676	0.82	43-77	97-97	68
BARD score [41]	NAFLD		669		30*	≥2	0.81	n.a.	n.a.	n.a.

HCV, chronic hepatitis C; HBV, chronic hepatitis B; NAFLD, non-alcoholic fatty liver disease; AUROC, area under ROC curve; Se, sensitivity; Sp, specificity; CC, correctly classified: true positive and negative; n.a., not available.
*F3F4.
**HCV patients.

Source: European Association for the Study of the Liver (EASL), *J of Hepatology,* 2015

2.10 CLINICAL IMPLICATIONS OF LIVER FIBROSIS

The onset and progression of liver fibrosis is insidious and can be clinically silent for decades. If untreated, liver fibrosis can worsen to cirrhosis, hepatocellular carcinoma and ESLD, which is incurable.[3,53] Liver transplantation is not a solution because not only is recurrence of HCV post-transplant almost universal,[103,104] but also, this scarce resource is not available to co-infected individuals in Canada. Successful HCV treatment would lead to removal of HCV, consequently reducing one of the main instigators of the inflammatory environment causing fibrosis. However, lack of access,[105] high costs, drug-drug interactions and possibility of re-infection in co-infected individuals[3,49] means that HCV treatment cannot be relied upon as the only approach to reducing

morbidity and mortality from HCV. Furthermore, even with HCV cure, advanced cirrhosis can be irreversible and other risks associated with HCC and portal hypertension remain.[65,66]

Thus, earlier detection and a better understanding of factors related to fibrosis progression are essential to not only identify persons at risk for liver disease progression but also to develop and target intervention strategies to reduce the development of end-stage liver disease and maximize treatment efficacy.

In summary, due to the accelerated progression of liver fibrosis resulting from immune impairment, heightened inflammation, and interrupted HIV treatment in co-infected individuals, characterizing the etiologic and predictive determinants in this group before onset of critical outcomes is especially pressing.

2.11 HCV TREATMENT

2.11.1 Benefits of HCV treatment and cure

Unlike HIV, HCV can be fully cured with treatment, as indicated by sustained virological response (SVR), traditionally defined as being HCV RNA-negative for 6 months after end of treatment. Recently, studies have shown that SVR at 12 weeks after completing treatment is an equally good indicator of cure. Attaining SVR can reduce mortality by 75-80% even in those who are co-infected or have cirrhosis.[106] Other benefits of SVR include improved liver histology such as regression of fibrosis[63] and cirrhosis[62,107] and reduced necroinflammation,[108] which lowers the probability of clinical complications like portal hypertension and splenomegaly. SVR

also can reduce the risk of liver cancer by 70% and the risk of liver-related mortality and transplantation by as much as 90%.[109,110] In addition to the liver, SVR can also improve symptoms and mortality due to extra-hepatic complications of HCV, such as cryoglobulinemic vasculitis.[111] Finally, HCV cure can also have public health benefits, by reducing transmission and thus reducing HCV incidence and prevalence.[112]

2.11.2 Challenges

Despite the multiple benefits associated with cure, until recently, HCV treatment uptake and success has remained low due to unpleasant side effects and length of treatment.[105] Since the early 2000s, standard treatment consisted of pegylated interferon alfa (PEG-IFN) and ribavirin administered for 24- to 48-weeks depending on the HCV genotype. Success rates were low at 42%–46% for genotype 1, the most common genotype in Europe and North America.[113,114] SVR with IFN based regimens is even lower in co-infected individuals (16-38%).[115] For genotypes 2 and 3, SVR was higher: 70-80% in HCV-monoinfected individuals but only 62% in their HIV-HCV co-infected counterparts.[116,117]

2.11.3 New era of HCV treatment

In 2011, direct-acting antiviral agents (DAAs) were approved which sidestepped many of these problems. Not only do the DAAs have much higher viral cure rates across genotypes (>80% in most of the trials), they also showed similarly high efficacies in co-infected individuals with fewer side effects. SVR remains somewhat lower in those with more advanced liver disease.[118,119] However, extremely high costs (between $50,000 and $120,000 for a single treatment course)[119,120] remain a major hurdle to widespread treatment uptake and HCV

eradication. Reimbursement by public and private payers has been restricted to date to those with advanced fibrosis in order to rationalize costs. In fact, the number of individuals treated for HCV in Canada has not changed much, remaining at approximately 5,000 patients per year.[29] Without increasing treatment uptake, modeling studies predict that the incidence of advanced liver disease (decompensated cirrhosis, HCC and other liver-related deaths) is expected to rise and peak between 2031 and 2035, which could potentially have implications for treatment efficacy as well, as treatment success can be lower in very advanced liver disease.[29]

Furthermore, the possibility of HCV re-infection among the men who have sex with men (MSM) and the active injecting drug users who make up the majority of the co-infected population in Canada will be a challenge. Therefore, characterizing etiologic and prognostic factors for liver disease can help identify higher risk individuals, thus aiding in making treatment decisions and prioritizing them for HCV treatment.

2.12 IFNL GENOTYPES

While some host genetic factors (*HLA-II DRB1/DQB1, HLA-II KIR-2DL3/2DL3*)[54,121,122] have been associated with favourable treatment responses or HCV outcomes, very few have been as strongly and consistently associated as the single nucleotide polymorphisms (SNPs) in *Interferon Lambda 3* (*IFNL3*, formerly *Interleukin 28B or IL28B* and also referred to as *IFNL4* in the literature). These SNPs have served as markers to both spontaneous and treatment induced HCV clearance as well as risk of liver disease, and are the main "exposures" of interest in Manuscripts 1 and 2.

2.12.1 *IFNL3* and HCV treatment response

The effect of *IFNL3* SNPs have been verified by 4 separate genome-wide association studies (GWAS) which examined over 500,000 SNPs in genes of patients enrolled in randomized trials for treatment of chronic HCV.[123,124,125,126,127] Though the racial makeup of their study population and their chip technology varied, all the studies found that SNPs around the *IFNL3* gene and its regulatory sequences were strongly predictive of favourable response to treatment, as indicated by sustained virologic response (SVR), i.e. being HCV RNA-negative for 6 months after treatment.

Ge et. al. found, for example, that in the SNP rs12979860, the C allele, which was more common among white than black Americans, was strongly associated with better treatment response, being most pronounced in individuals who had 2 copies of the allele. People with the CC genotype had SVR rates more than 2-fold higher than those who had 1 or 0 copies.[124] Among Japanese and Swiss patients, the T allele in the SNP rs8099917 was linked to better response, so that those with 1 or 0 copies of the T allele were 5-fold less likely to respond to treatment. [125,126] The strength of association of these SNPs with favourable treatment response also held in HIV-HCV co-infected populations,[128-131] though some researchers contend that the beneficial effect is most pronounced with HCV genotype 1.[131,123] The SNPs also continue to be a determinant of treatment response with the newer, more efficacious direct acting antiviral agents (DAAs).[132] For example, SNP rs12979860 can be used to identify candidates for shorter treatment duration[133] and *IFNL3* SNPs are predictive of SVR with interferon-free regimens.[134,135]

2.12.2 *IFNL3* and spontaneous clearance

Of those who are infected with HCV, approximately 20-45% spontaneously clear the infection without treatment, though this proportion is lower in co-infected individuals because the HCV-specific immune response is not as robust.[136] Identifying factors related to spontaneous clearance can aid in making treatment decisions, targeting public health interventions, and provide etiological clues about the host response to HCV infection. Some factors shown to be independently associated with higher rates of HCV clearance include female sex, Aboriginal or East Asian ethnicity, history of past clearance and infection with HCV genotypes 2 and 3.[8,51,52,126,137-139]

Recent studies have also found that *IFNL3* SNPs are associated with spontaneous clearance of HCV, in both mono[136-138,140] and co-infected[126] populations. The effect estimates indicated that 2 copies of the treatment responder alleles were also linked to higher rates of spontaneous clearance. With rs12979860, various studies in those with HCV infection only have found odds of spontaneous clearance over 3 times higher in those with 2 copies of the beneficial alleles compared to those with just 0 or 1 copy in adults and in children.[136,138,140,141] In HIV-HCV co-infected Swiss individuals, Rauch et. al. found that individuals who had TT genotype at rs8099917 were over 5 times more likely to spontaneously clear than those who had GT or GG genotypes, and this effect was similar when compared to mono-infected people.[126] Generally, rates of spontaneous clearance are lower in co-infected individuals, so strong associations in even this population point to the robustness of these findings.[123,142] However, none of the studies on *IFNL3* and spontaneous clearance have been done in HCV-infected Canadian populations,

which differ in their genetic makeup with an over representation of peoples of Aboriginal ancestry.

2.12.3 *IFNL3* and fibrosis progression

The relationship between IFNL genotypes and fibrosis progression is mixed and contradictory, owing to heterogeneity in study designs, measurement of outcomes and study populations. Some postulate no relationship[143-146] while others found, interestingly, a more rapid progression to fibrosis[1,12,147,148] and cirrhosis in those with the "responder" genotype (that is, the homozygous genotype linked with spontaneous clearance and improved treatment response).[11,149-152] Conversely, other studies found a more rapid progression or greater risk of severe fibrosis[153-155] or cirrhosis[155,156] associated with the nonresponder genotype or a link with steatosis, which is strongly correlated with liver fibrosis progression and inflammation.[146,157] Given the role of inflammation in liver fibrosis progression, the high degree of inflammation that persists in co-infected individuals despite HIV treatment,[74,158] and that IFNL SNPs have been consistently associated with pro-inflammatory immune responses, it is possible that interferon responses may play a role in the natural history of HCV infection and can affect necroinflammation, thus driving liver fibrosis progression.

Furthermore, as detailed in the next section, genotypes at IFNL might be linked to expression of specific genes involved in immune, cell death and survival pathways.[159] While these responses could make hepatocytes more efficient in clearing HCV, they could also, paradoxically, drive hepatic inflammation and liver disease upon viral persistence.

2.13 PUTATIVE MECHANISMS OF *IFNL3* SNPS AND ITS PROTEIN

It is not fully clear how these *IFNL3* SNPs affect HCV outcomes, though it may involve serving as markers for the direct antiviral activity of the protein IFNL3 or its pro-inflammatory manipulation of the immune system.

2.13.1 Antiviral activity of the IFNL3 protein

The gene, *IFNL3,* encodes the protein Interferon Lambda 3 (IFNL3), which belongs to a family of potent cytokines (Type 3 Interferons, referred to as Interferon Lambdas or IFNL) with antiviral properties against many viruses including HCV and HIV.[160-163] The antiviral effect of IFNL3 and the limited expression of its receptors have even made it a very promising therapeutic molecule. In Phase IIb trials, IFNL not only showed significant reductions in HCV viremia, but also had fewer side effects.[164] (Due to the rapid advance of DAAs, however, its development as a potential therapeutic agent for HCV has been halted).

Results from studies on the IFNL proteins themselves and effects on HCV outcome have been mixed, however. Since IFNL inhibits viral replication, it is logical to expect that favourable genotypes (those leading to better response to treatment and spontaneous clearance) would lead to higher levels of endogenous IFNL3 and lower HCV viral load. Some clinical and epidemiological studies have shown that higher amounts of IFNL are indeed associated with HCV clearance[127,125] but others have not found clear trends.[124,165,166]

2.13.2 Other causal variants

The SNPs rs12979860 and rs8099917 are not located in coding sections of the *IFNL3* gene and thus cannot directly cause changes in the IFNL protein structure. They could, however, be acting as markers for other causal SNPs by being in linkage disequilibrium (i.e. close enough to be traveling together) with them. These causal variants could actually affect function, stability, or responsiveness of the IFNL3 protein. For example, the variant rs368234815 (previously designated ss469415590) is linked to the encoding of protein IFNL4, which affects IFNL3 responsiveness.[167-170] Another SNP, rs4803217, could influence IFNL3 protein levels by affecting stability of mRNA transcripts.[171] Finally, two SNPs that might affect function or protein stability include rs28416813 and rs8103142, which might potentially affect receptor binding.[142] These last polymorphisms could also conceivably impact the stages where the RNA is translated to proteins (ie. mRNA splicing, translation, etc.)—these could be missed by microarray and RT-PCR analyses that have been performed. In populations where there is strong correlation (i.e. linkage disequilibrium or LD) between these causal SNPs and markers such as rs12979860 and rs8099917, there is a higher likelihood of HCV clearance, indicating a strong and responsive immune mechanism.

For this thesis, we were interested in the relationship between marker SNPs rs12979860 and rs8099917 and functional variant rs8103142 and two different HCV outcomes-spontaneous clearance for Manuscript 1 and significant liver fibrosis following HCV persistence in Manuscript 2. We analyzed the effects of the SNPs individually as well as as their joint inheritance as haplotypes.

2.13.3 Immune response regulation

Most relevant to us, IFN-lambda proteins may affect the course of HCV infection by their involvement in the innate and adaptive host response. Though IFNL3 has its own unique receptor, it shares many downstream signaling pathways with Type 1 interferons[172,131] and the two may act synergistically to exert antiviral effects.[131,173] Generally, IFNL3 signaling is pro-inflammatory,[123] leading to turning on of specific genes (interferon stimulated genes or ISGs) that are involved in early innate host immune responses.[173,160,174] Some studies have found that hosts with the major responder alleles at IFNL SNPs are more likely to turn on genes that are involved in antiviral immune responses or cell-death pathways.[159,175] While this gene expression may lead to effective HCV clearance, it may also lead to inflammation and liver damage if HCV persists.

The IFNL family has also been shown to be indirectly involved in the adaptive immune response, via upregulating IFN-gamma,[160] which recruits T-lymphocytes to the liver[176] and which has been found in higher levels in HIV-HCV co-infected individuals.[177,178] Furthermore, IFNL proteins have been found to promote Type 1 T-helper cell response (Th1),[123,179,180] which is critical in the resolution of HCV infection.[181-183] As a result, the *IFNL3* polymorphisms associated with poor treatment response and viral persistence may be indicative of an impaired/inappropriate activation of the adaptive immune response.

In summary, while the same *IFNL3* genotypes have been associated with both favourable treatment response and higher rates of spontaneous HCV clearance, the underlying biological mechanism is unclear and likely to be pro-inflammatory. Since the inflammatory environment in

co-infected individuals is heightened and since so little data exist on IFNL genotypes and onset of significant liver fibrosis, we are in a unique position to address this question.

2.14 INFLAMMATION AND FIBROSIS IN CO-INFECTED INDIVIDUALS

Chronic inflammation due to HIV infection is not just due to viral replication and cannot be reversed by effective cART which can suppress HIV replication completely.[184,185] Several mechanisms drive the inflammatory and fibrogenic response in HIV-HCV co-infected individuals and these include the following: the direct effect of specific virus gene products; immune dysregulation caused by depletion of CD4 T cells which play a critical role in both humoral and cytotoxic T-cell responses; bystander activation of T and B lymphocytes caused by an increased level of production of proinflammatory cytokines; the presence of opportunistic diseases; increased apoptosis of lymphocytes and hepatocytes; the preferential infection of central memory CD4 T cells which leads to the concentration of the antigenic load in central lymphoid tissues; and microbial translocation following disruption of the gut epithelial barrier.[78,184] (See also Figure 2.2)

2.14.1 Immune Activation due to microbial translocation

Microbial translocation, the movement of gut bacteria and microbial products through the disrupted intestinal epithelial barrier into systemic circulation,[184] has been linked to higher immune activation and disease progression in both HIV monoinfected[184,186] and co-infected populations.[187-189] One postulated mechanism involves CD4 depletion in the gut due to HIV which then leads to microbial translocation and HCV disease progression. Studies have shown

that this immune activation could lead to AIDS in HIV monoinfected individuals, and liver disease in co-infected individuals.[189]

2.14.2 Predicting Fibrosis using Immune Markers

Most of the data on inflammatory biomarkers comes from studies on HIV or HCV mono-infected individuals, where a biomarker was defined as a "characteristic that is objectively measured and evaluated as an indicator of normal biologic processes or pharmacological responses."[190,191] The inflammatory markers were studied using various study designs, but the main objective was to assess the association between the marker and different outcomes (e.g. clinical AIDS and/or mortality), effects of therapy or the ability to improve diagnosis of fibrosis. Because co-infected individuals have an accelerated liver disease profile and higher inflammatory response, we would expect similar and perhaps heightened, relationships between the markers and liver outcomes. Most of the prior research has focused on diagnosis (detection of existing outcome or disease stage) rather than prognosis (probability of developing outcome later).

However, strong associations (e.g. high odds ratios) in etiologic models do not automatically mean that a marker improves ability to classify or predict risk for individuals, which is most relevant for guiding both clinical and policy decisions.[192,193] That is, while measuring the association between specific immune markers and significant liver fibrosis is important, from a clinical perspective, what is also useful is assessing whether measures of the significantly associated markers truly improve ability to predict risk of the outcome in co-infected individuals beyond traditional predictors alone. Predicting risk of liver fibrosis is especially important as

HCV DAAs are highly expensive and very advanced liver disease could potentially lower treatment efficacy.

2.15 CLASSES OF IMMUNE MARKERS OF INTEREST

2.15.1. Cytokines

Cytokines are signalling molecules that control host immune response and can directly inhibit viral replication.[3] Both persistence of HCV and liver disease progression have been attributed in part to cytokine production, with several studies showing that increased concentrations of pro-inflammatory cytokines in HCV-infected patients are associated with an increase in liver disease.[3,177,194]

One of the key mediators of fibrogenesis is transforming growth factor beta (**TGF-β**), which not only triggers production of collagen but also prevents breakdown of extracellular matrix.[53,194] In hepatic stellate cells, TGF-β also catalyzes transition to myofibroblasts,[194] and is present at higher levels in cirrhosis,[195] all of which support its role in fibrosis onset and progression. Interventions that disrupted TGF-β synthesis or signaling also decreased fibrosis in experimental models.[53,77] This cytokine is also directly triggered by HIV proteins,[79,196] providing the biological mechanism that would lead to higher levels in co-infected individuals.

Serum levels of tumor necrosis factor alpha (**TNF-α**) are higher in HIV-positive individuals[3,197-199] and these cytokines act by upregulating viral replication.[3,200,201] TNF-α is produced by liver macrophages in response to viruses, alcohol, and LPS.[53] In livers of HIV-HCV co-infected

individuals, levels of TNF-α were even higher than in HCV monoinfected persons and these levels were reduced with HIV therapy,[177] demonstrating the heightened inflammatory effect of HIV on HCV infection. TNF-α promotes production of C-reactive protein (CRP) from the liver..

2.15.2. High-sensitivity C-reactive protein (hsCRP)

Widely studied in the context of cardiovascular disease, this protein is produced in the liver and is released during the acute phase of an infection, where it binds to damaged or infected cells to aid in their removal or repair.[202] CRP concentration has been described as a reproducible, dynamic reflection of ongoing tissue injury.[203] In HIV-positive individuals, higher levels of hsCRP were associated with a 3.5 fold higher odds of opportunistic disease than those with lower levels,[204] as well as increased risk of all-cause mortality.[205] In co-infected individuals, however, higher hsCRP levels might not necessarily correlate with more advanced liver fibrosis. In fact, one study found that lower CRP levels were associated with HCV co-infection.[206] This is biologically plausible as the liver may be producing less CRP as the fibrosis progresses to terminal stages.

2.15.3. Chemokines

Chemokines are a subset of cytokines that are chemoattractant, directing leukocytes and other immune cells through the body, such as through the inflamed liver. Of special interest to us (see below) are interleukin-8 (**IL-8**); monocyte chemotactic protein-1 (**MCP-1** or CCL2); macrophage inflammatory protein 1 (**MIP1α** or CCL3; **MIP1β** or CCL4); Regulated upon

Activation, Normal T cell Expressed and Secreted protein **(RANTES** or CCL5)**; CXCL9;** and **CXCL11.**

Studies have shown that in patients with chronic HCV, IL-8 serum levels are associated with disease progression and relate to interferon unresponsiveness.[207,208] Experiments have also shown that aside from being a chemoattractant, IL-8 has direct profibrogenic properties. In cell cultures, HCV proteins stimulated liver cells to produce IL-8 in a dose-dependent manner,[209] and the levels were heightened when HIV proteins were present,[210] providing biological evidence of the pro-inflammatory effect of HIV and HCV coinfections. MCP1 and RANTES have been found to stimulate fibrogenesis,[53,211] and while MCP1 mRNA is not expressed in healthy livers, it is secreted in livers of those with HCV.[212] RANTES and MIP1α levels are also higher in co-infected livers than monoinfected ones.[178] MIP-1α, MIP-1β and RANTES bind CCR5, receptors needed for HIV entry, and these pairs dominate during HCV infection. Data from cell cultures also show that HIV proteins, after binding with CCR5, directly stimulate profibrogenic activity in liver cells.[213] CXCL9 and CXCL10 have been correlated to histologic disease activity.[214]

2.15.4. Endothelial activation markers

This is a class of adhesion molecules that promote recruitment of immune cells and are elevated in HIV-infected individuals. This increase has been associated with HIV disease progression and higher HIV viral load.[53,215] Levels of these markers are also elevated in cirrhotic patients compared to healthy controls.[216] Of interest to us are soluble Vascular Cell Adhesion Molecule-1 (sVCAM-1) and soluble Intercellular Adhesion Molecule-1 (sICAM-1).

2.15.5 Soluble CD14 (sCD14)

One marker of microbial translocation is lipopolysaccharide (LPS), a component of Gram-negative bacterial cell walls.[184,217] LPS binds to sCD14, which acts as a co-receptor and is released in a soluble form upon monocyte activation, triggering further cytokine production and immune activation.[184,217] In HIV-infected individuals, those with the highest sCD14 levels had a 6-fold higher risk of mortality compared with those with lowest levels of sCD14, even after adjusting for other inflammatory markers, CD4 count and HIV viral load.[218] In HCV mono-infected individuals, higher sCD14 levels were linked to a lower HCV treatment response,[186] and this was also seen in HIV-HCV co-infected patients, where higher sCD14 levels were also linked with cirrhosis.[219] Another study in co-infected individuals reported an eight-fold higher odds of cirrhosis in those in the highest quartile of sCD14 compared to the lowest quartile.[189]

Therefore, after careful review of the literature summarized above, we selected these 13 pro-fibrogenic markers (TNF-α, TGF-β1, hsCRP, sCD14, IL-8, MIP-1α, MIP-1β, MCP-1, RANTES, CXCL9, CXCL 11, sICAM-1 and sVCAM-1) as best representative of underlying inflammatory mechanisms in liver fibrosis progression and thus potential candidates as predictive markers.

CHAPTER 3. Objectives and Overview of Methods

3.1 CANADIAN CO-INFECTION COHORT: THE SOURCE POPULATION

Research aims 1-3 were all nested in the Canadian Coinfection Cohort (CCC), which served as the source population. The CCC was established in 2003 and began recruitment at three university-based HIV clinics in Montreal, Canada. As an open prospective cohort of HIV-HCV co-infected individuals, the CCC has since grown to include 19 sites from 6 provinces across Canada, representing approximately 23% of the co-infected population under care.[43] Participating centres now include community-based clinics and out reach programs from both large and small urban centres across the country, as well as university-based HIV treatment programs. As of January 2015, the CCC consisted of 1,423 individuals. Manuscript 1 included data from the April 2013 data cut, Manuscript 2 analyzed data collected through January 2015, and Manuscript 3 data from July 2012. Funding is provided by the Canadian Institutes of Health Research (CIHR), the Fonds de recherche du Québec-Sante (FRQ-S) and the CIHR Canadian HIV Trials Network (CTN).

To be included in the CCC, patients must be over 16 years or older, give informed consent, and have serologic evidence for both HIV and HCV. For HIV, evidence includes ELISA with western blot, and for HCV, it can include detection of HCV- specific antibodies (positive by ELISA with recombinant immunoblot assay II (RIBA II) or enzyme immunoassay (EIA)) or detection of HCV–RNA (that is, eligibility does not differentiate between active or past HCV infection). At visits every six months, socio-demographic and behavioural information is collected using validated questionnaires, while medical and treatment information is collected by trained personnel along with plasma, serum and peripheral blood mononuclear cells (PBMC).

The study has been approved by research ethics boards at each of the participating institutions and the Community Advisory Committee of the CTN.[43]

3.2 *IFNL3* AND HAPLOTYPE ANALYSIS

3.2.1 *IFNL3* genotype: Independent variable for Manuscripts 1 and 2

Genotypic information was not routinely collected as part of the CCC protocol. It was measured specifically for Manuscripts 1 and 2, using never thawed plasma, serum or PBMC samples from visits with the most specimens as it is a time-fixed variable. The plasma (n=827) and serum (n=95) were processed the same day as the blood was drawn, while the PBMCs (n=10) were either processed the same day or the day after the blood draw, followed by storage in liquid nitrogen tanks. While samples were extracted at different time points and at different locations, they were all centrally stored at the Montreal Chest Institute at -80°C. The specimens were genotyped at the Bay Area Genetic Laboratory (BAGL) by laboratory staff who were blinded to the case or ethnicity status of each specimen. Genotyping was successful at each SNP (<3% failure at each SNP genotyped) and did not vary significantly by sample type or SNP. Details of the DNA extraction and genotyping process are outlined in the "Methods" section of Manuscript 1.

3.2.2 Hardy-Weinberg Equilibrium (HWE)

The principle of Hardy-Weinberg Equilibrium (HWE) states that if the population is very large and random mating occurs, allele and genotype frequencies remain unchanged (i.e. are in equilibrium) over time. Assumptions underlying HWE include random mating, lack of selection according to genotype, and absence of mutation, migration or gene flow.[220,221] Testing for HWE

has been suggested as a quality control measure as departures can be a sign of genotyping errors and can also interfere with correct haplotype estimation, since most software use algorithms that assume HWE.[221,222] Tests for HWE involve examining goodness-of-fit with Pearson's chi-squared or exact tests, which compares observed allelic frequencies to expected frequencies under HWE.

In Manuscript 1, HWE was assessed separately in Aboriginal peoples (n=140) and Canadian-born whites (n=620), and then among the white and Aboriginal non-cases (i.e. those who never cleared) separately, using the command –genhwi- in Stata 12. For Manuscript 2, HWE was assessed separately among white and Aboriginal non-cases as well. The software Haploview was also used to test HWE (results were concordant with Stata).

3.2.3 Haplotype Analysis

Since the markers at *IFNL3* are tightly linked to each other, analysing the cumulative effect of their inheritance as a haplotype instead of individual genotypes may serve as a better marker for the true causal variant. A haplotype is a sequence of alleles in neighbouring genes on the same chromosome that tend to be inherited together. Haplotype analysis, which estimates the joint effect of multiple SNPs as haplotypes on outcomes, can also help improve power and may define functional units whose effects cannot be predicted from what is known of the individual effect of each separate variant.[223] Because the SNPs we examined in Manuscripts 1 and 2 are tightly linked to each other, we reconstructed haplotypes and examined frequencies (Manuscript 1) as well as associations with the outcomes (both Manuscripts 1-2) using the HAPSTAT software.

Haplotypes are not directly observed and are probabilistically inferred from genotypes, especially for heterozygous subjects, who carry different alleles at two or more loci. For these individuals, the problem of "phase ambiguity" arises, which is often treated as a missing data problem. Based on study designs, several likelihood methods have been proposed to deal with this, which can be classified as prospective or retrospective likelihoods. Prospective likelihood, which pertains to the probability of a phenotype given a genotype, is most appropriate for cross-sectional and cohort designs. On the other hand, for a traditional case-control study, where sampling is conditional on outcome, a retrospective likelihood, which is the probability of a genotype given a phenotype, is more appropriate.[224]

The risk or causal haplotype is the haplotype constructed from the alleles of interest. For Manuscripts 1 and 2, the causal haplotype was TCT, which has the major alleles at each SNP, ordered by its location on chromosome 19: T at rs8103142, C at rs12979860 and T at rs8099917. For the haplotype-outcome relationship, several genetic models can be fitted based on the number of risk haplotypes inherited. Under the additive model, having two copies of a causal haplotype has twice the effect on the outcome versus having only a single copy, while in the dominant model, having one or two copies has the same effect. Under the recessive model, only having two copies of the causal haplotype will cause the outcome. And finally, in the codominant model, the effect of having two copies can be arbitrarily different from that of having a single copy. In the HAPSTAT results, the codominant effects are decomposed into additive and recessive components.

3.3 MEASURING SIGNIFICANT LIVER FIBROSIS

For both Manuscripts 2 and 3, the outcome of interest was time to significant liver fibrosis, measured with the aspartate AST-to-platelet ratio index (APRI). The APRI is an indirect serum measure that reflects alterations in hepatic function.[92] An APRI cutoff of 1.5 has been validated in our study population for detecting significant liver fibrosis (METAVIR stages F2-F4) with a sensitivity of 52%, specificity and positive predictive value (PPV) of over 99% and an AUC of 0.85 ± 0.06.[93] In addition, APRI cutoffs of 1.5 and 2 have also been shown to be associated with cirrhosis, other adverse liver and clinical outcomes, and death in our study cohort,[94] as well as in others.[95,96]

For Manuscript 2, APRI score is missing in approximately 5% of the visits in the study population and in Manuscript 3, values were missing in 4% of the analytic sample (subcohort and cases outside subcohort).

3.4 STATISTICAL ANALYSIS

Data in all three manuscripts were analyzed with Cox proportional hazards models. Details specific to each research aim are outlined in the manuscripts themselves and in Chapters 4-6. Below is a highlight of issues common to all three research aims.

3.4.1 Left truncation and selection bias

Since half of the cohort had been HCV infected for over 18 years at first visit, late entry was used to address the problem of left truncation. Left truncation occurs when time at risk does not

coincide with time in the study, i.e. subjects enter a study after their time zero or origin, and this may reduce power or cause selection bias.[225] For our study, time at risk for spontaneous clearance or liver fibrosis begins with HCV infection, which happened many years prior to study entry. The time axis was calendar time with the estimated date of HCV infection as the origin. With late entry, risk sets were thus matched on HCV duration and follow-up time. HCV duration was based on date of HCV seroconversion, if known, or on the year of first injection drug use or blood product exposure.[226]

In Manuscript 1, selection bias may occur if those with the favourable genotype cleared HCV more often and were less likely to enroll in the CCC. However, since eligibility was independent of RNA status (that is, entry to the CCC could be based on *either* positive viral load *or* positive HCV antibody test), selection bias due to recruitment criteria is unlikely. Furthermore, we also address this potential bias with a sensitivity analysis using inverse probability weights (IPW) to account for exclusion due to insufficient RNA tests or HCV treatment. The IPW approach thus also addresses selection bias caused by missing data[227] and is detailed further in Chapter 4, section "***4.4.2.1 Inverse Probability Weights (IPW).***"

With Manuscript 2, exclusion of those who have fibrosis at visit 1 may cause selection bias that may move the effect estimate towards the null. This is because those who are excluded may have more rapid progression to severe disease (e.g. before cohort entry) or be "exposed" cases, that is, cases with the responder genotype. Sensitivity analysis to address this potential limitation is described in more detail in Chapter 5.

3.4.2 Interval censoring and pooled logistic regression

For all our research aims, we used Cox proportional hazards models which treat time continuously and assume that exact event time is known. However, with the CCC data, data are collected at visits every six months so time of outcomes or time-varying covariates can only be narrowed to intervals of varying lengths—six months with no missing visits and longer with skipped visits. This is not a problem with the primary exposures and main confounders of Manuscripts 1 and 2, because host genotype and ethnicity do not vary with time. It is also not an issue with Manuscript 3, because our exposure is collected at first available visit, simulating data available in most clinical settings.

However, to address the possible impact of discrete time on the outcomes- spontaneous HCV clearance in Manuscript 1 and significant liver fibrosis for Manuscript 2- we conducted the following sensitivity analysis. In Manuscript 1, the pooled logistic regression with IPWs not only addressed missing data, but also included visits as indicator variables. The results are presented in Appendix Table 4.3 in Chapter 4 and did not differ greatly from those in Manuscript 1. Similarly, in Manuscript 2, the pooled logistic regression with IPWs addressed missing data, but also included visits as indicator variables, treating time as a discrete variable. Additionally, we used a polynomial transformation of visit numbers and included them in pooled logistic regression models in a study population similar to Manuscript 2. Results are presented in Appendix Tables 5.2 and 5.3 in Chapter 5 and are similar to those in Manuscript 2.

3.4.3 Proportional hazards assumptions

The proportional hazards assumption was tested in the full Cox model using the –stphtest-command in Stata. This command uses the unscaled Schoenfeld residuals to test proportionality for each individual predictor, and the scaled Schoenfeld residuals as a global test for the full model. A p-value>0.05 indicated that neither the model nor any of the covariates violated the proportionality assumption in any of the Cox models in Manuscripts 1-3.

CHAPTER 4. Distribution of *IFNL3* Genotypes in Canadian Aboriginal Peoples and Association with HCV Spontaneous Clearance (Manuscript 1)

4.1 PREFACE TO MANUSCRIPT 1

Previous studies have linked genetic markers near the *Interferon Lambda 3 (IFNL3)* gene to higher spontaneous and treatment-induced HCV clearance rates. Genotyping can thus help in making clinical and public health decisions, especially with new expensive HCV treatments and in populations at high-risk for HCV re-infection. The distribution of these markers also vary by ethnicity but have not been characterized in the Canadian co-infected population, which has a disproportionate number of Canadian Aboriginal peoples, who have a distinct and complex genetic history and are also reported to clear HCV more often.

Results from a few published studies indicated that Canadian Aboriginal peoples also have differences in immunological responses to HCV, though none had characterized distribution or function of *IFNL3* SNPs.[8-10] Because our data from the CCC was generalizable to the co-infected population under care and because Canadian Aboriginal peoples face many health challenges, we wanted to a) measure the association of specific IFNL markers with spontaneous HCV clearance and b) compare the frequency of IFNL SNPs between Canadian Aboriginal peoples and Canadian-born whites.

For these questions in Manuscript 1, we chose two SNPs that had been well described in other populations (rs12979860, rs8099917) as well as a third SNP (rs8103142) that some studies had suggested as a potential causal variant, since it led to amino acid changes in the IFNL3 protein.

We did not have any data on other more recently identified potential causal variants such as the IFNL4 SNP rs368234815 (previously designated ss469415590) or SNP rs4803217, which affects IFNL mRNA stability and is described in more detail in the section "**2.13 Putative Mechanisms of *IFNL3* SNPs and its Protein"** because these SNPs were identified after we began our studies. We analyzed the IFNL genotypes as a binary exposure—that is, if they have the "favourable" genotype consisting of 2 copies of the major alleles or not. We did not find any evidence of a dose-response where two copies of the major allele would have double or higher the effect of having only one copy (i.e. effect of rs12979860CC vs. rs12979860TT > rs12979860CT vs. rs12979860TT). Not only did we lack the power to detect such an effect, but also, the mechanism of the response of the beneficial alleles is likely recessive.[136]

We were able to successfully genotype at least 1 SNP in 80% or more of the study sample for each of the research questions. In addition to looking at the effects of each SNP individually, we also examined their distribution and joint effects as haplotypes.

This manuscript was published in the *International Journal of Molecular Sciences* in March 2015.

4.2 STUDY POPULATIONS IN MANUSCRIPT 1

1a. IFNL and spontaneous clearance: Because we wanted to verify spontaneous HCV clearance with higher certainty, we used 2 consecutive RNA-negative tests instead of just one. We also excluded all individuals who had ever been treated for HCV or initiated treatment under follow-up to ensure that clearance was not treatment-related. Because of the strict eligibility

criteria, at least half the source population was excluded. This could lead to selection bias if those excluded had a different exposure-outcome relationship or distribution from the included study sample. We assessed selection bias using various methods detailed in section "**4.4 Sensitivity Analysis for Manuscript 1.**" The effect estimates from that sensitivity analyses are very similar to those reported in the Manuscript and indicate that selection bias is likely not a problem.

1b. *IFNL3* Frequency Comparisons: From the April 2013 data cut of the CCC, 181 individuals reported any Aboriginal identity. However, to correct for population stratification (confounding by ethnicity), we restricted analysis to only those Aboriginal individuals that did not report any other ethnicity (n=157) and only those whites who were born in Canada (n=788). Samples were missing for 17 Aboriginal peoples and 168 Caucasians so our final study sample included 140 Aboriginal individuals and 620 whites. While misclassification is possible due to inaccuracies in using self-reported ethnicity, our allele frequency estimates for Canadian-born whites are very similar to those of Europeans in other study populations.[127,136]

We had only one category for Aboriginal peoples and did not break it down further into groups such as First Nations, Metis and Inuit. While the CCC is representative of a fraction of the co-infected population under care, it is not representative of the full national Canadian Aboriginal population. For example, while the majority of the Aboriginal persons in the CCC identified as First Nations (82% of the 181 who reported any Aboriginal ancestry), this proportion is higher than the national population where only 61% identified as First Nations.[228] Furthermore, we did not have any individuals who identified as Inuit, who make up 4.2% of the Canadian Aboriginal

population and mainly live in Nunavut and also in Labrador and the Northwest Territories, areas which are not included in the CCC.[228] Nevertheless, because the CCC includes centres from the provinces where most Aboriginal peoples live (Ontario, British Columbia, Alberta and Saskatchewan), our data captures frequency estimates from the majority of the co-infected Aboriginal peoples under care, and thus is relevant to treating clinicians.

4.3 DATA ANALYSIS IN MANUSCRIPT 1

4.3.1 Allele frequencies and Linkage disequilibrium measures

Stata and Haploview were both used to measure IFNL allelic frequencies and linkage disequilibrium (LD) separately in Aboriginal peoples and Canadian-born whites.

LD is said to exist between two alleles when they are correlated or travel together more than would be expected by random chance. It is often a function of distance between two loci, so when a mutation first occurs, it is often in perfect LD with its closest marker (D'=1). Pairwise measures include the Lewontin standardized disequilibrium coefficient D' and the squared correlation coefficient r^2 -- value of 1 indicates perfect LD, i.e. the two alleles are perfectly linked, while value of 0 means no LD or independence.

Patterns of LD can differ between two populations due to genetic events but also due to differing ancestries and histories.[229,230] In Stata 12, LD measures were determined using the command –pwld. The software Haploview was also used to measure LD values and their confidence intervals, and generate the heatmaps in Appendix Figure 4.1.[231,232]

The heatmaps show that LD is stronger in Canadian-born whites than in Aboriginal peoples, especially between SNPs rs12979860 and rs8103142, a potential functional variant since it leads to actual changes in the amino acid of the IFNL3 protein. This is reflective of the distinctive ancestry of Aboriginal peoples and also consistent with the findings in Manuscript 1 which suggest that the increased clearance in Aboriginal peoples is due to mechanisms that are not fully captured by the IFNL genotypes.

4.3.2 Haplotype Analysis

4.3.2.1 Inferring Haplotype Frequencies

Frequencies of haplotypes were estimated separately in Canadian-born whites and Aboriginal peoples using PHASE v.2.1. This software implements a Bayesian statistical method for reconstructing haplotypes from population genotype data.[233,234]

4.3.2.2 Comparing Haplotype Frequencies between Whites and Aboriginal Peoples

For Manuscript 1, the –hapipf- command in Stata 12 was used to compare haplotype frequencies in whites and Aboriginal peoples. This function uses the Expectation-Maximization (EM) algorithm to resolve phase uncertainty[235] and log-linear modeling and the likelihood-ratio test to test for association of haplotypes with group status (in our case, ethnicity). A p-value <0.05 indicates a rejection of the null of no association, that is, rejecting that haplotypic frequency is equal in the two groups.

4.3.2.3 Haplotype Association with HCV Spontaneous Clearance

For the haplotype-disease analysis, the Stata command -haplologit- was used to test the effect of TCT on spontaneous clearance, after adjusting for ethnicity. The command –haplologit- implements a modified retrospective semiparametric profile-likelihood method best suited to the following conditions: rare outcome, a single candidate gene in Hardy–Weinberg equilibrium, independence of genetic and environmental factors and a case-control design. It implements expectation-maximization (EM) and Newton-Raphson algorithms to estimate haplotype effects, similar to the HAPSTAT software, which was also used for Manuscript 2.[224,236]

Only results from the Stata analysis are reported in Manuscript 1. A comparison with the same analysis with the HAPSTAT software is displayed in Appendix Table 4.1). HAPSTAT software can provide estimates using a prospective likelihood more suitable for the cohort design used in Manuscript 1. The results are almost identical, as the underlying methods are very similar, but there are differences in the width of the confidence intervals possibly due to the different likelihood methods used for the study designs. More details on the likelihood methods are summarized in section "**3.2.3 Haplotype Analysis."** Both analyses also indicate that the additive model fit the data best, as demonstrated by the Akaike information criterion (AIC) values.

Details and results from sensitivity analyses for Manuscript 1 are in **Appendix 4A**, enclosed after the manuscript.

MANUSCRIPT 1: Distribution of *IFNL3* Genotypes in Canadian Aboriginal Peoples and Association with HCV Spontaneous Clearance

This manuscript was published in the *International Journal of Molecular Sciences* in March 2015.

Int. J. Mol. Sci. **2015**, *16*, 1-x manuscripts; doi:10.3390/ijms160x000x

OPEN ACCESS
International Journal of
Molecular Sciences
ISSN 1422-0067
www.mdpi.com/journal/ijms

Article

Favourable *IFNL3* Genotypes are Associated with Spontaneous Clearance and are Differentially Distributed in Aboriginals in Canadian HIV-Hepatitis C Co-Infected Individuals

Nasheed Moqueet [1], Claire Infante-Rivard [1], Robert W. Platt [1], Jim Young [2,3], Curtis Cooper [4], Mark Hull [5], Sharon Walmsley [6], Marina B. Klein [1,3,*]and the Canadian Co-infection Study Investigators [7]

[1]Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal H3A 1A2, Canada; E-Mails: nasheed.moqueet@mail.mcgill.ca (N.M.); claire.infante-rivard@mcgill.ca (C.I.R.); robert.platt@mcgill.ca (R.W.P.)
[2]Basel Institute for Clinical Epidemiology and Biostatistics, University Hospital Basel, Basel 4031, Switzerland; E-Mail: jyoung@uhbs.ch
[3] Department of Medicine, Division of Infectious Diseases/Chronic Viral Illness Service, Royal Victoria Hospital, McGill University Health Centre, 3650 Saint-Urbain Street, Montreal QC H2X 2P4, Canada
[4] The Ottawa Hospital—Research Institute, Ottawa K1Y 4E9, Canada; E-Mail: ccooper@toh.on.ca
[5] BC Centre for Excellence in HIV/AIDS, Vancouver BC V6Z 1Y6, Canada; E-Mail: mhull@cfenet.ubc.ca
[6] Toronto General Research Institute, University Health Network, University of Toronto, Toronto, ON M5G 2M9, Canada; E-Mail: sharon.walmsley@uhn.ca
[7] the Canadian Co-infection Study Investigators (see Appendix 1)

* Author to whom correspondence should be addressed; E-Mail: marina.klein@mcgill.ca; Tel.: +514-394-1934 (ext. 32306); Fax: +514-943-2092.

Academic Editor: Tatsuo Kanda

Received: 27 January 2015 / Accepted: 13 March 2015 / Published:

Abstract: Canadian Aboriginals are reported to clear Hepatitis C (HCV) more frequently. We tested the association of spontaneous clearance and three single nucleotide polymorphisms (SNPs) near the *Interferon-lambda 3 (IFNL3)* gene (rs12979860, rs8099917, functional variant rs8103142) and compared the SNP frequencies between HIV-HCV co-infected whites and Aboriginals from the Canadian Co-infection Cohort. HCV treatment-naïve individuals with at least two HCV RNA tests were included (n = 538). A spontaneous clearance case was defined as someone with two consecutive HCV RNA-negative tests, at least six months apart. Data were analyzed using Cox proportional hazards adjusted for sex and ethnicity. Advantageous variants and haplotypes were more common in Aboriginals than Caucasians: 57% *vs.* 46% had the rs12979860 CC genotype, respectively; 58% *vs.* 48%, rs8103142 TT; 74% *vs.* 67%, the rs12979860 C allele; and 67% *vs.* 64% the TCT haplotype with three favourable alleles. The adjusted Hazard Ratios (95% CI) for spontaneous clearance were: rs12979860: 3.80 (2.20, 6.54); rs8099917: 5.14 (2.46, 10.72); and rs8103142: 4.36 (2.49, 7.62). Even after adjusting for rs12979860, Aboriginals and females cleared HCV more often, HR (95% CI) = 1.53 (0.89, 2.61) and 1.42 (0.79, 2.53), respectively. Our results suggest that favourable *IFNL3* genotypes are more common among Aboriginals than Caucasians, and may partly explain the higher HCV clearance rates seen among Aboriginals.

Keywords: Aboriginals; genetic factors; Hepatitis C spontaneous clearance; Hepatitis C Epidemiology; HIV-hepatitis C co-infection; interferon lambda 3

---

1. INTRODUCTION

Among those infected with Hepatitis C (HCV), approximately 20%–45% spontaneously clear the infection without treatment.[136] This proportion is lower in HIV co-infected individuals due to weaker HCV-specific immune responses. Factors associated with higher HCV clearance include female sex, East Asian ancestry, and infection with HCV genotypes 2 and 3.[126,137-139,237-239] In Canada, Aboriginal ancestry has also been associated with spontaneous clearance.[8,240] HIV co-infection, on the other hand, is associated with lower rates of spontaneous clearance.[78] Of the host genetic factors linked with favourable HCV outcomes, the most consistent have been the

single nucleotide polymorphisms (SNPs) in the *Interferon-lambda 3* (*IFNL3*) gene, formerly *IL28B*.[38,39,54,241]

The SNPs near the *IFNL3* gene (rs12979860 and rs8099917) are strongly predictive of spontaneous clearance of HCV and favourable treatment response, in both mono[121,124,125,136-138,140,141,242,243] and co-infected [78,126] populations. The odds of spontaneous clearance are three times higher in those inheriting two copies of the beneficial alleles versus those with one or no copies. Other SNPs that have been proposed as the causal mechanism include ss469415590 (*IFNL4*) which impairs HCV clearance by turning on hepatic interferon-stimulated genes (ISGs) and reducing responsiveness to IFNL3[167,168,170] or rs8103142, which leads to amino acid substitutions in the IFNL3 protein.[142,244] The SNPs also continue to be a strong determinant of treatment response with the more efficacious direct acting antiviral agents (DAAs).[132] For example, SNP rs12979860 can be used to identify candidates for shorter treatment duration[133] and *IFNL3* SNPs are predictive of SVR with interferon-free regimens.[134,135]

Allele frequencies of the three *IFNL3* SNPs vary among ethnicities, with the favourable variants being almost universal in East Asians and rare in those of African ancestry.[124,125,136,140] However, no studies on *IFNL3* have been conducted in HCV-infected Canadian populations, which are unique in their genetic makeup. Aboriginals accounted for 15.2% of new HCV infections between 1999–2004,[34,35] despite representing only 4.3% of the Canadian population.[228] In the HIV-HCV co-infected population in the Canadian Co-infection Cohort (CCC), 16% of participants report Aboriginal descent. Despite higher HCV seroprevalence in mono-infected Aboriginals,[36] less than 5% had detectable HCV RNA compared with 75% of non-Aboriginal Canadians in one study.[37]

This finding suggests Aboriginals may have markedly increased rates of HCV clearance possibly due to specific immunological responses.[8,10,37-39,245] Aboriginals also respond better to HCV treatment in some studies[35,44] but not all.[246] The distribution of *IFNL3* SNPs in Aboriginals, which could differ from Caucasians due to their separate historical and genetic ties to populations in central Asia as well as their distinct migration and mixing patterns,[247,248] has never been reported. Given the burden of infection in the Aboriginal population, information on *IFNL3* would be valuable for tailoring treatment for co-infected Aboriginals who face numerous challenges in accessing care.

We studied the association between *IFNL3* SNPs (rs12979860, rs8099917 and rs8103142) and spontaneous clearance and compared the distribution of *IFNL3* SNPs in Canadian whites and Aboriginals.

2. MATERIALS AND METHODS

2.1. Source Population

The Canadian Co-infection Cohort Study (CCC, n = 1176), established in 2003, is an open prospective cohort of HIV-HCV co-infected individuals recruited from 18 centres across Canada, representing approximately 20% of the co-infected population under care.[43] For our study, we included data collected up until April 2013. To be included in the CCC, patients had to be over 16 years or older, give informed consent, be HIV infected (confirmed via ELISA with western blot), and have HCV infection or evidence of HCV exposure (HCV-antibody positive by ELISA with recombinant immunoblot assay II (RIBA II) or enzyme immunoassay (EIA) or if serologically false negative, HCV–RNA-positive). At visits every six months, socio-demographic,

medical and behavioural information was collected using validated questionnaires and biological samples were obtained and stored. The study has been approved by research ethics boards at each of the participating institutions.

2.2. Study Population and Covariates

For the spontaneous clearance study, we included individuals who had never been treated for HCV and who had at least two HCV RNA tests available (n = 538). Visits after HCV treatment initiation were censored. A spontaneous clearance case was defined as an individual who was HCV-RNA negative on two consecutive PCR tests, at least six months apart (Figure 3a). HCV RNA levels were measured at most visits (COBAS AMPLICOR HCV Test, version 2.0, Roche Diagnostics, Hoffmann-La Roche Ltd, Laval, Canada, lower limit of detection <50 $IU \cdot mL^{-1}$).

To compare the genotype distribution of the three *IFNL3* SNPs of interest between Canadian whites and Aboriginals, self-reported ethnicity was used. Participants self-identified as being of Caucasian, black, other (Asian or Hispanic Latino), or Aboriginal (First Nations, Metis, or Inuit) ethnicity. In the CCC, 15.6% reported some Aboriginal ancestry (n = 181), but analysis was restricted to those who did not report any other ancestry (n = 140). These results were compared to those from 620 genotyped Canadian-born whites (Figure 3b).

2.3. IFNL3 SNP Genotyping

Never thawed plasma and serum samples were processed and genotyped using a real-time PCR assay (Bay Area Genetic Lab). DNA was extracted using a modified Qiagen Mini Blood extraction protocol and the genotyping assay was developed for the LightCycler® 480 (Roche

Diagnostics, Laval, Canada) to cover bi-allelic SNPs rs12979860, rs8099917, and rs8103142 separately. Oligos were designed in-house and synthesized by TibMolBiol. Each real-time assay consists of a primer set to amplify the specific gene region and a set of hydrolysis probes representing each allelic variant (common and rare allele). Probes are dual-labeled with a 5' reporter dye (6' FAM or HEX) and 3' quencher dye (BBQ). Results were analyzed using Endpoint Genotyping and Abs Quant/2nd derivative analysis software to determine the genotype at each SNP.

Samples with ambiguous or no result for an individual SNP were retested using more of the extracted DNA. For samples with ambiguous or no results for all the SNPs, more DNA was extracted and the SNP assays repeated using the newer extraction.

2.4. Statistical Analysis

The software PHASE v2.1 was used for haplotype inference and distribution.[233,234] Haploview[231] and Stata version 12 were used to determine allele frequency distributions, Hardy-Weinberg Equilibrium measures, and linkage disequilibrium (LD) measures in Aboriginals and Canadian-born whites. Pearson's Chi-squared test was used to compare allelic and genotypic frequencies between the two subpopulations. The Stata command -hapipf- was used to test presence of LD in each group and also compare haplotypic frequencies between the two.

Due to the presence of censoring and left truncation, data were analyzed using Cox proportional hazards with adjustments for sex and ethnicity. The time axis was calendar time with the estimated date of HCV infection as the origin. Time in the analysis for each patient starts with

cohort entry. This method of late entry was used to address the problem of left truncation since half the Cohort had been HCV infected for over 18 years at first visit. HCV duration was estimated based on date of HCV seroconversion, if known, or on the year of first injection drug use or blood product exposure. Interaction terms were tested between the *IFNL3* genotype and sex as such interactions have previously been reported.[249] Due to the absence of HCV RNA in some available samples, HCV genotype information was missing in approximately 20% of the study population. Thus, HCV genotype was not included in the main multivariate analysis. In sensitivity analyses, Multiple Imputation by Chained Equations (MICE) was used to impute missing HCV genotypes.[250] Missingness for HCV genotype and other variables such as RNA tests, plasma samples, or *IFNL3* genotype was assumed to be at random. Stata 12 was used for all analyses (StataCorp LP, College Station, TX, USA).

A dominant model was used in the association analyses between genotype and spontaneous clearance. Subjects with one or two copies of the variant allele were grouped and contrasted with the wild-type genotype. For all three SNPs the homozygous wild-type genotype is considered favourable. Therefore, for rs12979860, genotype CC was compared with the CT and TT genotypes, whereas for rs8099917 and rs8103142 the TT genotype was compared with the TG and GG genotypes or with the TC and CC genotypes, respectively.

For the haplotype analysis, the Stata command-haplologit- was used to test the effect of TCT, the haplotype with the favourable alleles at all the SNPs (T at rs8103142, C at rs12979860 and T at rs8099917), after adjusting for ethnicity.

3. RESULTS

For the spontaneous clearance study, the study population included 46% of those enrolled in the CCC. The distribution of the sociodemographic and clinical features of included patients were similar to the CCC as a whole (Table 1). Half the participants had been infected with HCV for over 18 years. About 90% were HCV RNA positive at their first available test and very few were ever treated for HCV. Twenty-two of the 79 (28%) spontaneous clearance cases were HCV RNA positive at cohort entry and cleared HCV infection while under observation. The majority were male, older, with a history of injection drug use and were therefore likely infected with HCV before acquiring HIV. The majority were receiving antiretroviral therapy and had well-controlled HIV with good CD4 recovery (median CD4 count at baseline >350 cells·μL^{-1}). The 85 Aboriginals in the study population for the *IFNL3*-spontaneous clearance study were more likely to be female (60% *vs.* 32%) and their median CD4 count was slightly lower at 330 cells·μL^{-1} (*vs.* 365 cells·μL^{-1} in the study population). Spontaneous clearance cases were more likely to be female, Aboriginal, less likely to be infected with genotype 1, and more likely to possess the advantageous genotypes at all the *IFNL3* SNPs than those who did not clear spontaneously (Table 2).

3.1. Clearance and IFNL3

Eighty individuals (15%) cleared spontaneously; one was missing HCV duration information, so 79 cases were used in the analysis. *IFNL3* alleles were in Hardy-Weinberg equilibrium in both whites and Aboriginals ($p > 0.01$). The favourable genotypes at all the SNPs were associated with higher rates of clearance at a statistically significant level, with hazard ratios >3 in both

univariate and multivariate analyses (Tables 2 and 3). The rates of clearance did not change appreciably after adjustment for ethnicity or sex, indicating that the effect of the SNP is likely not related to, or mediated by, either of these variables.

In both univariate and multivariate analyses, being of Aboriginal descent was linked to a higher likelihood of spontaneous HCV clearance (Tables 2 and 3). The estimated clearance rates per 100 person-years (95% CI) for Aboriginals was 8.20 (5.17, 13.01) compared to 4.24 (3.28, 5.48) for whites, supporting the univariate estimate that Aboriginals were almost twice as likely to clear HCV compared with whites. There was no evidence that the *IFNL3*-spontaneous clearance relationship varied by ethnicity. Interaction terms between each SNP or haplotype with sex were not statistically significant (p-value = 0.9, results not shown). Females also appeared to have higher likelihood of clearing spontaneously, with a 50% higher clearance rate than males. Infection with HCV genotype 1 or 4 tended to be associated with a lower chance of clearance compared to infection with type 2 or 3 (HR = 0.56, 95% CI 0.29, 1.10 after multiple imputation for missing genotypes).

Patients carrying a haplotype with advantageous alleles from all three SNPs had a much greater rate of clearance than those lacking the haplotype, regardless of the mode of inheritance. A recessive model assumes that only having two copies of the beneficial haplotype raises the rate of clearance, while an additive model is closer to a dose response where two copies of the beneficial haplotype has a two-fold effect on the outcome compared to only one copy. Based on the AIC values, the additive model fit best, indicating that those with two copies of the haplotype

were over three times more likely to clear than those with a single copy of the haplotype (OR = 3.23, 95% CI: 2.66, 3.92).

3.2. IFNL3 Distribution in Aboriginals and Whites

The prevalence of the beneficial alleles and the genotypes (Figure 1) was more common in Aboriginals than white Canadians, especially at rs8103142 and rs12979860 ($p < 0.05$). At rs8099917, the differences in the allelic and genotypic frequencies did not reach statistical significance ($p = 0.53$ and $p = 0.24$, respectively).

The SNPs were in linkage disequilibrium in both Aboriginals and whites ($p < 0.001$ in each), though the magnitude of the measures differed. Among whites, higher r^2 and D' values indicated the SNPs were much more tightly linked to rs12979860 than in Aboriginals. In whites, D' values between favourable alleles in rs12979860 and rs8099917 were 0.85 (95% CI 0.79, 0.90) and between rs12979860 and rs8103142 were 0.93 (95% CI 0.89, 0.96). The corresponding values among Aboriginals were 0.68 (95% CI 0.53, 0.80) and 0.75 (95% CI 0.64, 0.83), respectively. In whites, r^2 between rs12979860 and rs8099917 was 0.39 and between rs12979860 and rs8103142 was 0.81, while in Aboriginals, the values were 0.32 and 0.55, respectively.

Haplotype analysis with PHASE estimates the prevalence of the most common haplotypes in each population. The top five most common haplotypes are presented in Figure 2. Estimates indicate that the haplotype containing the beneficial allele at all three SNPs (TCT corresponds to T at rs8103142, C at s12979860 and T at rs8099917) was most frequent in both Aboriginals and whites, though more common in the former: 67% of Aboriginals *vs.* 64% of whites. On the other

hand, CTG, which contains the unfavourable allele at each SNP, was more common in whites (18%) than Aboriginals (14%). Likelihood ratio test results also indicate that haplotypic frequency differed between the two groups ($p < 0.05$).

4. DISCUSSION

Beneficial SNPs near the *IFNL3* gene, linked to both spontaneous HCV clearance and treatment response, are distributed differentially in populations of different ancestry, and have never been characterized in the Canadian HIV-HCV co-infected population where Aboriginal people are overrepresented. As in other studies, we found clearance rates in our co-infected cohort to be lower than in HCV mono-infected populations.[126,251,252] Aboriginals cleared HCV infection more often than whites and also possessed higher frequencies of the advantageous genotypes, alleles and haplotypes. Furthermore, after adjusting for beneficial *IFNL3* genotypes, Aboriginals still tended to clear more often suggesting there may be additional factors that explain higher rates of spontaneous clearance in this population. Our results could have implications for treatment-related decisions, especially since Aboriginals are disproportionately affected by both HCV and HIV and more often carry favourable haptotypes that predict favourable HCV treatment responses.

SNPs of interest included those at rs12979860 and rs8099917, which are located in the non-coding region of the IFNL3 protein. We also studied the SNP at rs8103142, which leads to a nonsynonymous amino acid change (K70R), that could affect IFNL3 protein function[253,254] or interactions with other unknown factors involving viral control.[123] As in other populations, the favourable genotypes at all these SNPs were linked to higher rates of spontaneous clearance in

HIV-HCV co-infected Canadians. The beneficial genotypes were more frequent in Aboriginals and the differences reached statistical significance at rs12979860 and rs8103142. Canadian Aboriginals, like other Native American indigenous peoples, have complex ancestries but share links with Asian populations (i.e. Siberians and Mongolians), where the beneficial genotypes are almost universal.[39,247] We found the *IFNL3* allele frequencies in Canadian-born whites similar to those reported for European populations in other studies.[255,256]

The frequency of *IFNL3* haplotypes also differed between Aboriginals and Canadian-born whites. The TCT haplotype, which contains beneficial alleles at all three SNPs, was more common among Aboriginals than whites while the reverse was true for the CTG haplotype, which includes the unfavourable allele at each SNP. As with allelic frequencies, the haplotypic frequencies in whites were similar to those reported in other studies of Caucasian populations.[167] The frequencies of the beneficial alleles and haplotypes in Aboriginals were not as high as those reported in East Asians,[39,167,247] suggesting genetic divergence or reflecting mixing with European or other populations over time—for example, 93% of East Asians carry the TCT haplotype vs. 67% of Canadian Aboriginals, and only 5.5% carry the unfavourable CTG version vs. 14% of Aboriginals.[167]

As in prior studies,[8,35,37,257] in univariate analyses, Aboriginals were more likely to clear than whites. This association weakened after adjustment for *IFNL3* and sex. However, Aboriginals still appeared more likely to clear HCV than whites suggesting that *IFNL3* may be only one factor that contributes to increased clearance among Aboriginals. Differences in killer-cell immunoglobulin-like receptors (KIR) or IL-10 variants may explain higher HCV resolution

among Aboriginals.[38,39] We could not examine these factors, or test if they interact with the SNPs near *IFNL3*.

The effect of rs8103142 cannot be separated from rs12979860 at a population level, making it difficult to conclude if the SNPs were directly linked to the biological mechanism of spontaneous clearance or rather were behaving as markers for the true causal variant. The lysine-arginine (K70R) substitution caused by the rs8103142 polymorphism did not affect IFNL3 protein function in *in vitro* studies,[166,258] but since these studies involved a single experimental model within a short time frame (24 h), the authors could not rule out a major role for the rs8103142 variant in treatment response.[166,259] It is also possible that rs8103142 alleles are in LD with other causal variants such as ss469415590, which encodes Interferon Lambda 4 (IFNL4).[260] Some linkage ($r^2 \geq 0.6$) between the IFNL4 SNP and both rs12979860 and rs8103142 have been reported in other studies.[167,168]

The strengths of our study include a study sample that is representative of the co-infected population in Canada and included a large number of Aboriginal persons, so our findings will be directly relevant to treating clinicians. Our cases were carefully defined and sampled to reduce measurement error, thus providing reasonable estimates of relative rates of clearance.

Although ours is the largest study to date of spontaneous clearance in co-infected Canadians, we could not study other host or viral factors that could impact the IFNL3-spontaneous clearance relationship. As expected, infection with HCV genotype 1 or 4 was linked to a much lower rate of clearance, but accounting for HCV genotype did not affect the estimates for SNPs or

Aboriginal ancestry. Reduced power could also explain the lack of statistical significance of Aboriginal ancestry in the multivariate analysis. We also lacked the power to detect interactions between IFNL3 SNPs and sex which have been previously reported.[249]

Our study population was very similar to the CCC overall so our results should be generalizable to co-infected individuals receiving care in Canada. However, although the CCC attempts to recruit from diverse populations including patients with various risk factors and who are marginalized, persons not accessing care may differ from those included in our analyses. Those not under care may be more unstable, active injection drug users and more likely to be Aboriginal. Furthermore, the CCC does not represent the full diversity of all Aboriginal people in Canada, but is most reflective of co-infected Aboriginals in the most populous Canadian regions (Ontario, British Columbia, Quebec and Alberta). Our results may not be fully generalizable to Aboriginal persons outside these regions.

Another potential limitation is that the date of HCV infection used as the origin was approximate in most instances. If the error of this estimation was differential by ethnic group, that is, greater in Aboriginals than whites, for example, then it could bias the effect estimate. However, when we modeled time at risk using age, which is known with better accuracy, or used different modeling strategies (discrete time survival analysis or conditional logistic in a nested case control study), we did not obtain different results (not shown). Since most individuals enrolled in the CCC many years after they were infected, we cannot know with certainty whether this was a first or repeat clearance. Nevertheless, our findings still address whether *IFNL3* genotype is associated with spontaneous clearance in our study population.

In conclusion, HIV-HCV co-infected Aboriginals were more likely than whites to clear HCV infection and to carry the beneficial *IFNL3* genotypes and alleles linked to increased HCV clearance. Future studies should explore the mechanisms behind enhanced clearance among Aboriginals, including functional studies of *IFNL3 and IFNL4* genes or any other host factors that might enhance the immune response to HCV infection. Understanding the underlying biology of HCV clearance will ultimately help in making treatment decisions for Aboriginals who have urgent clinical needs.

ACKNOWLEDGMENTS

Special thanks to Rhyan Pineda, Laurence Brunet and Kathleen Rollet, for sample and data preparation; Connie Lisle and Ronald Carter, Bay Area Genetic Laboratory, McMaster University

This study was funded by the Canadian Institutes of Health Research (HEO-115694). The Canadian Co-infection Cohort is funded by Fonds de recherche Québec- Santé, Réseau SIDA/maladies infectieuses (FRQ-S), the Canadian Institutes of Health Research (MOP-79529), and the CIHR Canadian HIV Trials Network (CTN222); Marina B. Klein is supported by a Chercheur National career award from the FRQ-S. Nasheed Moqueet is supported by a doctoral research award from the Canadian Institutes of Health Research (CIHR DRA; 201010MDR) and National CIHR Research Training Program in Hepatitis C (NCRTP-Hep C).

Author Contributions

Nasheed Moqueet and Marina B. Klein designed the study. Nasheed Moqueet drafted the manuscript and conducted all the analyses with help from Claire Infante-Rivard, Robert W. Platt, Jim Young, and Marina B. Klein. All the authors have commented on the manuscript.

[7]Appendix 1: The Canadian Co-infection cohort investigators (CTN222) are: Jeff Cohen, Windsor Regional Hospital Metropolitan Campus, Windsor, ON; Brian Conway, PENDER Downtown Infectious Diseases Clinic, Vancouver, BC; Curtis Cooper, The Ottawa Hospital—Research Institute, Ottawa ON; Pierre Côté, Clinique du Quartier Latin, Montréal, QC; Joseph Cox, MUHC IDTC- Montréal General Hospital, Montréal, QC; John Gill, Southern Alberta HIV Clinic, Calgary, AB; Shariq Haider, McMaster University Medical Centre—SIS Clinic, Hamilton, ON; Aida Sadr, Native BC Health Center, St-Paul's Hospital, Vancouver, BC; Lynn Johnston, QEII Health Science Center for Clinical Research, Halifax, NS; Mark Hull, BC Centre for Excellence in HIV/AIDS, Vancouver, BC; Julio Montaner, St. Paul's Hospital, Vancouver, BC; Erica Moodie, McGill University, Montreal, QC; Neora Pick, Oak Tree Clinic, Children's and Women's Health Centre of British Columbia, University of British Columbia, Vancouver, BC; Anita Rachlis, Sunnybrook & Women's College Health Sciences Centre, Toronto, ON; Danielle Rouleau, Centre Hospitalier de l'Université de Montréal, Montréal, QC; Roger Sandre, Health Sciences North—The HAVEN/Hemophilia Program, Sudbury, ON; Joseph Mark Tyndall, Department of Medicine, Infectious Diseases Division, University of Ottawa, Ottawa ON; Marie-Louise Vachon, Centre Hospitalier Universitaire de Québec, Québec, QC; Steve Sanche, SHARE University of Saskatchewan, Saskatoon, SK; Stewart Skinner, Royal University Hospital & Westside Community Clinic, University of Saskatchewan, Saskatoon, SK; and David Wong, University Health Network, Toronto, ON.

Conflict of Interests

Marina B. Klein received grants from the Canadian Institutes of Health Research, Fonds de recherche en santé du Québec, Réseau SIDA/maladies infectieuses (FRQ-S), the National Institute of Health Research, Merck, ViiV Healthcare and Schering-Plough, consulting fees from Glaxo-Smith Kline and ViiV Healthcare, and lecture fees from Bristol-Meyers Squibb, Glaxo-Smith Kline and ViiV Healthcare. She also received fees for the development of educational presentations from Gilead, Glaxo-SmithKline and ViiV Healthcare. Curtis Cooper reports grants from Merck and Abbott; consulting fees from Merck and Vertex and lecture fees from Merck and Roche. Sharon Walmsley received grants, consulting fees, lecture fees, nonfinancial support and fees for the development of educational presentations from Merck, ViiV Healthcare, Gilead, Abbott, Tibotec, Janssen, Bristol-Myers Squibb and Boehringer Ingelheim. No conflict of interests were declared by the other authors.

Table 1. Baseline characteristics of *IFNL3*-spontaneous clearance study population compared with the Canadian Co-infection Cohort (CCC) source population.

Variables	Study Population *n* = 538	CCC *n* = 1176
Median follow-up time, years (IQR)	3.2 (1.7–4.6)	3.0 (1–4.4)
Mean age at baseline, years (SD)	44 (8.2)	45 (8.6)
Male, *n* (%)	368 (68)	864 (74)
Ethnicity, *n* (%)		
White	418 (78)	891 (77)
Black	15 (3)	45 (4)
Aboriginal	85 (16)	181 (16)
Other	15 (3)	44 (4)
Injection drug use ever, *n* (%)	472 (87)	944 (81)
Median HCV duration, years (IQR)	19 (11–25)	18 (10–26)
HCV RNA positive at first available test	481 (90)	889 (76)
HCV genotype 1, *n* (%)	304 (74) [a]	683 (73) [b]
Median CD4 counts, cells/μL (IQR)	365 (230–530)	420 (270–604)
On HIV therapy	415 (77)	957 (81)

[a] HCV genotype available in 410 individuals at visit 1; [b] HCV genotype available in 935 individuals at visit 1.

Table 2. Characteristics of spontaneous clearers compared to chronically HCV infected patients and univariate Cox proportional hazards analyses of spontaneous HCV clearance.

Variables	Spontaneous Clearers *n* = 79	Chronically Infected *n* = 462	Univariate HR (95% CI)
Aboriginal	18 (23%)	67 (15%)	1.91 (1.12, 3.25) [a]
Female	32 (41%)	142 (31%)	1.62 (1.02, 2.57)
HCV genotype, *n* (%)			
1	11 (52%)	293 (75%)	0.56 (0.29, 1.10) [b]
2	1 (5%)	17 (4%)	
3	9 (43%)	67 (17%)	
4	0	12 (3%)	
***IFNL3* genotypes**			
rs12979860			
CC, *n* (%)	53 (75%)	180 (43%)	3.89 (2.28, 6.63) [c]
CT, *n* (%)	15 (21%)	186 (44%)	
TT, *n* (%)	3 (4%)	57 (13%)	
rs8099917			
TT, *n* (%)	68 (88%)	256 (60%)	4.65 (2.32, 9.32)
GT, *n* (%)	9 (12%)	138 (33%)	
GG, *n* (%)	0 (0)	30 (7%)	
rs8103142			
TT, *n* (%)	59 (78%)	186 (45%)	4.23 (2.46, 7.28)
CT, *n* (%)	15 (20%)	182 (44%)	
CC, *n* (%)	2 (2%)	50 (12%)	

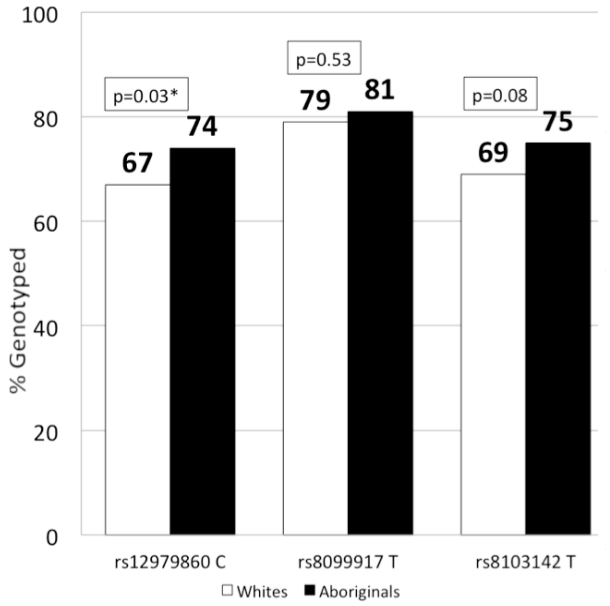
[a] The comparison shown is for Aboriginals *vs.* White Canadians (other ethnicities excluded); [b] HCV genotypes 1, 4 *versus* genotypes 2, 3, after multiple imputation; [c] CC *versus* non-CC.

Table 3. Multivariate results of the association of HCV spontaneous clearance with *IFNL3* genotypes.

Characteristic	Adjusted HR (95% CI) by *IFNL3* Genotype		
	rs12979860 CC	**rs8099917 TT**	**rs8103142 TT**
IFNL3 Genotype	3.80 (2.20, 6.54)	5.14 (2.46, 10.72)	4.36 (2.49,7.62)
Aboriginal Ethnicity *vs.* White	1.42 (0.79, 2.53)	1.50 (0.88, 2.57)	1.40 (0.79, 2.48)
Other Ethnicity *vs.* White	1.04 (0.21, 5.17)	1.19(0.25, 5.75)	1.01 (0.20, 5.05)
Female	1.53 (0.89, 2.61)	1.58 (0.97, 2.57)	1.55 (0.94, 2.56)

Figure 1. Distribution of favourable *IFNL3* genotypes and alleles in Canadian-born Whites and Aboriginals: (**a**) Frequency of favourable *IFNL3* alleles is higher in Aboriginals than Whites; (**b**) Frequency of favourable *IFNL3* genotypes is higher in Aboriginals than Whites. * $p < 0.05$.

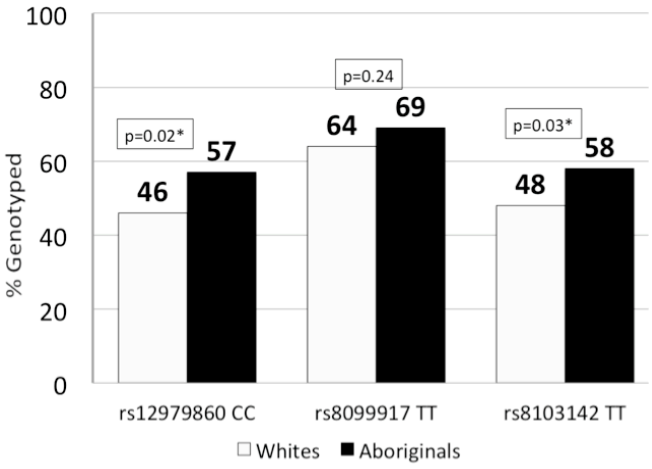
(a)



	Whites, n=620	Aboriginals, n=140
rs12979860 C	821/1220 (67%)*	206/278 (74%)
rs8099917 T	975/1228 (79%)	227/280 (81%)
rs8103142 T	846/1224 (69%)	197/264 (75%)

*Two alleles per person so denominator= 2 x # individuals

(b)



	Whites, n=620	Aboriginals, n=140
rs12979860 CC	281/610 (46%)	79/139 (57%)
rs8099917 TT	393/614 (64%)	97/140 (69%)
rs8103142 TT	293/612 (48%)	77/132 (58%)

Figure 2. Haplotype distribution in Canadian-born Whites and Aboriginals. Haplotypes containing the favourable alleles at all three SNPs (TCT) are more common in Aboriginals than whites while the opposite is true about haplotypes with the disadvantageous alleles (CTG). TCT = T at rs8103142, C at rs12979860 and T at rs8099917.

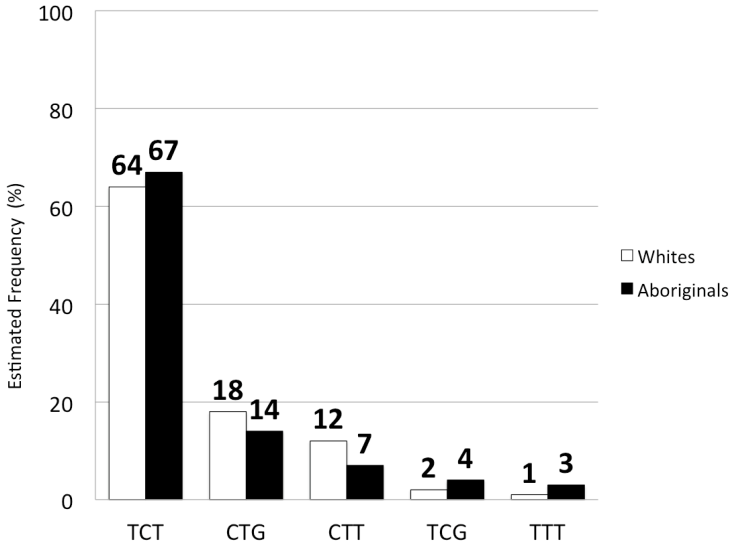
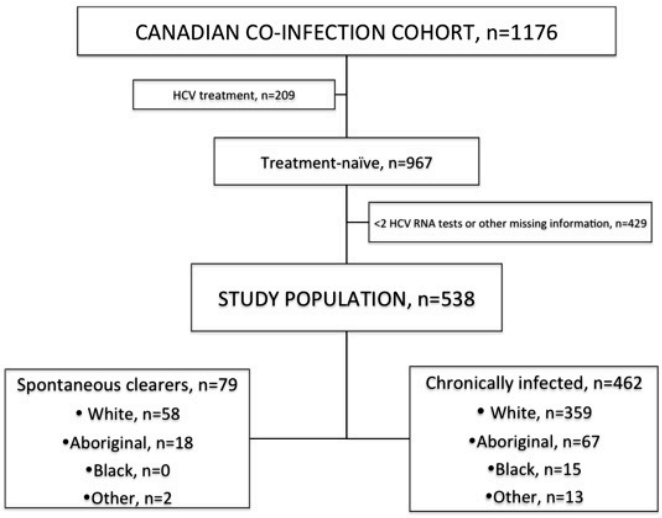
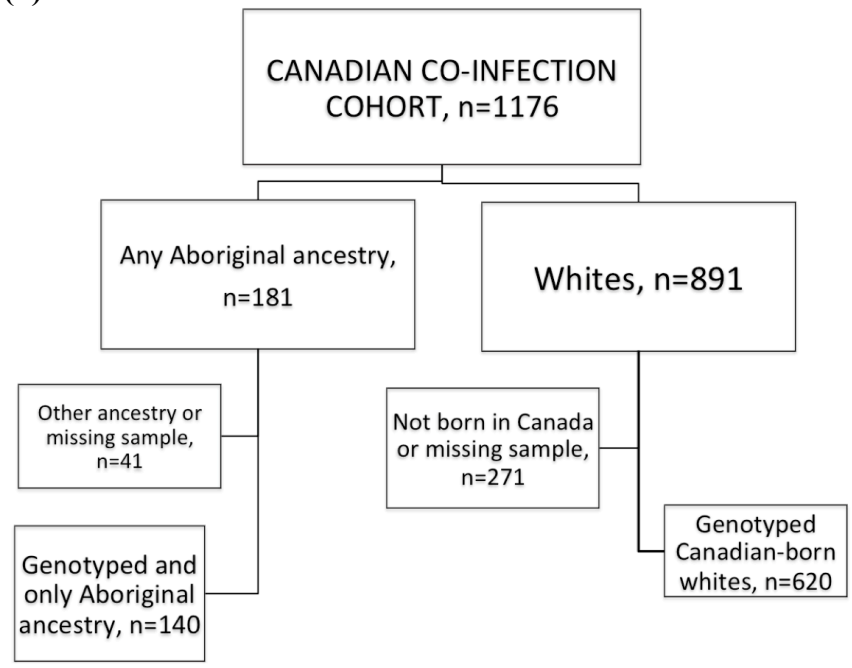


Figure 3. Study and source population: (**a**) Selection of study population for evaluating the association of *IFNL3* genotypes and rates of spontaneous clearance; (**b**) Selection of study population for comparison of *IFNL3* frequency distribution between Canadian Aborginals and Whites.

(a)



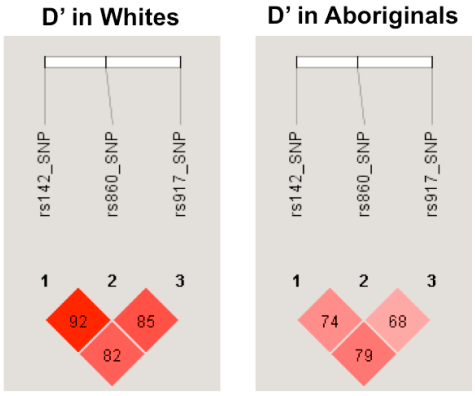
(b)



4A. APPENDIX FOR CHAPTER 4 (Manuscript 1)

Appendix Figure 4.1 Linkage Disequilibrium (LD) in IFNL SNPs between Canadian-Born Whites and Aboriginal Peoples from the Canadian Co-infection Cohort

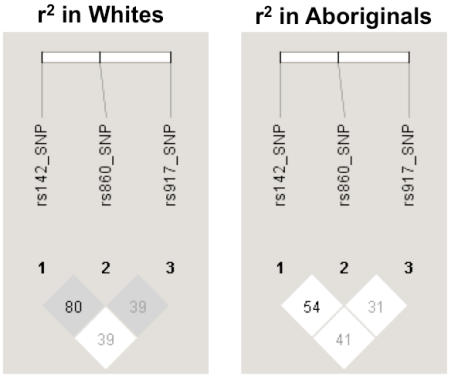
a) D' as LD measures between Canadian whites and Aboriginal peoples



LD with rs12979860 D' (95% CI)	rs8099917	rs8103142
Whites	0.85 (0.79, 0.9)	0.93 (0.89, 0.96)
Aboriginals	0.68 (0.53, 0.8)	0.75 (0.64, 0.83)

Dark red= High LD, Shades of pink= lower LD

b) r^2 as LD measures between Canadian whites and Aboriginal peoples



LD with rs12979860	rs8099917	rs8103142
Whites	0.39	0.80
Aboriginals	0.31	0.54

white=strong evidence of recombination, dark gray= strong evidence of LD

Appendix Table 4.1 Comparison of Haplotype Effect of TCT with Spontaneous HCV Clearance using Stata and HAPSTAT Software

Inheritance	Stata OR (95% CI)	HAPSTAT HR (95% CI)
Additive	3.23 (2.66, 3.92)	3.05 (2.03, 4.60)
Recessive	4.15 (3.30, 5.22)	3.90 (2.37, 6.40)
Dominant	3.37 (2.08, 5.46)	5.59 (1.75, 17.89)

Abbreviations: OR, odds ratio; CI, confidence interval; HR, hazard ratio.

4.4 SENSITIVITY ANALYSIS FOR MANUSCRIPT 1

4.4.1 Selection bias

Only 46% of the full Canadian Coinfection Cohort (CCC, n=1176) is included in the final study population (n=538) due to insufficient RNA tests or exclusion due to being on or having received HCV treatment. If those excluded have a different IFNL-clearance relationship or distribution than those included in the analysis, selection bias could result. This does not seem to be the case, as most variables, including IFNL genotypes, did not differ between the study population and the excluded sample (Appendix Table 4.2). Ethnicity, which is the main confounder of the *IFNL3*-spontaneous clearance association, also did not differ significantly between those included and excluded. Nevertheless, we used inverse probability weights (IPW) to account for potential selection bias as well as missing data (insufficient RNA tests).[227] More details are provided in section "***4.4.2.1 Inverse Probability Weights (IPW).***"

Appendix Table 4.2 Baseline Features of Source (CCC), Study, and Excluded Population, Manuscript 1

Variable	CCC, n=1176	Study population, n=538	Excluded, n=638
Follow-up time, years	3 (1-4)	3 (2-5)	1 (0.4- 3)
Age at baseline, years	45 (39-50)	44 (39-49)	45 (40-51)
Male	864 (74)	368 (68)	492 (77)
Ethnicity			
White	891 (77)	418 (78)	473 (75)
Black	45 (4)	15 (3)	30 (5)
Aboriginal	181 (16)	85 (16)	96 (15)
Other	44 (4)	15 (3)	29 (5)
HCV genotype 1	683 (73)[a]	304 (74)[b]	379 (72)[c]
HCV duration, years	18 (10-26)	19 (11-25)	18 (9- 27)
CD4 count, cells/μl	420 (270-604)	365 (230-530)	400 (257-580)
History of injection drug use	944 (81)	472 (87)	472 (75)
Ever treated for HCV	208 (18)	NA	208 (33)
HCV RNA positive at first available test	889 (76)	481 (90)	408 (83)
On HIV antiretroviral therapy	957 (81)	415 (77)	542 (86)
Homeless or Shelter	144 (12)	83 (16)	61 (10)
Born in Canada	981 (90)	490 (93)	491 (88)
Missing HCV genotype	240 (20)	132 (24)	108 (17)
ESLD	110 (9)	36 (7)	74 (12)
APRI ≥ 1.5	236 (21)	96 (18)	140 (23)
Died	141(12)	63 (12)	78 (12)
		IFNL genotypes	
rs12979860CC	437 (48)	234 (47)	204 (48)
rs8099917TT	599 (65)	325 (65)	275 (65)
rs8103142TT	451 (50)	246 (50)	206 (50)

Abbreviations: CCC, Canadian Co-infection Cohort; HCV, hepatitis C virus; NA, not applicable; ESLD, end-stage liver disease; APRI, aspartate aminotransferase (AST) to platelet ratio index, calculated as follows: [(AST/upper limit of normal)/platelet count (109 /L)] x 100.

Presented as n(%) or Median (Interquartile Range).

a. HCV genotype available in 935 individuals at visit 1 (80% of source populaton)
b. HCV genotype available in 410 individuals at visit 1 (76% of study population)
c. HCV genotype available in 525 individuals at visit 1 (82% of excluded population)

4.4.2 Missing Data

In Manuscript 1, missing data was addressed using inverse probability weights (IPW) and multiple imputation (MI). Both MI and IPW methods give very similar estimates of the IFNL-

spontaneous clearance link. The MI results also support a higher rate of clearance among Aboriginal peoples compared to whites.

4.4.2.1 Inverse Probability Weights (IPW)

IPW can account for missing data as well as potential selection bias due to exclusion from the study population of those with missing RNA tests and previous HCV treatment. In addition, using pooled logistic regression with indicator variables for study visits and IPWs accounted for study time as a discrete variable.

Weights for Censoring

The probability of having less than 2 RNA tests was estimated using a pooled logistic regression model with indicator variables for study visits and the following variables: centre ID, dichotomous *IFNL3* genotype, Aboriginal ancestry, sex, whether missing HCV genotype or not, history of injection drug use or not, homelessness, monthly income <=$1500, whether born in Canada or not, liver disease status (end-stage liver disease (ESLD) and significant liver fibrosis as measured by APRI ≥1.5), whether dead or alive, and whether using HIV antiretrovirals or not. The numerators of the stabilized inverse probability weights used fixed baseline variables in the regression and denominators used the time-updated counterparts. The stabilized weights were well-distributed when checked via histogram and had a mean of approximately 1. The variables were selected based on their potential to predict likelihood of missing 2 RNA tests.

Weights for HCV treatment

The probability of being treated for HCV was estimated using a pooled logistic regression model with indicator variables for study visits and the following variables: centre ID, dichotomous *IFNL3* genotype, Aboriginal ancestry, sex, monthly income <=$1500, BMI underweight or not, age as a restricted cubic spline with 4 knots, whether has a history of injection drug use or not, homelessness, whether born in Canada or not, and liver disease status (end-stage liver disease (ESLD) and significant liver fibrosis as measured by APRI ≥ 1.5). The numerators of the stabilized inverse probability weights used fixed baseline variables in the regression and denominators used the time-varying counterparts. As with the censoring weights, distribution was checked via histogram and the mean of the stabilized treatment weights was also approximately 1. These variables were selected based on their ability to predict likelihood of receiving HCV treatment.

Analysis with IPW

The final weights (mean=1), which are the products of the stabilized censoring weights and stabilized treatment weights, were used in a pooled logistic regression model for spontaneous clearance with binary IFNL genotype, sex and ethnicity and visits as indicator variables. The results were the same as those in Table 3 in Manuscript 1, except the effect of female sex on spontaneous clearance was statistically significant with IPWs, indicating that females have more than double the rates of spontaneous clearance as males.

Appendix Table 4.3 Sensitivity Analysis with Inverse Probability Weights in Pooled Logistic Models to Account for Missing and Excluded Data in Manuscript 1 (*IFNL3* and spontaneous clearance)

	rs12979860 CC	rs8099917 TT	rs8103142 TT
Adjusted OR (95% CI)	3.23 (1.41, 7.38)	5.30 (1.67, 16.79)	6.30 (2.66, 14.90)
Female	2.28 (1.05, 4.97)	2.35 (1.08, 5.13)	2.58 (1.21, 5.51)
Aboriginal vs. White	1.63 (0.71, 3.75)	1.63 (0.70, 3.77)	1.70 (0.71, 4.02)

Abbreviations: *IFNL3*, Interferon Lambda 3; OR, odds ratio; CI, confidence interval.

4.4.2.2 Missing Data in HCV genotypes: Multiple Imputation with MICE

To account for missing HCV genotypes a potential predictor of the outcome (missing in 25% of the study sample), Multivariate Imputation using Chained Equations (MICE) was implemented using Stata commands –ice-, -mim- and –mi estimate. Factors predictive of HCV genotype as well as the mechanism behind why it was missing were included. The following variables were used: centre ID, ethnicity, mode of HCV transmission, HCV genotype, binary *IFNL3* genotype, spontaneous clearance status, age, sex, whether injection drug user or not, homelessness, monthly income <=$1500 and the Nelson-Aalen estimate of the cumulative hazard function.[250] The results are similar to those in Table 3 in Manuscript 1, but those infected with HCV genotype 1 or 4 were 38% less likely to clear than those with types 2 or 3.

Appendix Table 4.4 Sensitivity Analysis with MICE to account for Missing Data in Manuscript 1 (*IFNL3* and spontaneous clearance)

	Adjusted HR (95% CI)
rs12979860 CC	3.15 (1.83, 5.42)
HCV genotype 1 or 4 (vs. 2 or 3)	0.62 (0.31, 1.26)
Female	1.47 (0.89, 2.41)
Aboriginal vs. White	1.50 (0.85, 2.64)

4.4.3 Uncertainty in HCV duration estimates for time axis

The time axis in the Cox model was calendar time with the estimated date of HCV infection as the origin. HCV duration was estimated based on date of HCV seroconversion, if known, or on the year of first injection drug use or blood product exposure.[226] As there was uncertainty in estimating date of HCV infection, we conducted a sensitivity analysis using age as the time axis (Appendix Table 4.5). Results are similar to those presented in Table 3 of Manuscript 1.

Appendix Table 4.5. Sensitivity Analysis using Age as the Time Axis in Manuscript 1 (*IFNL3* and spontaneous clearance) to Address Uncertainty in HCV Duration Estimates

	rs12979860 CC	**rs8099917 TT**	**rs8103142 TT**
Adjusted HR (95% CI)	3.84 (2.18, 6.76)	5.21 (2.51, 10.82)	4.29 (2.46, 7.47)
Female	1.43 (0.84, 2.43)	1.43 (0.85, 2.41)	1.48 (0.87, 2.51)
Aboriginal vs. White	1.43 (0.80, 2.56)	1.42 (0.83, 2.43)	1.33 (0.75, 2.37)

CHAPTER 5. IFNL Genotypes and Significant Liver Fibrosis (Manuscript 2)

5.1 PREFACE TO MANUSCRIPT 2

In Chapter 4, we concluded that specific genotypes at *IFNL3* were not only linked to higher rates of spontaneous HCV clearance but were also more common in Canadian Aboriginal peoples, who cleared more often, but possibly via other pathways as well. The mechanism behind *IFNL3* is still under investigation but is considered pro-inflammatory, which is why it could be linked to liver disease, which is caused by chronic inflammation. Because the risk of liver disease is higher in co-infected individuals, identifying the etiologic determinants in this group is especially pressing, as it can enable better screening and clinical decision-making, such as prioritizing higher risk individuals for expensive HCV treatments.

Several studies have examined the link between IFNL genotypes and liver disease such as significant liver fibrosis[1,11,152] or cirrhosis[149] but have found contradictory results. Some prior studies also had problems like low power[143] or selection bias.[144] Because eligibility criteria for the Canadian Co-infection Cohort was independent of HCV treatment or HCV RNA status, we did not enrich or deplete our study sample with those more likely to have the outcome or more likely to have the nonresponder genotypes, like studies nested in clinical trials. We also had longitudinal follow-up in individuals who had been infected with HCV for a long time, thus enabling us to capture the outcome in an etiologically relevant time window in those most at risk. And because the CCC is representative of the co-infected population under care, our results will be directly relevant to clinicians.

We used a prospective cohort study design to examine the link between time to significant liver fibrosis (APRI≥1.5) and the IFNL SNPs (rs12979860, rs8099917 and rs81013142) examined in Manuscript 1.

This manuscript has been accepted for publication in the *Journal of Infectious Diseases*.

5.2 STUDY POPULATION IN MANUSCRIPT 2

Because we wanted to examine the relationship of *IFNL3* genotypes on risk of liver fibrosis progression, we excluded participants who were no longer at risk (those who already had the outcome or had cleared HCV) or who were on HCV treatment, because treatment affects AST and platelet count, thus affecting the accuracy of APRI, the outcome measure. Of the 767 individuals who were potentially eligible, 63% (n=485) had genotypes available in all 3 SNPs, and were thus included in the final analytic sample. In the analysis reported in Manuscript 2, we used MICE to impute other missing covariates. Results where *IFNL3* genotype was also imputed from the full eligible study population were very similar to those reported.

5.3 DATA ANALYSIS IN MANUSCRIPT 2

5.3.1 Haplotype Associations

The software HAPSTAT[224,261,262] was also used to test the effect of TCT, the haplotype with the major alleles at all the SNPs (T at rs8103142, C at rs12979860 and T at rs8099917), on time to significant liver fibrosis, after adjusting for ethnicity. HAPSTAT also uses the EM and Newton-

Raphson algorithms to estimate haplotype effects,[224] followed by Cox proportional hazards regression for cohort designs.

Details and results from sensitivity analyses for Manuscript 2 are in **Appendix 5A,** enclosed after the manuscript.

MANUSCRIPT 2: IFNL GENOTYPES AND SIGNIFICANT LIVER FIBROSIS

This manuscript has been accepted for publication in the *Journal of Infectious Diseases*.

Responder Interferon Lambda genotypes are associated with higher risk of liver fibrosis in HIV-Hepatitis C Virus Co-infection

Running Title: IFNL & fibrosis in HIV-HCV Co-infection

Authors: Nasheed Moqueet,[1] Curtis Cooper,[2] John Gill,[3] Mark Hull,[4] Robert W. Platt,[1] Marina B. Klein,[1,5] Canadian Co-infection Cohort

[1]Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, QC Canada; [2] The Ottawa Hospital-Research Institute, Ottawa, ON Canada; [3] Southern Alberta HIV Clinic, Calgary, Alberta, Canada; [4] BC Centre for Excellence in HIV/AIDS, Vancouver, BC Canada; [5] Division of Infectious Diseases and Chronic Viral Illness Service, McGill University Health Centre, Montreal, QC Canada

Corresponding Author:

Marina B. Klein
Division of Infectious Diseases and Chronic Viral Illness Service
McGill University Health Centre – Glen Site
1001 Decarie Boulevard D02.4110
Montreal, Qc, Canada H4A 3J1
email: marina.klein@mcgill.ca
fax: 1-514- 843-2092
phone: 1-514- 843-2090

Word count: Abstract: **200** words Text: **3430** words

FOOTNOTE PAGE

Conflict of Interest Statement

None of the authors perceives a conflict of interest with regards to this study. Nevertheless, several co-authors have received either research support or speaker honoraria or have acted as consultants for a pharmaceutical company (details below). There was no pharmaceutical company support for this study.

Marina B. Klein received grants from the Canadian Institutes of Health Research, Fonds de recherche du Quebec – Santé (FRQ-S), Réseau SIDA/maladies infectieuses, the National Institute of Health Research, Merck, ViiV Healthcare, Janssen, Gilead, and Schering-Plough; consulting fees from ViiV Healthcare, Bristol-Meyers Squibb and AbbVie; and lecture fees from ViiV Healthcare and Gilead. Curtis Cooper reports grants from Abbvie and Gilead; consulting fees from Gilead, Merck and Abbive and lecture fees from Abbvie and Gilead. No conflict of interests were declared by the other authors.

Funding Statement

This work was supported by the Canadian Institutes of Health Research (HEO-115694). The Canadian Co-infection Cohort is funded by Fonds de recherche Québec- Santé, Réseau SIDA/maladies infectieuses (FRQ-S), the Canadian Institutes of Health Research (MOP-79529), and the CIHR Canadian HIV Trials Network (CTN222); Dr. Marina B. Klein is supported by a Chercheur National career award from the FRQ-S. Nasheed Moqueet is supported by a doctoral research award from the Canadian Institutes of Health Research (CIHR DRA; 201010MDR) and the Canadian Network on Hepatitis C (CanHepC, formerly National CIHR Research Training Program in Hepatitis C or NCRTP-Hep C).

Meetings where information was previously presented

July 2015 8th Conference of the International AIDS Society (IAS) Vancouver, BC, Canada

May 2015 Annual meeting of the Canadian Association for HIV Research (CAHR)

Toronto, ON, Canada

Corresponding Author:

Marina B. Klein
Division of Infectious Diseases and Chronic Viral Illness Service
McGill University Health Centre – Glen Site
1001 Decarie Boulevard D02.4110
Montreal, QC, Canada H4A 3J1
email: marina.klein@mcgill.ca
fax: 1-514- 843-2092

phone: 1-514- 843-2090

ABSTRACT

Background: Liver fibrosis progresses faster in HIV-HCV co-infected individuals. Interferon Lambda 3 (IFNL3) protein has both antiviral and pro-inflammatory properties. Genotypes at IFNL SNPs (rs12979860CC, rs8099917TT) are linked to higher HCV clearance, potentially via rs8103142. We examined the relationship between IFNL genotypes and significant liver fibrosis in HIV-HCV co-infection.

Methods: From the prospective Canadian Co-infection Cohort (n=1423), HCV RNA-positive participants with IFNL genotypes, free of fibrosis, end-stage liver disease and chronic Hepatitis B at baseline (n=485) were included. Time to significant fibrosis (AST-to-platelet ratio index (APRI) ≥1.5) by IFNL genotypes was analyzed using Cox proportional hazards, adjusting for age, sex, ethnicity, alcohol use, CD4 count, HCV genotype, GGT and baseline APRI. Haplotype analysis was performed, adjusting for ethnicity.

Results: 125 participants developed fibrosis over 1595 person-years (7.84/100 person-years, 95% CI: 6.58, 9.34). Each genotype increased fibrosis risk (aHR [95% CI]): rs12979860CC, 1.37 [0.94, 2.02]; rs8103142TT, 1.34 [0.91, 1.97]; rs8099917TT, 1.79 [1.24, 2.57]. Haplotype TCT was also linked with higher risk, 1.14 [0.73, 1.77].

Conclusions: IFNL SNPs rs12979860, rs8099917 and rs81013142 were individually linked to higher rate of fibrosis in HIV-HCV co-infection. IFNL genotypes may be useful to target HCV treatments to those who are at higher risk of liver disease.

INTRODUCTION

Despite effective HIV treatment and control, liver disease is especially serious in those with hepatitis C virus (HCV) and HIV co-infection. In co-infected individuals, liver fibrosis progression is accelerated, leading to cirrhosis, hepatocellular carcinoma (HCC) or end-stage liver disease (ESLD).[158] While curing HCV is now increasingly possible with the latest direct acting antiviral agents (DAAs), treatment uptake remains low and costs are high. HCV re-infection after cure remains a real problem among active injection drug users as well as among men who have sex with men (MSM), who make up the majority of the co-infected population in Canada. Furthermore, even with HCV cure, advanced cirrhosis can be irreversible and other risks associated with HCC and portal hypertension remain.[65] Because the risk of liver disease is higher in co-infected individuals, identifying etiologic determinants in this group is especially pressing, as it can enable better screening and treatment decision-making.

Among host genetic factors that could potentially affect liver fibrosis progression are the single nucleotide polymorphisms (SNPs) around the Interferon Lambda-3 (*IFNL3*) gene (formerly referred to as the *IL28B* gene)**.** In several studies, these SNPs (rs12979860 and rs8099917) were strongly predictive of favourable HCV treatment response and spontaneous clearance, in both mono[124,125] and co-infected[244,263] populations. The odds of spontaneous clearance or favourable treatment response were over 3 times higher in those inheriting 2 copies of the responder alleles compared to those with one or no copies. *IFNL3* genotypes are also predictive of treatment responses with DAAs, though the association is weaker.[132]

Reports on the relationship between IFNL genotypes and fibrosis progression are contradictory, owing to heterogeneity in study designs, measurement of outcomes and study populations. Some postulate no relationship[144] while others found, interestingly, a more rapid progression to fibrosis and cirrhosis in those with the "responder" genotype (that is, the genotype linked with spontaneous clearance and improved treatment response).[1,11] Conversely, other studies found a greater risk of severe fibrosis[153] associated with the nonresponder genotype. Given the role of inflammation in liver fibrosis progression, the high inflammation that persists in co-infected individuals despite HIV treatment,[158] and that IFNL SNPs have been linked with pro-inflammatory immune responses, it is possible that interferon responses play a role in the natural history of HCV infection and can affect necroinflammation, thereby driving liver fibrosis progression.[1,123]

We examined the association between the homozygous genotypes at 3 IFNL SNPs rs12979860, rs8099917, rs8103142 and the risk for developing significant liver fibrosis (defined as an APRI score ≥1.5) in the Canadian Co-infection Cohort.

METHODS

Source population: The Canadian Co-infection Cohort Study (CCC, n=1,423), established in 2003, is an open prospective cohort of HIV-HCV co-infected individuals recruited from 19 centres across Canada, representing approximately 23% of the co-infected population under care.[43] At visits every six months, socio-demographic, medical and behavioural information is

collected using validated questionnaires, along with plasma, serum and peripheral blood mononuclear cells (PBMC). For this analysis, we included data collected up until January 2015.

To be included in the CCC, patients must be 16 years or older, give informed consent, be HIV infected (confirmed via ELISA with western blot), and have HCV infection or evidence of HCV exposure--HCV-antibody positive by ELISA with recombinant immunoblot assay II (RIBA II) or enzyme immunoassay (EIA) or if serologically false negative, HCV–RNA-positive. The study has been approved by research ethics boards at each of the participating institutions.

Study population and covariates: Seven hundred and sixty-seven HCV RNA-positive participants free of fibrosis, end-stage liver disease and chronic Hepatitis B at baseline were eligible. HCV RNA levels were tested using qualitative tests (COBAS AMPLICOR HCV Test, version 2.0, Roche Diagnostics, Hoffmann-La Roche Ltd, Laval, Canada, lower limit of detection <50 IU ml-^{1}) and were available at most visits. Presence of existing significant fibrosis was determined using APRI≥1.5 at visit 1, while end-stage liver disease was a clinical diagnosis defined as presence of cirrhosis, ascites, portal hypertension, spontaneous bacterial peritonitis, encephalopathy, oesophageal varices or hepatocellular carcinoma. These diagnoses were verified using specific case report forms and validated centrally.[43] The presence of Hepatitis B surface antigen was used to determine chronic Hepatitis B infection.

Of 767 potentially eligible patients, analyses were restricted to those with available genotypes at all 3 SNPs (n=485) (Fig 1). The aspartate aminotransferase (AST) to platelet ratio index (APRI), a validated marker of liver fibrosis was calculated as follows: [(AST/upper limit of

normal)/platelet count (109 /L)] x 100. It consists of routinely available and non-invasive lab markers and was available for almost every visit. Significant fibrosis was defined as developing an APRI≥1.5 during follow-up.

Self-reported ethnicity, the strongest expected confounder, was used to categorize patients as white, black, other (Asian or Hispanic Latino), or Aboriginal (First Nation, Metis, or Inuit). Along with ethnicity, we adjusted for other risk factors including sex, alcohol use (currently drinking or not), baseline age (per decade), dichotomized CD4 count (>=350 vs. <350 cells/ul), HCV genotype 3 vs. non-3 (i.e. HCV genotypes 1, 2 and 4), log-transformed time-updated gamma-glutamyl transferase (GGT)[1] and baseline log-transformed APRI. Variables such as CD4 and HCV genotype were dichotomized at values found clinically relevant in other studies.[70,264]

IFNL genotypes: Never thawed plasma and serum samples were processed and genotyped using a real-time PCR assay developed by the Bay Area Genetic Lab (BAGL), as described previously.[263]

Statistical Analysis:

Data were analyzed using Cox proportional hazards after multiple imputation using Stata version 12.[265] The time axis was calendar time with the estimated date of HCV infection as the origin. Time in the analysis for each patient starts with cohort entry. This method of late entry was used to address the problem of left truncation since half the cohort had been HCV infected for 17 years at first visit. HCV duration was estimated based on date of HCV seroconversion, if known,

or depending on route of HCV acquisition, self-reported year of first injection drug use or first blood product exposure as a proxies for HCV infection acquisition.[226] Individuals were censored at last visit if they were lost to follow-up (defined as ≥1.5 years without visit, equivalent to missing 3 consecutive visits) or on date of death if prior to January 2015. Visits after HCV treatment initiation were censored.

Multiple Imputation by Chained Equations (MICE) was used to account for missing HCV genotype and other covariates. The reason for missing HCV genotype (~18%) and other variables such as RNA tests or plasma samples (<=15%) was assumed to be at random, as distribution of variables did not vary in the analytic population with all 3 SNPs compared to the source or study population. Before imputation, missingness of all other covariates in the final model was also <15%.

A dominant model was used in the association analyses between genotype and significant liver fibrosis. Subjects with one or two copies of the variant allele were grouped and compared with the wild-type genotype. For all 3 SNPs, the homozygous wild-type genotype has been linked to higher HCV clearance and better treatment response. Therefore, for rs12979860, genotype CC was compared with the CT and TT genotypes, whereas for rs8099917 and rs8103142, the TT genotype was compared with the TG and GG genotypes or with the TC and CC genotypes, respectively.

As reported in other studies, we also tested 3 types of interaction: product terms between sex and each genotype,[1] age (dichotomized at 40) and each genotype,[1] and HCV genotype and each

genotype.[152,266] To account for population stratification, we also tested product terms between genotype at each SNP and ethnicity. A sensitivity analysis was performed using age as the time axis to address uncertainty in the estimated date of HCV acquisition.

Stata 12 was used to evaluate Hardy-Weinberg Equilibrium and the software HAPSTAT[224] was used to test the effect of TCT, the haplotype with the major alleles at all the SNPs (T at rs8103142, C at rs12979860 and T at rs8099917), on liver fibrosis, after adjusting for ethnicity.

RESULTS

Four hundred and eighty-five patients representing 34% of those enrolled in the CCC met final inclusion criteria. One third of the exclusions were due to presence of existing liver disease at baseline (n=275 APRI≥1.5 and n=58 with prevalent ESLD). Sociodemographic and clinical factors of those included were similar to the Cohort as a whole (Table 1). Participants were estimated to have been infected with HCV for over 17 years on average, predominantly with HCV genotype 1. The majority of the study participants were Canadian-born white males with a median age of 44 years and a history of injection drug use (83%). Almost half reported drinking some alcohol at baseline. Most were on HIV antiretroviral therapy and had well-controlled HIV with good CD4 recovery (median CD4 count at baseline >350 cells/μl) and undetectable HIV viral load. The frequency of the responder IFNL genotypes was somewhat lower in the study population compared to the CCC as a whole. Alleles at each SNP were in Hardy-Weinberg Equilibrium ($p>0.05$).

One hundred twenty-five participants developed fibrosis, over 1595 person-years of risk (7.84 per 100 person-years, 95% CI: 6.58, 9.34). Those who developed significant liver fibrosis were more likely to be female, white, and injection drug users than those who never developed fibrosis (Table 2). They also were more likely to drink alcohol and had poorer HIV control, as evidenced by the higher proportion of antiretroviral interruptions and lower CD4 count at baseline. Individuals who developed fibrosis were also more likely to carry rs8099917 TT (i.e. the responder genotype), but not rs12979860 CC or rs8103142 TT. At baseline, fibrosis cases (i.e. those who eventually developed fibrosis) already had higher median APRI scores compared to those who never reached an APRI>1.5.

In univariate analysis, rs8099917 had the strongest association. Since both rs12979860 and rs8103142 are very close together and tightly linked (r^2=0.81 among whites in our cohort[263] and r^2>=0.85 in other studies[167,244]), their effect estimates were almost identical and thus were analysed separately. None of the other product terms between the SNPs and age, sex, HCV genotype or ethnicity indicated effect measure modification (p>0.05).

As with the individual SNPs, the results from the haplotype analysis indicated that those with a haplotype with the major alleles from all 3 SNPs (TCT) had a higher risk of fibrosis than those lacking the haplotype, regardless of the mode of inheritance. The log likelihood values are very similar (Table 4), but based on the dominant model, those with one or two copies of TCT had a 14% higher risk of fibrosis compared to those with no copies (HR 1.14, 95% CI: 0.73, 1.77).

Other than IFNL genotype, log-transformed GGT, which is a marker of oxidative stress, and baseline APRI had the strongest effects, with each log increase associated with a tripling of risk. Being female or of younger age was associated with a higher risk of fibrosis, after accounting for HCV duration. Being infected with HCV genotype 3 or currently drinking alcohol also raised risk of fibrosis by over 40% compared to non-drinkers or those infected with non-3 HCV genotypes (types 1, 2, or 4), respectively. Those with well-controlled HIV, evidenced by CD4 counts 350 cells/ul or over, were 30% less likely to develop significant fibrosis than those with counts less than 350 cells/ul (Table 3).

DISCUSSION

We found that co-infected persons without liver fibrosis at baseline who carry major alleles in any of the IFNL SNPs were at increased risk of developing significant liver fibrosis even after accounting for major known risk factors of fibrosis progression. This relationship between IFNL SNPs that have been associated with pro-inflammatory responses and risk of liver fibrosis was present independently of the baseline APRI score, suggesting the SNPs may be valuable markers for identifying patients who could benefit from curative HCV therapy.

SNPs of interest included those at rs12979860 and rs8099917, which are located in the non-coding region of the IFNL3 protein, and one at rs8103142, which leads to a nonsynonymous amino acid change. While genotypes at all SNPs were linked with higher fibrosis, rs8099917 had the strongest association, with the TT genotype almost doubling the risk of fibrosis. Previous studies examining the association between the SNPs and fibrosis progression, mainly in mono-

infected populations, have yielded mixed results, possibly due to low power[153] or selection bias, where study recruitment depended on failing previous HCV treatment or having advanced liver disease and thus possibly enriching by non-responder genotypes.[144] In contrast, ours is the largest longitudinal study of co-infected individuals where recruitment was independent of eligibility for HCV treatment.

Several cross sectional and cohort studies have suggested effect sizes similar to those we observed. In these other studies, the strongest relationships (OR=1.93 with rs8099917T) were those reported in co-infected patients,[11] though odds ratios for rs8099917 from mono-infected populations were also above 1.5.[152] A large study in patients without HIV infection (n=3,129 patients with 1500 fibrotic outcomes) reported that rs12979860CC and rs8099917TT were associated with over 60% increase in risk of fibrosis.[1] In this study, IFNL genotypes were predictive of fibrosis independent of disease etiology, as they were equally predictive in HCV as well as in those with Hepatitis B infection or non-alcoholic fatty liver disease (NAFLD).[1] This suggests that even with clearance of HCV after therapy, individuals with the IFNL responder genotypes may remain at higher risk for fibrosis, especially in the presence of other hepatotoxic behaviors or insults, many of which exist in drug-using or co-infected populations. Thus, these individuals may warrant closer follow up post-treatment for liver disease outcomes.

The mechanism of IFNL SNPs remains an active research area. The encoded protein, Interferon-Lambda 3 (IFN-λ3), is a Type III Interferon effective against viruses like HCV and HIV.[162,163] Additionally, IFN-λs turn on interferon-stimulated genes (ISGs) needed for viral control.[123] Some variants that have been proposed as the causal mechanism include rs368234815 (*IFNL4*)[167]

which affects IFN-λ3 responsiveness or rs8103142, which leads to amino acid substitutions in the IFNL3 protein and could affect interactions with other unknown factors involving viral control.[123,166] In populations where there is strong correlation (i.e. linkage disequilibrium or LD) between these causal SNPs and markers such as rs12979860 and rs8099917, there is a higher likelihood of HCV clearance, indicating a strong and responsive immune mechanism.

It is difficult to untangle the effects of these SNPs on protein function or fibrosis development without a functional study, though our findings are consistent with several possible mechanisms. The genotypes linked with higher fibrosis risk in our study could be serving as markers for inflammatory pathways as fibrosis progression is caused by heightened inflammation rather than actual HCV replication. For example, several studies have shown that the responder *IFNL3* polymorphisms are predictors of elevated histological inflammatory activity.[1,266] In addition, responder alleles at rs12979860 turn on genes involved in Natural Killer (NK) cell activation, resulting in apoptosis of infected hepatocytes and a pro-inflammatory environment.[267] Though these events are potent for clearing HCV, in the presence of viral persistence, as seen in the patients included in our analysis, they can also cause and exacerbate liver injury.

Results from a few in vitro studies[166,258] indicate that the IFNL3 protein function is unaffected by the K70R (lysine-arginine) amino acid substitution tagged by rs8103142. However, one of the studies only examined this in a single experimental model within a short time frame (24 hours), and thus the authors did not rule out a major role for the Lys70Arg variant in treatment or immune response.[166] Nevertheless, it is also possible that the alleles from rs8103142 or rs12979860 are linked with other causal variant(s) in the *IFNL3* region, such as rs368234815,

which encodes Interferon Lambda 4 (IFNL4).[167] The antiviral activity of IFNL3 and IFNL4 is linked with higher levels of interferon-stimulated gene (ISG) expression,[167] and though this has been linked with improved HCV clearance, it is not known whether this vigorous immune response contributes to fibrosis progression. It is also possible that as the SNP at rs8099917 is less tightly bound to both IFNL3 and IFNL4 (r^2=0.39 among whites in our Cohort[263] and r^2~0.40 in other studies[167]), it could be tagging the effect of other causal variants from fibrogenic pathways.

Our study population was selected from one of the largest prospective cohorts of HIV-HCV coinfected individuals in the world and one which was representative of the Canadian co-infected population, including marginalized groups such as people who inject drugs and Aboriginal people. Because of regular longitudinal follow-up, we were able to measure risk factors of fibrosis progression that could potentially be combined into a progression index and investigated in future studies. By including those who acquired HCV remotely, we were also able to avoid the "referral bias" of previous retrospective studies. Those studies overestimated the severe outcomes of HCV infection (cirrhosis, hepatocellular carcinoma, and death) and did not allow for the examination of those who spontaneously recover from their infection or have milder forms of the disease.[268] Our study thus has characteristics of a prospective cohort but allows long-term follow-up that rarely can be achieved in prospective studies. Our study population was very similar to the CCC overall so our results should be reasonably generalizable to co-infected individuals receiving care in Canada. However, although the CCC attempts to recruit from diverse populations including patients with various risk factors and who are marginalized, persons not accessing care may differ from those included in our analyses.

A potential limitation in our study is the possibility of selection bias induced through the exclusion of cases of spontaneous HCV clearance and prevalent liver disease, both of which are more likely to have the genotypes of interest. However, results from an analysis in datasets with limited or no exclusions were very similar to those reported in Table 3, indicating that any selection bias was likely negligible (results not shown).

Another important issue is the possibility of residual confounding by ethnicity because of the diverse ancestries of our participants. We attempted to address this by adjusting for self-reported ethnicity. Furthermore, after restricting analyses to whites only or stratifying for ethnicity in sensitivity analysis, the effect estimates obtained were similar to those in Table 3 although less precise (not shown), suggesting that such confounding is unlikely to have been great.

Another possible limitation in our study is the use of the APRI score to measure the outcome. Liver biopsy, the gold standard for measuring liver fibrosis, is invasive, risky and subject to measurement error, thus making it impractical for longitudinal research purposes. Without biopsy samples, we were also unable to assess degree of hepatic necroinflammation. Fibroscan data was not available in sufficient numbers of patients to permit longitudinal analyses. However, an APRI cutoff at 1.5 (corresponding to F2 in the METAVIR scoring system) has been validated in our study population for detecting significant liver fibrosis with a sensitivity of 52%, specificity and positive predictive value (PPV) of over 99% and an AUC of 0.85 ± 0.06.[93] In addition, APRI cutoffs of 1.5 and 2 have also been shown to be associated with cirrhosis, other adverse liver and clinical outcomes, and death in our cohort,[94] as well as in other studies.[95,96]

Using an APRI ≥1.5 also allowed us to potentially capture the etiologically relevant transition to fibrosis stage F2, as IFNL genotypes were reportedly more important in earlier fibrosis transitions (F0-F1 and F1-F2) rather than later ones (F2-F3 and F3-F4).[1] Alternative non-invasive markers such as FIB-4 (> 3.25) correspond to more advanced fibrosis stage (F3 and higher) missing the F1-F2 transition.[97] Applying FIB-4 to these analyses therefore resulted in lower sample size and less precise estimates with the strongest association at rs8103142, aHR (95% CI) 1.14 (0.71, 1.83). Finally, measurement error in APRI is likely independent of IFNL genotype (the main exposure group) and would be non-differential and thus potentially underestimate a true causal effect.

There was uncertainty in the estimate of the date of HCV acquisition, which was approximate in most instances and used as the origin. Because this date preceded cohort entry by many years, it also led to left truncation, which we addressed by using delayed entry in our analysis. Moreover, results from the sensitivity analyses using age as a time axis were very similar to those in Table 3 (not shown). Finally, while other risk factors in our results such as alcohol and HCV genotype 3 are consistent with other studies,[54,55,70,264] we lacked the power to detect the interaction of IFNL genotypes with sex, age, HCV genotypes or ethnicity.[1,152]

In conclusion, our results suggest that the homozygous genotypes at the IFNL SNPs rs8099917, rs12979860 and rs8103142 individually increased the risk of significant liver fibrosis in HIV-HCV co-infected Canadians. The association of rs8099917-TT was strongest, with almost a doubling of risk. Our findings are consistent with a heightened inflammatory profile and could

help identify higher risk individuals who would benefit the most from expensive HCV DAAs, before liver disease advances to the "point of no return."

Funding

This work was supported by the Canadian Institutes of Health Research (HEO-115694). The Canadian Co-infection Cohort is funded by Fonds de recherche Québec- Santé, Réseau SIDA/maladies infectieuses (FRQ-S), the Canadian Institutes of Health Research (MOP-79529), and the CIHR Canadian HIV Trials Network (CTN222); Dr. Marina B. Klein is supported by a Chercheur National career award from the FRQ-S. Nasheed Moqueet is supported by a doctoral research award from the Canadian Institutes of Health Research (CIHR DRA; 201010MDR) and the Canadian Network on Hepatitis C (CanHepC, formerly National CIHR Research Training Program in Hepatitis C or NCRTP-Hep C).

Acknowledgments

Special thanks to Rhyan Pineda and Kathleen Rollet, for sample and data preparation; Claire Infante-Rivard, for expertise on statistical and genetic epidemiology analyses; Connie Lisle and Ronald Carter, Bay Area Genetic Laboratory, McMaster University.

The Canadian Co-infection cohort investigators (CTN222) are: Drs. Jeff Cohen, Windsor Regional Hospital Metropolitan Campus, Windsor, ON; Brian Conway, Vancouver Infectious Diseases Research and Care Centre, Vancouver, BC; Curtis Cooper, The Ottawa Hospital Research Institute, Ottawa ON; Pierre Côté, Clinique du Quartier Latin, Montréal, QC; Joseph Cox, Montréal General Hospital, Montréal, QC; John Gill, Southern Alberta HIV Clinic, Calgary, AB; Shariq Haider, McMaster University, Hamilton, ON; Marianne Harris, St. Paul's Hospital, Vancouver, BC; David Haase, Capital District Health Authority, Halifax, NS; Mark

Hull, BC Centre for Excellence in HIV/AIDS, Vancouver, BC; Julio Montaner, St. Paul's Hospital, Vancouver, BC; Erica Moodie, McGill University, Montreal, QC; Neora Pick, Oak Tree Clinic, Children's and Women's Health Centre of British Columbia, University of British Columbia, Vancouver, BC; Anita Rachlis, Sunnybrook & Women's College Health Sciences Centre, Toronto, ON; Danielle Rouleau, Centre Hospitalier de l'Université de Montréal, Montréal, QC; Roger Sandre, HAVEN Program, Sudbury, ON; Joseph Mark Tyndall, Department of Medicine, Infectious Diseases Division, University of Ottawa, Ottawa ON; Marie-Louise Vachon, Centre Hospitalier Universitaire de Québec, Québec, QC; Sharon Walmsley, University Health Network, Toronto, ON; and David Wong, University Health Network, Toronto, ON

Author Contributions

N.M. and M.B.K. designed the study. N.M. drafted the manuscript and conducted all the analyses with help from R.W.P. and M.B.K. All of the authors have commented on the manuscript.

Table 1. Baseline Characteristics of the Canadian Co-Infection Cohort and Study Population

Characteristic	Study population, n=485	CCC, n=1423
Follow-up time, years	5 (3-6)	3 (1-6)
Age at baseline, years	44 (38-49)	45 (39-50)
Male	335 (69)	1,026 (72)
White	369 (76)	1,054 (74)
Born in Canada	420 (91)	1,114 (90)
History of injection drug use	403 (83)	1,143 (81)
Current alcohol drinker	234 (48)	744 (52)
APRI	0.52 (0.37- 0.78)	0.63 (0.39-1.24)
Undetectable HIV viral load[a]	280 (59)	876 (63)
On HIV antiretroviral therapy	382 (79)	1,178 (83)
CD4 count, cells/µl	381 (260-550)	398 (250-575)
HCV duration, years	17 (10-25)	18 (10-26)
HCV genotype 1	313 (77)	852 (74)
IFNL genotypes		
rs12979860CC	202 (42)	437 (48)
rs8099917TT	297 (61)	599 (65)
rs8103142TT	214 (44)	451 (50)

Abbreviations: APRI, aspartate aminotransferase (AST) to platelet ratio index, calculated as follows: [(AST/upper limit of normal)/platelet count (109 /L)] x 100; HCV, Hepatitis C virus; IFNL, Interferon Lambda.

a. Limit of detection, <=50 copies/ml

Presented as n(%) or Median (Interquartile Range).

Table 2. Baseline characteristics of patients developing and not developing fibrosis in follow-up

		Fibrosis (n=125)	No Fibrosis (n=360)
Age, years		43(37-48)	44 (38-49)
Female		42 (34)	103 (29)
Ethnicity			
White		100 (80)	269 (75)
Black		3 (2)	14 (4)
Other		4 (3)	12 (3)
Aboriginal		18 (14)	65 (18)
History of injection drug use		111 (89)	292 (81)
Current alcohol use		69 (55)	165 (46)
Baseline APRI		0.73 (0.45- 0.98)	0.49 (0.35- 0.71)
CD4 ≥350 cells/µl		62 (50)	214 (60)
HIV ARV interruption		11 (9)	20 (6)
HCV duration, years		18 (12-24)	17 (9-25)
HCV genotype 3		18 (17)	45 (15)
		IFNL genotypes	
rs12979860			
	CC	53 (42)	149 (41)
	CT	52 (42)	160 (44)
	TT	20 (16)	51 (14)
rs8099917			
	TT	84 (67)	213 (59)
	GT	37 (30)	117 (33)
	GG	4 (3)	30 (8)
rs8103142			
	TT	56 (45)	158 (44)
	CT	47 (38)	159 (44)
	CC	22 (18)	43 (12)

Abbreviations: APRI, aspartate aminotransferase (AST) to platelet ratio index, calculated as follows: [(AST/upper limit of normal)/platelet count (109 /L)] x 100; ARV, antiretrovirals; HCV, hepatitis C virus; IFNL, Interferon Lambda.

Presented as n(%) or Median (Interquartile Range).

Table 3. Univariate and multivariate analysis (Hazard Ratio, 95% CI) of the association of IFNL genotypes with the development of significant liver fibrosis

	rs12979860 CC	**rs8099917 TT**	**rs8103142 TT**
Univariate	0.98 (0.68, 1.40)	1.41 (0.97, 2.03)	1.03 (0.72, 1.46)
Multivariate	1.37 (0.94, 2.02)	1.79 (1.24, 2.57)	1.34 (0.91, 1.97)
Female	1.38 (0.92, 2.08)	1.44 (0.97, 2.13)	1.42 (0.95, 2.14)
Baseline age, per 10 years	0.70 (0.54, 0.89)	0.71 (0.55, 0.91)	0.69 (0.54, 0.88)
Ethnicity, Aboriginal vs. White	0.90 (0.51, 1.57)	0.89 (0.52, 1.53)	0.90 (0.52, 1.56)
Alcohol use	1.44 (0.98, 2.11)	1.39 (0.94, 2.04)	1.40 (0.97, 2.10)
Baseline APRI	2.92 (1.82, 4.70)	2.90 (1.78, 4.72)	2.81 (1.76, 4.49)
GGT	2.93 (2.37, 3.61)	2.89 (2.35, 3.56)	2.90 (2.36, 3.58)
CD4 count, >=350 vs. <350	0.70 (0.48, 1.02)	0.70 (0.48, 1.02)	0.69 (0.48, 1.00)
HCV genotype 3, vs. non-3	1.44 (0.80, 2.57)	1.42 (0.79, 2.54)	1.48 (0.83, 2.62)

Abbreviations: CI, confidence interval; IFNL, Interferon Lambda; APRI, aspartate aminotransferase (AST) to platelet ratio index, calculated as follows: [(AST/upper limit of normal)/platelet count (109 /L)] x 100; GGT, gamma-glutamyl transferase; HCV, Hepatitis C virus.

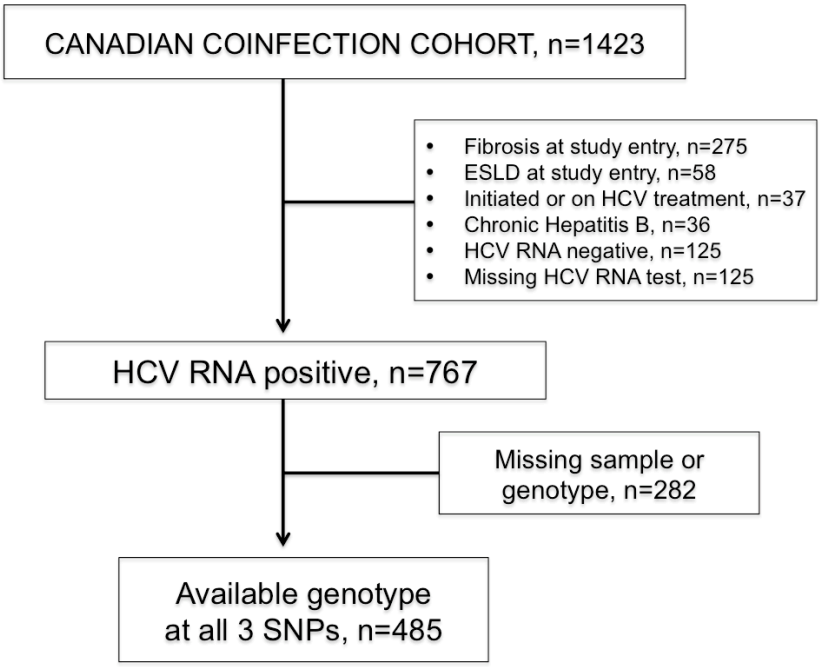
Table 4. Haplotype Analyses with IFNL SNPs and Significant Liver Fibrosis

Mode of inheritance	Hazard Ratio (95% CI)*	p-value	Log-likelihood
Dominant	1.14 (0.73, 1.77)	0.56	-6246.57
Additive	1.05 (0.83, 1.33)	0.66	-6246.65
Recessive	1.04 (0.72, 1.48)	0.85	-6246.73
Co-Dominant	1.15 (0.71, 1.86)	0.57	-6246.57
	0.86 (0.41, 1.79)	0.69	

Abbreviations: IFNL, Interferon Lambda; CI, confidence interval

* haplotype containing the major allele at all 3 SNPs (TCT corresponds to T at rs8103142, C at rs12979860 and T at rs8099917), adjusted for ethnicity.

Figure 1. Source and Study population for examining link between IFNL genotypes and significant liver fibrosis



5A. APPENDIX FOR CHAPTER 5 (Manuscript 2)

5.4 SENSITIVITY ANALYSIS FOR MANUSCRIPT 2

5.4.1 Missing Data

Approximately 66% of the CCC was excluded to derive the final study population for Manuscript 2. The distribution of most of the important variables are similar, except IFNL genotypes (Appendix Table 5.1), which could induce selection bias and is discussed in section "**5.4.2 Selection bias**."

Appendix Table 5.1 Baseline Features of Source (CCC), Study, and Excluded Population, Manuscript 2

Variable	CCC, n=1423	Study population, n=485	Excluded, n=938
Follow-up time, years	3 (1-6)	5 (3-6)	2 (0.3-5)
Age at baseline, years	45 (39-50)	44 (38-49)	45 (40-51)
Male	1,026 (72)	335 (69)	691 (74)
White	1,054 (74)	369 (76)	685 (73)
Born in Canada	1,114 (90)	420 (91)	694 (90)
History of injection drug use	1,143 (81)	403 (83)	740 (79)
Current alcohol drinker	744 (52)	234 (48)	512 (55)
APRI	0.63 (0.39-1.24)	0.52 (0.37- 0.78)	0.75 (0.39-1.77)
Undetectable HIV viral load[a]	876 (63)	280 (59)	596 (65)
On HIV antiretroviral therapy	1,178 (83)	382 (79)	795 (85)
CD4 count, cells/μl	398 (250-575)	381 (260-550)	407 (247-590)
HCV duration, years	18 (10-26)	17 (10-25)	18 (10-26)
HCV genotype 1	852 (74)	313 (77)	539 (72)
Ever treated for HCV	252 (18)	46 (10)	206 (22)
		IFNL genotypes	
rs12979860CC	437 (48)	202 (42)	237 (54)
rs8099917TT	599 (65)	297 (61)	304 (69)
rs8103142TT	451 (50)	214 (44)	239 (56)

Abbreviations: CCC, Canadian Co-infection Cohort; APRI, aspartate aminotransferase (AST) to platelet ratio index, calculated as follows: [(AST/upper limit of normal)/platelet count (109 /L)] x 100; HCV, Hepatitis C virus; IFNL, Interferon Lambda.

a. Limit of detection, <=50 copies/ml

Presented as n(%) or Median (Interquartile Range).

For exclusion due to missing data, especially missing RNA or IFNL genotype, we conducted sensitivity analyses using inverse probability weights (IPW). The weights for censoring were modeled as the probability of having missing RNA or IFNL genotype at visit 1 using a pooled logistic regression model with indicator variables for study visits and including factors most likely to explain why someone was missing sample for IFNL genotype or an RNA test. The following predictors were included: centre ID, currently drinking alcohol or not, log-transformed baseline APRI, age, Aboriginal ancestry, sex, whether missing HCV genotype or not, history of injection drug use, homelessness, monthly income <$1500, whether born in Canada, liver disease status (end-stage liver disease (ESLD) and significant liver fibrosis as measured by APRI ≥ 1.5), and whether dead or alive.

The stabilized weights (mean=1) were derived as detailed in Chapter 4 and included in a pooled logistic analysis for significant liver fibrosis with visits as indicator variables, along with the factors included in the final model of Manuscript 2. Results indicate that estimates are very similar to those reported from the Cox model in Manuscript 2 (Appendix Table 5.2). In addition to accounting for missing data, this pooled logistic regression with IPW also serves as a sensitivity analysis for interval censoring.

Appendix Table 5.2 Sensitivity Analysis with Inverse Probability Weights in Pooled Logistic Models to Account for Missing and Excluded Data in Manuscript 2, Adjusted Effect Estimates (95% CI)*

	Pooled logistic*	Cox (Manuscript 2)*
rs12979860 CC	1.31 (0.92, 1.87)	1.37 (0.94, 2.02)
rs8099917 TT	1.34 (0.92, 1.96)	1.79 (1.24, 2.57)
rs8103142 TT	1.46 (1.02, 2.10)	1.34 (0.91, 1.97)

*Adjusted for all variables in the final model in Manuscript 2. Pooled logistic model also includes visits as indicator variables

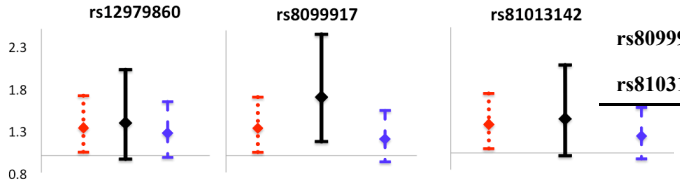
5.4.2 Selection bias

Of the 938 excluded individuals, 333 of them were prevalent liver disease cases (significant liver fibrosis or ESLD at visit 1). If these cases were also more likely to have the inflammatory *IFNL3* genotypes (i.e. exposed cases), excluding them could induce selection bias. Based on Appendix Table 5.1, the at-risk genotypes are more frequent among the excluded group.

To account for potential selection bias, the following sensitivity analyses was performed with age as the time axis: the final Cox model was run on a dataset that included all the prevalent cases with and without adjustment for baseline liver diseases. This sensitivity analyses was run in a dataset with no excluded individuals (CCC, n=1423) and one where we excluded only those who initiated or were on treatment (Treatment-naïve, n=1374). The effect estimates (Appendix Figure 5.1) were very similar to those reported in Manuscript 2, regardless of adjustment for prevalent liver diseases. This indicates that any selection bias induced by removing cases of existing liver disease from the analysis was likely negligible. Using age as the time axis also serves as sensitivity analysis to address the uncertainty in estimates of HCV duration, which is used as the time axis in Manuscript 2.

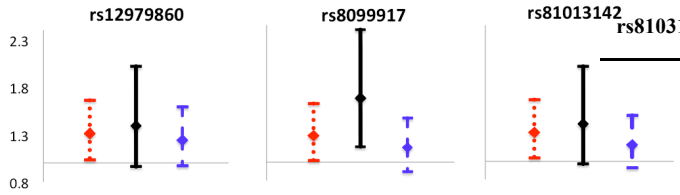
Appendix Figure 5.1 Sensitivity Analyses using Multivariate Cox Models* to Address Selection Bias in Manuscript 2 (IFNL genotypes and Significant Liver Fibrosis), aHR (95% CI)

a) No adjustment for prevalent liver diseases



	CCC, n=1423	Study, n=485	Treatment-naïve, n=1374
rs12979860CC	1.33 (1.04, 1.71)	1.39 (0.96, 2.02)	1.27 (0.98, 1.64)
rs8099917TT	1.33 (1.04, 1.70)	1.70 (1.17, 2.45)	1.20 (0.93, 1.54)
rs8103142TT	1.34 (1.05, 1.71)	1.41 (0.97, 2.05)	1.20 (0.93, 1.54)

b) After adjustment for prevalent liver diseases



	CCC, n=1423	Study, n=485	Treatment-naïve, n=1374
rs12979860CC	1.32 (1.03, 1.68)	1.39 (0.96, 2.02)	1.26 (0.98, 1.62)
rs8099917TT	1.30 (1.02, 1.66)	1.70 (1.17, 2.45)	1.16 (0.90, 1.49)
rs8103142TT	1.34 (1.05, 1.70)	1.41 (0.97, 2.05)	1.20 (0.93, 1.53)

Abbreviations: IFNL, Interferon Lambda; CCC, Canadian Co-infection Cohort

*Age as time axis; adjusted for all variables in the model as in Manuscript 2

Red dotted lines: Full Canadian Co-infection Cohort (CCC), n=1423

Black solid lines: Study population as in Manuscript 2, n=485

Blue dashed lines: Treatment-naïve, n=1374

5.4.3 Interval Censoring

To account for the fact that exact event time was unknown as data was captured at semi-annual visits, indicator variables for visits was used in the analysis in Appendix Table 5.2 with IPWs, which also accounted for those excluded due to missing sample. Furthermore, to address interval censoring specifically in the study population from Manuscript 2, a polynomial transformation of visit numbers was included in a pooled logistic regression model along with all the variables in the final model. The effect estimates of the IFNL SNPs with significant liver fibrosis were

similar to those reported in Manuscript 2 and derived from Cox. This is not surprising as the methods are approximately equivalent when the time between the intervals is short and the probability of the event is low within the interval.[269]

Appendix Table 5.3 Sensitivity Analyses using Pooled Logistic Regression with Polynomial Transformation of Visits to Account for Interval Censoring in Manuscript 2 (IFNL and Significant Liver Fibrosis), Adjusted Effect Estimates (95% CI)*

	Pooled logistic*	Cox (Manuscript 2)*
rs12979860CC	1.46 (0.95, 2.24)	1.37 (0.94, 2.02)
rs8099917TT	1.94 (1.30, 2.89)	1.79 (1.24, 2.57)
rs8103142TT	1.48 (1.00, 2.18)	1.34 (0.91, 1.97)

*Adjusted for all variables in the final model in Manuscript 2.

5.5 STATISTICAL SIGNIFICANCE OF THE EFFECTS OF THE SNPS

The effects of SNPs rs12979860 and rs8103142 were weaker than rs8099917 and lacked statistical significance. Some possible biological reasons for this are outlined in the Discussion section of Manuscript 2. Statistically, this could also be due to lack of power (i.e. insufficient number of events for the lower effect size at these SNPs) or due to the lower accuracy of APRI to detect fibrosis. In the latter scenario, nondifferential misclassification by APRI could lead to a bias towards the null.

CHAPTER 6. Developing a Prognostic Model for Significant Liver Fibrosis using Immune and Genetic Markers (Manuscript 3)

6.1 PREFACE TO MANUSCRIPT 3

As described in 5, genetic markers can help identify those who are at higher risk for significant liver fibrosis by capturing the heightened inflammation that drives faster liver disease progression in co-infected persons. Direct immune markers of the fibrogenesis process could also help identify higher-risk individuals. Using markers can thus help target expensive HCV treatments to prevent progression to advanced liver disease, which may be irreversible and reduce treatment efficacy. However, immune markers are relatively expensive to measure and not routinely assessed in clinic or study visits. Furthermore, not much is known about their ability to improve prognosis, defined as estimating the likelihood of developing the outcome in the future, which differs from diagnosis or the detection of existing outcome or disease stage (diagnostic factors are assessed at the time of outcome). Risk assessment for prognosis thus differs from diagnosis in its stochastic nature, because data collection for prognosis precedes occurrence of the outcome, which is often unknown at the time the prognostic factors are assessed.[270]

We used a case cohort design to address some of these problems. Measuring the markers only in cases and a randomly selected subcohort that is representative of the full source population provides a practical, economical way to collect data that was not gathered during regular follow-up. Additionally, because the Barlow weights mimic the proportions observed in the full cohort sample, they can be used for estimation of absolute risk without any further rescaling of the

cumulative baseline hazard and survival functions.[271] While the case cohort design has been gaining popularity with the rise of new metabolic and genomic markers, ours is one of the first applications for prognosis of liver fibrosis in a co-infected population.

We also assessed a variety of markers from various stages of fibrogenesis, allowing us to potentially capture the outcome at different stages of development. Another strength of our study was a large source population that was representative of Canadian co-infected patients, thus making our results directly generalizable and clinically relevant. This is important as features such as marker and disease prevalence or marker correlations with other known risk factors can all affect the estimates of the discriminatory capacity of a marker.[13] Our study also simulated the clinical setting in that we included marker measures from only one time point (study entry) decades after HCV infection. In real life, symptoms of liver disease are largely silent, so many people seek care many years after acquiring HCV, something which our study accurately reflects.

To compare whether incorporating measures of markers improved ability to predict 3-year risk of significant liver fibrosis, we compared two models—Model 1 included clinical predictors only, while Model 2 included Model 1 plus selected markers. The measures of predictive accuracy we compared between the two models included discrimination, calibration and net reclassification improvement index (NRI).

6.2 STUDY POPULATION IN MANUSCRIPT 3

Because we wanted to assess immune and genetic markers to predict 3-year risk of liver fibrosis, we excluded participants who were no longer at risk (those who already had the outcome or had cleared HCV) or who were on HCV treatment, because treatment affects AST and platelet count, thus affecting the accuracy of APRI, the outcome measure. We also excluded those who were infected with Hepatitis B, as triply infected individuals have a different risk profile. Using a sampling fraction of 0.45 in the eligible subpopulation (n=679), we generated a random subcohort (n=236) and included all cases outside the subcohort (n=59) in the analysis. Based on sample availability, we had immune marker information on 171 individuals from the subcohort and 46 nonsubcohort cases. We used MICE to impute the missing data, as detailed in the "Methods" section of Manuscript 3.

6.3 CASE COHORT DESIGN

This study design was originally popularized by Ross Prentice[272] as a way to reduce data collection when sampling the full study population was not feasible or economical. It became most applicable to studies where biobanks of specimens were available, such as occupational exposure studies. Conceptually, a random subsample or "subcohort" was selected from the original sample at entry, which served as the source of comparison observations for each observed event. As the subcohort is selected randomly without regard to case status, it can be considered representative of the full eligible cohort and thus can be used in other studies with

different endpoints of interest, unlike a nested case-control study. The subcohort will also contain a fraction of the cases, while the cases outside the subcohort enter via delayed entry but otherwise do not contribute any reference person-time for analysis, i.e. cases outside the subcohort are not included in earlier risk sets.

The entry of nonsubcohort cases into the subcohort necessitates the use of a robust jackknife estimator for the variance as well as different weighting schemes to account for the sampling.[273] Barlow weights, which are used in Manuscript 3, assign weight of 1 to all cases at event time while all subcohort members are weighted by the inverse of the sampling fraction.[273] Therefore, as the original sampling fraction is 0.45 in Manuscript 3, each subcohort member has weight 2.2. Cases in the subcohort are also weighted by the inverse of the sampling fraction until failure time, at which point, they are assigned a weight of 1. This aims to mimic the proportions that would have been present in the full cohort analysis. The appropriate weights are also added to Harrell's C-index, our metric of discrimination, to account for the case cohort study design.[271]

6.4 DETAILS ON IMMUNE AND GENETIC MARKER ASSAYS

6.4.1 Immune marker assays

Chapter 2 provides an extensive background and summary on all the immune markers we considered for Manuscript 2. The method used to measure the markers in serum or plasma samples is summarized here.

Samples were first virally inactivated with 5% Triton in a 10:1 ratio for a final Triton concentration of 0.5%. The samples were then incubated at room temperature for 1 hour, then spun at maximum speed for 15 minutes to remove any debris. Each sample was aliquoted and re-frozen at -80°C. A subset of duplicates (n=30) representing individuals from the subcohort and nonsubcohort cases was run to ensure that the deactivation-freeze-thaw cycles did not alter results. Samples were run in 96-well plates with standards, controls, and reference samples from HIV-monoinfected individuals as well as healthy individuals for comparison. This was done to generate correct standard curves and provide realistic ranges. The plates were also arranged with a mix of samples from subcohort and cases to account for batch effects.

For each assay, one aliquot of the sample was used, melted at 4°C, spun at maximum speed for 10 minutes and then used as duplicates. Samples were diluted based on protocol recommendations for the kits: 3-plex (RANTES, sICAM1, sVCAM1; TGF-β1, 2, 3), 6-plex (IL6, IL8, MCP1, MIP1α, MIP1β, and TNFα) and 2-plex (CXCL9 and CXCL11). Standards were prepared in the same background as samples and run on a commercial assay from Millipore on a MAGPIX instrument (Millipore Corporation, Billerica MA) using a Luminex platform, according to the manufacturer's instructions. There was no difference in readings based on sample type (plasma or serum).

Commercially available ELISA kits were used to measure plasma and serum levels of soluble CD14 (sCD14, R&D Systems, Minneapolis, MN, USA) with dilutions of 1:300. High-sensitivity C-reactive protein (hsCRP) was tested based on manufacturer's instructions, using immunoassay kits from Synchron LX 20 PRO (Beckman Coulter, Ontario, Canada).

6.4.2 Genetic marker IFNL rs8099917

Details on the genotyping process are provided in the Methods section of Manuscript 1 in Chapter 4. Reasons for including this genotype as a marker for liver fibrosis are our findings from Manuscript 2 as well as results from other studies. Many studies have linked the homozygous genotype rs8099917 TT with a pro-inflammatory response that leads to higher HCV clearance.[125-127,263] However, if the virus persists, this same genotype may cause higher necroinflammation and fibrosis progression.[1,123,147] Several studies have shown that IFNL polymorphisms are linked with elevated histological inflammatory activity.[1,12,147,149,266,274] In addition, other markers in this region turn on genes that are involved in Natural Killer (NK) cell activation, resulting in cell death of infected liver cells and a pro-inflammatory environment.[255,267] Because this marker is not routinely collected as part of clinical care, we evaluated its predictive ability in Model 2.

6.5 DATA ANALYSIS

6.5.1 Correlation of Predictors in Subcohort

We examined the correlation of the clinical predictors and prospective log-transformed immune markers and IFNL genetic marker at baseline in the subcohort. We used the Spearman's rank correlation coefficient, which does not assume normality and is less sensitive to extreme values compared to the Pearson correlation coefficient. Among the clinical predictors we included in both models, none of them were strongly correlated with each other or with the prospective immune markers. This has some implications for the Harrell's C-index (see section **"6.5.3 Discrimination Measures: Harrell's C-index"**).

The final immune markers we selected for Model 2 (IL-8, sICAM-1, RANTES, hsCRP, and sCD14) are strongly correlated with the other markers we did not select. MIP1α, which we did not include, has weaker correlations than the other markers and also appears to differ between the subcohort and cases (Fig 2 in Manuscript 3). However, addition of this marker did not improve discrimination, calibration or fit.

6.5.2 Multiple Imputation

Variables used include all the predictors in the final models, all the immune markers, as well as the following: total bilirubin, gamma-glutamyltransferase (GGT), CD4 count, diabetes, BMI, ethnicity, centre ID, smoking status, homelessness, presence of an AIDS defining illness or not, injection drug user status, monthly income, end-stage liver disease status, HIV treatment interruptions or not, spontaneous clearer status, outcome variable and the Nelson-Aalen estimate of the cumulative hazard function.[250]

Misspecified multiple imputation models can lead to bias in estimating predictive ability, but this problem is especially acute if interaction terms are present.[275] In our study, because we did not have interaction terms and because the C-index estimates are so similar across the analyses, overt bias is likely not a problem. We cannot rule out all misspecification, however, as the imputed estimates are closer to the null than the non-imputed estimates.[250]

6.5.3 Measuring Predictive Accuracy

The stochastic nature of prognosis necessitates the use of specific measures for evaluating improvements in risk prediction. Strong associations (e.g. high odds ratios) in etiologic models

do not automatically mean that a marker improves ability to classify or predict risk for individuals, which is most relevant for guiding both clinical and policy decisions.[192,193] While risk predictors are almost always risk factors, the converse is not necessarily true,[276] because not only are *extremely* strong associations needed to improve utility in predicting future outcomes (magnitudes rare in epidemiologic studies), but also, the characteristics of the population or assay technique may themselves alter marker performance.[192,277] For example, higher breast density makes mammographic readings less accurate.[278]

Furthermore, other factors, such as validity of the multivariate measurement and analysis, the prevalence of the marker in the population, and its correlation with other known risk factors can all affect the discriminatory capacity of a marker.[13] For example, a factor might have a very strong association with the outcome, but if its prevalence is low in the population, it cannot substantially influence risk prediction.[279] Finally, measures such as the AUROC and risk stratification tables[280] offer a clearer picture of a marker's discriminatory ability (e.g. true positive fraction or sensitivity and false positive fraction or 1-specificity) than a specific odds ratio, which is a scalar association that can be of the same value from different pairs of sensitivities and specificities.[192]

Discrimination, which involves separating those with and without the disease, is an overlapping goal of prognostic and diagnostic models. Measures of interest in prognostic models only include calibration, which measures the agreement between observed and predicted risks in homogenous groups, and risk stratification, which involves assigning appropriate risk levels that can directly

impact clinical decisions. These metrics are discussed in more detail in the following sections, along with bootstrapping, which was used for internal validation.

6.5.4 Discrimination Measures: Harrell's C-index

As a rank parameter, Harrell's C-index only uses the order of the risk estimates, not the magnitude, to separate the events from the nonevents. As a generalization of the AUROC for censored data, the C-index can be interpreted as the proportion of all usable pairs where concordance is achieved, i.e. the higher risk estimate is assigned to the individual in the pair who had the event earlier.[281] A C-index value of 1 indicates perfect discrimination, while 0.5 means no better than random guessing. To put this in context of liver disease or HIV, a prognostic score for prediction of ESLD in HIV-HCV co-infected individuals on cART had a Concordance statistic (C-statistic) of 0.73, which can be considered good and of some clinical utility.[282] Other prognostic indexes which sought to predict risk of mortality in cirrhotics (Child-Pugh score, MELD score) or in HIV-infected individuals on cART (VACS Index)[283] had C-statistics around 0.80 or higher, which could be considered very good or excellent.[284,285] In other fields such as cardiovascular disease, one of the best known prognostic indices, the Framingham Risk Score, provided good discrimination in most cohorts with C-statistics ranging from 0.75-0.80.[286,287]

However, while the AUROC and C-index can be useful in measuring discrimination, it has some limitations. It can be insensitive to improvements in discrimination if the existing model is already strong or if the new predictor is correlated to the variables already in the model.[270] As such, values of C-index should be assessed in context with other measures of predictive accuracy and clinical impact, especially NRI.

6.5.5 Bootstrapping and Internal Validation

The models were internally validated using bootstrapping. In each imputed set, the C- index was estimated and bootstrapped 500 times. The estimates and standard errors were then combined using Rubin's rules.[250]

6.5.6 Calibration Measures

Calibration measures the extent to which predicted estimates of risk agree with observed probability of event.

6.5.6.1 Hosmer-Lemeshow

A commonly used measure is the Hosmer and Lemeshow test statistic, which is approximately distributed as a chi-square with k-2 degrees of freedom, where k=number of groups that predicted probabilities are divided into after ranking. A statistically significant test statistic indicates poor calibration, i.e. the model is a poor fit for the data, meaning there is a statistically significant difference for at least one group in the predicted number of outcomes compared to the observed number of outcomes. The Hosmer and Lemeshow test is best suited for logistic models and is sensitive to the number of groups as well as the cut points used for those groupings.

6.5.6.2 GB Test

The GB test is more appropriate for survival models as it is based on martingale residuals, and takes censoring and event times into account, unlike the Hosmer and Lemeshow test, which only calculates proportions at the end of follow-up. The GB test ranks and divides groups based on

the estimated risk score and compares the expected and observed events by summing the martingale residuals within each group and testing the null hypothesis that the sum is equal to zero in each group.[288] P-values from score and likelihood tests, as well as graphs of observed and predicted events in each risk subgroup in Arjas-like plots can be used in the GB test. P-values <0.05 or deviations between observed and expected number of events can reveal miscalibration. While the Arjas-like plots from the non-imputed data seem to indicate some deviation (Appendix Figure 6.1a), this is likely due to missing information, as the plots from imputed data display better agreement between observed and predicted number of events. (Appendix Figure 6.1b).

6.5.6.3 Calibration Plots

The Stata command –stcoxgrp- was used to visually assess calibration, by plotting the predicted survival functions with observed (Kaplan–Meier) curves in subgroups of risk.[289] As locations and number of cutpoints can affect estimates, calibration was assessed in two different scenarios (Supplementary Figure 3 in Manuscript 3). In one, subgroups included equal-sized quintiles, while in another 3 unequal groups cut at the 25th and the 75th percentiles of the failure times was used. For statistical reasons, Royston and Altman recommend 5 or fewer groups as large number of survival curves may be unstable with poor discrimination between neighboring groups.[290] They also suggest using groups of unequal sizes, because that enables identification of patients with more extreme prognoses and groups together patients with largely similar prognoses-individuals with the highest 25% predicted risk are considered separately from the middle 50% with intermediate risk, for example.[290] The command –stcoxgrp- derives population-averaged survival curves in each risk subgroup and compares them with Kaplan-Meier estimates with 95% CIs of observed risk in those risk subsets. That is, in each risk subgroup, individual survival

functions at the observed event or censoring times are averaged to obtain the population-averaged survival curves predicted by the Cox model. The command provides graphical assessment but not any statistical tests or results.

To minimize loss of information or power due to missing data, calibration plots were examined in multiply imputed data. All continuous predictors were centered at mean values derived from their distribution in the subcohort so that the baseline survivor function (which refers to survival when all predictors in the Cox model are set to zero) was meaningful. Based on the results (Supplementary Figure 3 in Manuscript 3), there was no large deviation between observed and predicted survival probabilities. Observed survival (i.e. likelihood of remaining fibrosis-free) in the worst prognosis group tended to be lower than predicted estimates, that is, predicted estimates in the worst outcome group underestimated observed risk, though the 95% CI of the Kaplan-Meier estimates overlap with the predicted values.

6.5.7 Net Reclassification Improvement (NRI)

This metric measures improvement in risk stratification, defined as the assigning of risk levels that can determine treatment decisions.[270] This index can be defined as the sum of improvements in risk classification in events and non-events. It is measured separately in those with the outcome and those without the outcome-- the sum of differences in proportions of individuals moving up minus the proportion moving down for those with the outcome, and the proportion of individuals moving down minus the proportion moving up for those without the outcome.[291]

$\mathbf{NRI_{events}}$= (# events moving to higher risk category/ # events) - (# events moving to lower risk category/ # events)

$\mathbf{NRI_{nonevents}}$= (# nonevents moving to lower risk category/ # nonevents) - (# nonevents moving to higher risk category/ # nonevents)

$\mathbf{NRI_{overall}}$= $\mathbf{NRI_{events}}$ + $\mathbf{NRI_{nonevents}}$

Category-based and the continuous NRIs were both calculated in Manuscript 3, as current guidelines and literature lack clinically relevant and established cutoffs.

For the category-based NRI, we used 3 clinically relevant risk categories: low risk, <=10%; medium risk, >10-25%; and high risk, >25%. The categories were determined based on estimates of mortality from liver disease in those with chronic HCV infection from published reports[292] as well as opinions of knowledgeable hepatologists and clinicians. For the continuous NRI, no categories were needed and any upward or downward movement in risk was considered, regardless of magnitude.

MANUSCRIPT 3: DEVELOPING A PROGNOSTIC MODEL FOR SIGNIFICANT LIVER FIBROSIS USING IMMUNE AND GENETIC MARKERS

A Prognostic Model for Significant Liver Fibrosis in HIV-Hepatitis C (HCV) Co-Infection

Nasheed Moqueet,[1] Cynthia Kanagaratham,[2] Danuta Radzioch,[2] Sahar Saeed,[1] Robert W. Platt,[1] Marina B. Klein[1,3]

Affiliations: [1] Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, QC Canada; [2]Department of Medicine and Department of Human Genetics, McGill University, Montreal, QC; [3]Division of Infectious Diseases and Chronic Viral Illness Service, McGill University Health Centre, Montreal, QC Canada.

Running title (not to exceed 45 characters and spaces): Prognostic Model for Fibrosis in HIV-HCV Co-infection

Corresponding Author:

Marina B. Klein
Division of Infectious Diseases and Chronic Viral Illness Service
McGill University Health Centre – Glen Site
1001 Decarie Boulevard D02.4110
Montreal, QC, Canada H4A 3J1
email: marina.klein@mcgill.ca
fax: 1-514- 843-2092
phone: 1-514- 843-2090

ABSTRACT

Background: Liver fibrosis advances faster in HIV-HCV co-infected individuals due to heightened inflammation. Immune markers targeting various stages of fibrogenesis could aid in prognosis of fibrosis.

Methods: A case-cohort study was nested in the prospective multicentre Canadian Co-infection Cohort (n=1119). Individuals who had not cleared HCV and were free of HCV treatment, fibrosis, end-stage liver disease and chronic Hepatitis B at baseline (n=679) were eligible. A random subcohort (n=236) was selected from the population at entry and included a few cases [APRI≥1.5 (significant fibrosis) over study follow-up]. Other cases entered via delayed entry.

Pro-fibrogenic markers (IL-8, MIP-1α, MIP-1β, MCP-1,TNF-α, RANTES, sICAM-1, sVCAM-1, CXCL9, CXCL11, TGF-β1, hsCRP, sCD14) and Interferon Lambda (IFNL) rs8099917 genotype were measured from first available sample in the cases and subcohort. Cox proportional hazards with Barlow weights was used. Discrimination, calibration and risk classification were compared between Model 1 (selected clinical predictors only) and Model 2 (Model 1 plus selected markers) for predicting 3-year risk of liver fibrosis. Discrimination was estimated using weighted Harrell's C index; calibration was assessed with the Hosmer-Lemeshow statistic and the Gronnesby and Borgan (GB) test. Continuous and category-based Net Reclassification Improvement (NRI) indices were used to compare risk classification. Bootstrapping was used for internal validation.

Results: 113 individuals developed significant liver fibrosis over 1300 years of risk [event rate (95% CI) 8.63 per 100 person-years (7.08, 10.60)]. Model 1 (sex, current alcohol use, HIV viral load, baseline APRI, HCV genotype, and age) was nested in Model 2, which also included IFNL rs8099917 genotype and immune markers IL-8, sICAM-1, RANTES, hsCRP, and sCD14. The C indexes (95% CI) for model 1 vs. model 2 were 0.720 (0.649, 0.791) and 0.756 (0.688, 0.825), respectively. Both models were well-calibrated. Model 2 classified risk more appropriately (overall NRI, $p<0.05$).

Conclusions: Including markers at IFNL rs8099917, IL-8, sICAM-1, RANTES, hs-CRP, and sCD14 enabled better prediction of the 3-year risk of significant liver fibrosis over clinical predictors alone. Whether this improvement justifies the additional cost of measuring these markers in the face of highly expensive HCV treatment requires further cost-benefit analyses.

Key words: case cohort; HIV-Hepatitis C (HCV) co-infection; liver fibrosis; prognosis.

Liver disease has become one of the leading non-AIDS causes of death among HIV-infected individuals in the developed world, mainly due to co-infection with hepatitis C (HCV). HCV infection is not cleared by HIV treatment, making HCV treatment a priority for this population.[293,294] Compared to HCV mono-infected individuals, liver fibrosis progression is accelerated in co-infected individuals, leading to cirrhosis, hepatocellular carcinoma or end-stage liver disease (ESLD).[3,68] Reasons for this acceleration include biological factors and possibly HIV therapy-related toxicity. HIV itself suppresses the immune response to HCV,[3,74] triggering a vicious cycle where inflammatory and fibrogenic cells continually stimulate each other, distorting the hepatic architecture, eventually leading to fibrosis. Liver fibrosis progression is thus caused by heightened inflammation rather than direct HCV replication.

In Canada, the majority of the HIV-HCV co-infected population is made up of current or former injection drug users for whom treatment access and adherence may be challenging. While international clinical guidelines recognize that co-infected individuals should be prioritized for HCV treatment,[295,296] the high cost of treatment (between $50,000 and $120,000 for a course of the new DAAs) in Canada[119,120] has meant that reimbursement by public and private payers has been restricted to those with advanced fibrosis (METAVIR stage F2 and higher). While biopsy is considered the gold standard to stage liver fibrosis, noninvasive indexes such as aspartate aminotransferase (AST) to platelet ratio index (APRI) and FIB-4 can also be used to determine treatment eligibility in some provinces, such as British Columbia.[297] Therefore, accurate markers of fibrosis could be essential for identifying persons at higher risk for liver disease progression in order to target effective intervention and treatment strategies in a cost-effective manner.

Some risk factors that have been associated with fibrosis progression include: alcohol intake (>50 g/day), infection with HCV genotype 3, male sex, excess weight, liver steatosis, presence of metabolic syndrome and/or type II diabetes, host genetic factors such as single nucleotide polymorphisms (SNPs) in the Interferon Lambda (IFNL) gene, hepatitis B co-infection, immunodeficiency related to HIV or the use of immunosuppressant drugs such as those used after liver transplantation, and HIV therapy interruption.[1,54,55,264]

We hypothesize that the addition of genetic markers of fibrosis progression and a panel of immune markers representative of the underlying inflammatory mechanisms would improve prediction of fibrosis risk beyond traditional clinical risk factors alone and would thus be of value in optimal use of new and expensive HCV therapies.

METHODS

Source population: The Canadian Co-infection Cohort (CCC) Study is an open prospective cohort of HIV-HCV co-infected individuals recruiting from 18 centres across Canada, representing approximately 23% of the co-infected population under care. At visits every six months, socio-demographic, medical and behavioural information are collected using pre-validated questionnaires, along with plasma, serum and peripheral blood mononuclear cells (PBMC). For our study, we included data from 1119 patients collected up until July 2012.

To be included in the CCC, patients must be over 16 years or older, give informed consent, be HIV infected (confirmed via ELISA with Western blot), and have HCV infection or evidence of

HCV exposure--HCV-antibody positive by ELISA with recombinant immunoblot assay II (RIBA II) or enzyme immunoassay (EIA) or if serologically false negative, HCV–RNA-positive. The study has been approved by research ethics boards at each of the participating institutions.[43]

Study population: For our study, HCV RNA-negative participants or those who had significant fibrosis (APRI≥1.5), end-stage liver disease or chronic Hepatitis B at study entry were excluded, as were individuals on HCV treatment. HCV RNA levels were tested using qualitative tests (COBAS AMPLICOR HCV Test, version 2.0, Roche Diagnostics, Hoffmann-La Roche Ltd, Laval, Canada, lower limit of detection <50 IU ml-^{1}) and were available at most visits. Presence of Hepatitis B surface antigen was used to determined Hepatitis B chronicity.

From an eligible study sample of n=679 (Fig 1), a random subsample or "subcohort" was selected from the population at entry, which served as the source of comparison observations for each observed event of significant liver fibrosis, defined as APRI≥1.5 over study follow-up. Because the subcohort was a representation of the full cohort, it also contained a few cases. Cases outside the subcohort were also included in the analysis.

Outcome measure: Progression to significant liver fibrosis (METAVIR stage F2-F4) was defined by an APRI≥1.5. The aspartate aminotransferase (AST) to platelet ratio index (APRI) is calculated as follows: [(AST/upper limit of normal)/platelet count(10^{9} /L)] x 100. It consists of routinely available and non-invasive measures that are available for almost every visit.

An APRI cutoff of 1.5 or higher has been validated against liver biopsies in our study population for detection of significant liver fibrosis (METAVIR stages F2-F4) with a sensitivity of 52%, specificity and positive predictive value (PPV) of over 99% and an AUC of 0.85 ± 0.06.[93] In addition, APRI cutoffs of 1.5 and 2 have also been shown to be associated with cirrhosis, other adverse liver and clinical outcomes, and death in our study cohort,[94] as well as in others.[95,96]

Clinical predictors: All values were time-fixed at first available visit and distribution was evaluated in the subcohort and compared to the full eligible study sample to mimic how a predictive score would be applied in practice. Predictors considered included sex; ethnicity; body mass index (BMI); age; HIV viral load; CD4 count; and baseline APRI. Variables included in fibrotic staging indexes from other studies such as total bilirubin[298,299] or gamma-glutamyltransferase (GGT)[298-300] were also considered.

Markers of interest: Of interest to us was the genetic marker at IFNL SNP rs8099917, which has been linked with elevated histological inflammatory activity[1,12,147,149,266,274] as well as Natural Killer (NK) cell activation, resulting in cell death of infected liver cells and a pro-inflammatory environment.[255,267] For immune markers, we chose the following based on their specific roles in liver fibrosis development (Supplementary Table 1): the cytokines transforming growth factor beta 1 **(TGF-β1)** and tumor necrosis factor alpha **(TNF-α);** the chemokines interleukin-8 **(IL-8);** monocyte chemotactic protein-1 (**MCP-1** or CCL2); macrophage inflammatory protein 1 (**MIP1α** or CCL3; **MIP1β** or CCL4); Regulated upon Activation, Normal T cell Expressed and Secreted protein (**RANTES** or CCL5); **CXCL9**; and **CXCL11**; endothelial activation markers soluble Intercellular Adhesion Molecule 1 (**sICAM-1)** and soluble Vascular Cell Adhesion

Molecule 1 (**sVCAM-1);** high-sensitivity C-reactive protein **(hsCRP);** and soluble CD14 **(sCD14),** a marker of microbial translocation.[184,186]

Immune markers were measured for patients with available samples from visit 1 or 2 from all the cases and the subcohort (n=171 from subcohort and 46 nonsubcohort cases, Fig 1). Frozen, never thawed plasma or serum samples were used. After thawing on ice, viral activity was inactivated, and then samples were aliquoted, refrozen and thawed before running with a commercial assay from Millipore on a MAGPIX instrument using a Luminex platform (Millipore Corporation, Billerica MA) according to the manufacturer's instructions. Samples were diluted according to kit recommendations: 3-plex (RANTES, sICAM1, sVCAM1; TGF- β1, 2, 3), 6-plex (IL6, IL8, MCP1, MIP1α, MIP1β, and TNFα) and 2-plex (CXCL9 and CXCL11). Standards were prepared in the same background as samples. More details are outlined in "Chapter 6. Developing a Prognostic Model for Significant Liver Fibrosis using Immune and Genetic Markers (Manuscript 3)."

Commercially available ELISA kits were used to measure plasma and serum levels of soluble CD14 (sCD14, R&D Systems, Minneapolis, MN, USA) with dilutions of 1:300. High-sensitivity C-reactive protein (hsCRP) was tested based on manufacturer's instructions, using immunoassay kits from Synchron LX 20 PRO (Beckman Coulter, Ontario, Canada). Hyaluronic acid, another direct marker of fibrogenesis[100] that has been included in other fibrotic indexes,[91,299] was also measured with a 1:30 dilution using the Hyaluronan Quantikine ELISA kit (R&D Systems, Minneapolis, MN, USA).

To measure IFNL genotypes, never thawed plasma and serum samples were processed using a real-time PCR assay developed by the Bay Area Genetic Lab (BAGL, Ontario, Canada), as previously described.[263]

Statistical analysis

Survival analysis: The subcohort (n=236) was generated with a random sampling fraction of 0.45 and included 54 cases. Cases that were not in the subcohort (n=59) entered via delayed entry. The time axis was follow-up time in study. Cox proportional hazards was used for analysis, with robust variance and Barlow weights to account for the case cohort design.[273]

Descriptive analysis was conducted in the subcohort using box plots, histograms, correlation matrices with Spearman's correlation coefficient, scatter plots and Q-Q plots. Values of the markers that were near the limits of detection were assigned the lowest detectable value.

Predictor selection and functional form: Predictors were considered based on their practical availability to physicians, strength of correlation with each other,[301] magnitude of associations in univariate analyses, ability to improve model fit as indicated by the Akaike Information Criterion (AIC) and ability to maximize discrimination as measured by Harrell's C.[302] For immune markers, variables were also included in models if they captured a different stage of the underlying etiology of fibrosis development or were linked to fibrogenesis in the literature. Statistical significance with p-values in univariate analyses was not a deciding factor.[301,303]

Univariate and multivariate analyses were conducted on untransformed variables, as well as after log-transformation or using median or quartile distributions of the immune markers. Functional forms for included continuous variables were determined using residual plots (deviance, Martingale), degree of freedom considerations, and clinically significant values. Continuous variables were centered at their mean values. Log transformation was used for all included continuous variables such as baseline APRI and immune markers IL-8, sICAM-1, RANTES, hsCRP, and sCD14 but not for age, which was modeled as a restricted cubic spline with 3 knots at the 10th, 50th and 90th percentiles corresponding to ages 33, 44, and 54. HIV viral load was dichotomized (undetectable or not at ≤ 50 copies/ml), as were alcohol use (currently drinking or not); HCV genotype (3 vs. non-3, i.e. types 1, 2, and 4); and host IFNL genotype rs8099917 (TT vs. non-TT).

Proportional hazards were assessed using the using Stata command –stphtest, detail- which uses scaled Schoenfeld residuals to check if proportional hazards holds globally and for included predictors.

Predictive accuracy: Discrimination, calibration and changes in reclassification were compared between Model 1 (selected clinical predictors only) and Model 2 (clinical predictors from Model 1 plus selected genetic and immune markers) for predicting 3-year risk of significant liver fibrosis.

Discrimination was measured with a weighted[271] Harrell's C or concordance index using Stata command –somersd- with robust jackknife estimator for standard errors.[302] A C-index value of 1

indicates perfect discrimination, while 0.5 means no better than random guessing. Calibration was assessed statistically (Hosmer-Lemeshow statistic and the Gronnesby and Borgan (GB) test),[288] as well as graphically with the Stata command –stcoxgrp using imputed data.[289] The statistical tests used 5 equal-sized subgroups (i.e. quintiles) as well as tertiles of predicted 3-year risk, as they can be affected by the number and threshold of the cutpoints. For the graphical checks, we plotted the Kaplan-Meier estimates against predicted risks derived in two different ways: from equal quintiles of the 3-year risk as well as 3 unequal groups cut at the 25th and the 75th percentiles of the failure times. The latter method allows identification of individuals with the most extreme prognosis.[289]

Change in reclassification was measured by the net reclassification improvement (NRI) summary index.[291] We calculated both the category-based and the continuous NRIs. For the category-based NRI, we used 3 clinically relevant risk categories: low risk, <=10%; medium risk, >10-25%; and high risk, >25%. The categories were determined based on estimates of mortality from liver disease in those with chronic HCV infection from published reports[292] as well as opinions of knowledgeable hepatologists and clinicians. For the continuous NRI, no categories were needed and any upward or downward movement in risk was considered, regardless of magnitude.

The models were internally validated using bootstrapping. All analyses were conducted using Stata 12.

Multiple Imputation: The case cohort study can be considered data missing by design, because immune marker measures are missing for the full cohort and available only for the subcohort and cases. As such, this missingness is missing at random, where the chances of being measured depend on the case status.[275] Missingness for plasma samples and other variables was also assumed to be at random. Multiple Imputation by Chained Equations (MICE) was used on the full cohort to account for all missing data, using all the predictors in the final models, all the immune markers, as well as variables that were possibly related to the reasons for missingness.

Two types of analyses were conducted following MICE: One, using the imputed full cohort data to measure predictive accuracy after unweighted Cox proportional hazards[275] (that is, immune markers were imputed for the full cohort and the imputed data was analyzed as a cohort design, using suitable estimators. This type of analysis can give unbiased estimates, provided the MI model is specified correctly).[304] This is referred to as Analysis 1 and results are included in Table 3. Analysis 2 involved generating a case cohort in each imputed set, followed by weighted predictive accuracy measures (that is, in each imputed set, a random subcohort was generated and then analyzed accordingly with Cox proportional hazards regression with Barlow weights, followed by weighted and bootstrapped C-indexes, which were combined with Rubin's rules).[271] A sensitivity analysis was also conducted where visit 2 clinical predictors were used for individuals who had samples and marker measures from visit 2 instead of visit 1 (n=47).

RESULTS

The variables in Table 2 in the subcohort were representative of both the CCC and the eligible subpopulation from which it was derived (Table 1). The biggest difference was in the median

APRI score at baseline, which was much higher in the CCC, likely due to the 276 prevalent cases of significant liver fibrosis and end-stage liver disease (ESLD). Half the participants had been infected with HCV for 18 years (Table 1) and almost half reported drinking alcohol. The majority of the study participants were white males with a median age of 44 years. Most were receiving HIV antiretroviral therapy and had well-controlled HIV with good CD4 recovery (>350 cells/μl) and undetectable HIV viral load.

113 individuals developed significant liver fibrosis over 1300 years of risk for an event rate of 8.63 per 100 person-years (95% CI: 7.08, 10.60 per 100 py). Significant liver fibrosis cases were much more likely to be female, currently drinking alcohol, infected with HCV genotype 3 and carriers of the rs8099917 TT genotype (Table 2). Surprisingly, they were also more likely to have undetectable levels of HIV RNA at first visit. As expected, even at baseline, the median APRI value was higher among cases than those in the subcohort.

Immune marker values were available for 74% of the individuals in the analytic sample. The median value of half of the log-transformed immune markers were higher in cases than the subcohort (Figure 2). These include IL-8, MIP1α, MIP1β, TNFα, CXCL9, sICAM-1 and sVCAM-1. The remainder of the markers (RANTES, TGF-β1, MCP-1, CXCL11, sCD14 and hs-CRP) showed the opposite, that is, were lower in cases than the subcohort. The genetic marker at IFNL rs8099917 was available in 92% of the individuals included in the analysis. The pro-inflammatory TT genotype was more common in cases than the subcohort (Table 2).

The final clinical predictors in Model 1 were sex, current alcohol use (yes or no), HIV viral load (undetectable or not at ≤ 50 copies), natural log-transformed baseline APRI, HCV genotype 3 and age (Table 3). Model 1 was nested in Model 2, which also included the genetic marker IFNL rs8099917 and the following 5 natural log-transformed immune markers: IL-8, sICAM-1, RANTES, hsCRP, and sCD14 (Table 3). These markers were selected because they target different stages of liver fibrosis development, are known risk factors for fibrosis or displayed the strongest associations with the outcome in univariate analyses (Supplementary Table 3). Even in multivariate analysis, most of the selected markers had a stronger association with the outcome than any of the clinical predictors with the exception of baseline APRI (Table 3). The selected markers were also strongly correlated with the other markers that were not included.

The bootstrapped Harrell's C indexes differed between models 1 and 2, regardless of which analysis was used and whether multiple imputation was used or not (Table 3, Supplementary Figure 2). The higher values for model 2 versus model 1 indicated that adding the six markers improved the discrimination capacity beyond that of traditional clinical factors. With both analyses 1 and 2, using only clinical risk factors indicated that there was a 73% probability that predicted risk is higher for cases than noncases. That probability rose to 76% using selected markers. Other clinical factors such as CD4 count or the remaining immune markers did not substantially increase the C-index or improve model fit as measured by the AIC. Variables included in indexes from other studies such as hyaluronic acid, total bilirubin, or GGT also did not increase the C-index, improve calibration, or improve model fit.

Both models 1 and 2 were well-calibrated as measured by the Hosmer-Lemeshow test and the GB test ($p>0.05$ in both tests, Table 3 and Supplementary Figure 3). This indicated that there were no significant deviations by observed risk from subgroups of predicted risk estimates, regardless of number or location of cutpoints. Risk stratification also appeared to improve with the inclusion of the genetic and immune markers in model 2. The category-based NRI indicated that the classification improved more in those not developing fibrosis (i.e. the nonevents) than in those with the outcome: 5.1% were correctly reclassified (i.e. moved to a higher risk category) in those with events, while almost 21% of the nonevents were correctly recategorized (i.e. correctly moved to a lower risk category). The continuous NRI, on the other hand, demonstrated improvement in both, with a greater movement among cases than noncases (Table 4). Results from the sensitivity analysis, post-imputation, were very similar to those from the main analysis.

DISCUSSION

Our results demonstrate that specific markers improved ability to predict 3-year risk of significant liver fibrosis over traditional clinical factors alone in a cohort representative of HIV-HCV co-infected Canadians. Because this is a population at high risk for accelerated liver fibrosis, there is a great need for better clinical prognosis and risk assessment, which can help with prioritizing expensive HCV treatment.

We tested those clinical predictors linked with fibrosis that were most likely to be available to clinicians. HCV duration, while informative, would be hard to estimate in a real-life clinical scenario and is only approximate in our study. We chose to use age instead to capture some of

the time element that HCV duration would have provided. Furthermore, for the injection drug users in our study, age at first injection is similar, on average, and approximately 18 years before cohort entry. Due to the limits of degrees of freedom, we modeled age with 3 knots instead of 4 and included sex, current alcohol use, HIV viral load, HCV genotype 3, and baseline APRI in the model with clinical predictors only (Model 1). Other factors like CD4 count or ethnicity did not improve predictive ability. Alcohol use could have been modeled differently, but to prevent overfitting and to simulate the level of detail clinicians would have access to, we chose to keep it dichotomous.

The immune markers we screened are not routinely collected at a clinical setting and are not cheap to measure (for the 6 additional markers included in Model 2, costs were approximately $110 for reagents alone). Many of the markers were correlated with each other, so we selected those most representative of different stages of fibrosis development, those with the strongest links to fibrosis, and those most likely to improve discrimination. Even without imputation, many of the selected markers displayed stronger relationships with fibrosis than any clinical predictor studied, except for baseline APRI. Model 2, which included all the variables from Model 1, also included log-transformed IL-8, sICAM-1, RANTES, hs-CRP, and sCD14 and IFNL genotype at rs8099917. In univariate and multivariate analysis, both IL-8 and sICAM-1 acted profibrogenically, being linked with a higher rate of significant fibrosis. These markers also appeared to behave in a dose-dependent manner when examined at the median and quartile level. RANTES and hs-CRP, on the other hand, appeared protective in univariate analysis (Supplementary Table 3). RANTES, which is a natural ligand of the HIV-co receptor CCR5, is HIV-suppressive, which might account for its anti-fibrotic effect. hs-CRP is produced by the

liver which is directly affected by HCV, and thus might be produced in lower amounts when the liver is more fibrosed. We cannot infer any causality from these trends, however, as our purpose in modeling them was solely for prediction.

Other factors, such as hyaluronic acid,[91,299] total bilirubin,[298,299] or GGT,[298-300] have been included in other fibrotic indexes but did not improve model fit or prognostic ability in our study. In previous studies,[91,299] these variables were collected not for predicting risk but for diagnostic and staging purposes in individuals who had already developed liver disease. As such, they were collected at the same time point as the liver biopsy sample, which was used to determine the outcome (fibrosis or cirrhosis). In our study, however, the markers were measured in disease-free individuals up to three years before significant fibrosis set in.

Nonetheless, strong associations (e.g. high odds ratios) in etiologic models do not automatically mean that a marker improves ability to classify or predict risk for individuals, which is most relevant for guiding both clinical and policy decisions.[192,193] Moreover, models for diagnosis, where the outcome has already occurred, are not always suitable for prognosis, as it involves a future outcome making it more stochastic in nature.[270] The ideal prognostic markers, then, would need to be fairly prevalent and strongly linked with the outcome, and also improve discrimination and calibration beyond traditional risk factors alone.[276] They would also help reclassify risk, moving more participants to the extremes of the risk distribution, where there are clearer implications for future actions.

Using these chosen six markers improved our ability to discriminate between those who develop significant liver fibrosis and those who do not, as indicated by the higher C-indexes in model 2 versus model 1. The improvement in discrimination due to addition of immune markers was minor and not statistically significant. Discrimination by both models as measured by Harrell's C-index was in the 0.70-0.80 range, similar to other comparable indexes. To put this in context, a prognostic score for prediction of ESLD in HIV-HCV co-infected individuals on cART had a C-statistic of 0.73, which can be considered good and of some clinical utility.[282] Other prognostic indexes which sought to predict risk of mortality in cirrhotics (Child-Pugh score, MELD score) or in HIV-infected individuals on cART (VACS Index)[283] had C-statistics around 0.80 or higher, which could be considered very good or excellent.[284,285]

However, while improvement in discrimination was modest, inclusion of immune markers improved the net reclassification improvement index, as indicated by both category-based and continuous NRI estimates. Results from the category-based NRI indicate that measuring the markers in model 2 correctly reduced risk estimates in those who did not develop liver fibrosis. While this does not enable identification of higher-risk individuals for treatment,[305] it could help reduce overtreatment in up to 21% of individuals who eventually did not develop fibrosis over 3 years. For these individuals, a different clinical plan such as delaying treatment or a less frequent follow-up, for example, might be reasonable. The continuous NRI, on the other hand, indicated that model 2 improved ability to identify higher-risk individuals for treatment. However, these changes in predicted risk might not be clinically significant, as values of continuous NRIs are often higher than their category-based counterparts.[306] NRIs are most useful not as a single measure, but in the context of other measures of discrimination and calibration.[291] When viewed

together with the Harrell's C-index, the NRI does seem to support improved discrimination and risk classification with inclusion of IL-8, sICAM-1, RANTES, hs-CRP, and sCD14 and IFNL genotype at rs8099917. A cost-benefit analysis that takes into account the expense of these markers as well as the high cost of HCV treatment might be useful as a next step. These results, however, must be interpreted with caution. Category-based NRIs are sensitive to the number of risk categories as well as the chosen cutpoints,[291] while continuous NRIs can have large values for even weak markers.[306]

The strength of our study includes a large source population that is representative of Canadian co-infected patients, thus making our results directly generalizable and clinically relevant. This is important as features such as marker and disease prevalence or marker correlations with other known risk factors can all affect the estimates of the discriminatory capacity of a marker.[13] Inclusion of specific markers in model 2 appeared to provide higher discrimination than the clinical predictors in model 1 without multiple imputation or whether the analysis for the imputed data was as a case cohort with weighted Harrell's C-index or as a full cohort with Rubin's rules (Supplementary Figure 2). Furthermore, our case cohort design provides an efficient, economical way to get a more representative picture of baseline distribution of immune markers for the full cohort and allows us to directly estimate baseline risk of liver disease in co-infected Canadians. Finally, the list of immune markers we tested target different stages of the underlying mechanism of fibrosis progression, enhancing our ability to capture the outcome at various stages of development.

The limitations of our study include having marker measures at only visit 1 or 2, which means we are unable to assess the predictive value of markers at other time points or measure the predictive value of changes of marker levels. However, as we were interested in prognosis, using the first available sample mirrors what would occur in clinical practice when a patient is first evaluated. We also lacked the power to assess interaction or effect measure modification of markers in different subgroups. Furthermore, our study modeled the markers specifically for prognostic assessment precluding causal inferences. Finally, external validation in an independent dataset should be performed before applying this model clinically. Data provided in Supplementary Tables 2 and 4 will be especially relevant to future external validation studies.

In conclusion, we found that, in an HIV-HCV co-infected population, incorporating a genetic marker from IFNL rs8099917 and the immune markers IL-8, sICAM-1, RANTES, hs-CRP, and sCD14 allowed us to better predict the 3-year risk of significant liver fibrosis over traditional clinical risk factors alone. While the improvement in discrimination was small, the model with the markers also classified risk and fit better than the one without the markers. To assess whether this improvement justifies the additional cost of measuring these markers in the face of highly expensive HCV treatment requires further cost-benefit analyses.

TABLES AND FIGURES

Table 1. Baseline Characteristics of the Canadian Co-infection Cohort and Study Population

Characteristic	CCC, n=1119	Eligible cohort, n=679
Age at baseline, years	45 (39-50)	44 (39-49)
White	855 (77)	521 (77)
Female	291 (26)	187 (28)
Currently drinking alcohol	566 (51)	333 (49)
APRI	0.63 (0.38-1.24)	0.52 (0.36- 0.78)
Receiving HIV antiretrovirals	903 (81)	538 (79)
CD4 count, cells/μl	380 (249- 550)	400 (270-568)
Undetectable HIV viral load, (<50 copies/ml)	682 (61)	395 (59)
HCV duration, years	18 (11-26)	18 (10-25)
HCV genotype 3	166 (19)[a]	87 (16)[b]

Abbreviations: CCC, Canadian Co-infection Cohort; APRI, aspartate aminotransferase (AST) to platelet ratio index, calculated as follows: [(AST/upper limit of normal)/platelet count (109 /L)] x 100; HCV, Hepatitis C virus.

Presented as n(%) or Median (Interquartile Range).

a. HCV genotype data available in only 874 individuals

b. HCV genotype data available in only 549 individuals

Table 2. Baseline Characteristics of the Subcohort and Cases Outside Subcohort

Characteristic	Subcohort, n=236	Cases outside subcohort, n=59
Age, years	44 (39-49)	44 (39-49)
White	182 (78)	48 (81)
Female	70 (30)	20 (34)
Currently drinking alcohol	114 (48)	32 (54)
APRI	0.52 (0.36-0.81)	0.70 (0.47-0.97)
IFNL genotype rs8099917 TT	127 (60)	41 (70)
Receiving HIV antiretrovirals	191 (81)	46 (78)
Undetectable HIV viral load, (<50 copies/ml)	135 (59)	37 (65)
CD4 count, cells/ μl	380 (250-540)	377 (230-540)
HCV duration, years	18 (11-26)	18 (12-24)
HCV genotype 3	30 (16)[a]	13 (26)[b]

Abbreviations: APRI, aspartate aminotransferase (AST) to platelet ratio index, calculated as follows: [(AST/upper limit of normal)/platelet count (109 /L)] x 100; IFNL, Interferon Lambda; HCV, Hepatitis C virus.

Presented as n(%) or Median (Interquartile Range).

a. HCV genotype data available in only 189 individuals

b. HCV genotype data available in only 50 individuals

Table 3. Predictive Accuracy and Multivariate Results of Cox Models Analyzing the Association of Significant Liver Fibrosis [HR (95% CI)] Before and After Multiple Imputation as Cohort with Rubin's Rules (Analysis 1)

	Before imputation		After imputation[a]	
Predictive Accuracy	**Model 1**	**Model 2**	**Model 1**	**Model 2**
Discrimination:				
Harrell's C-index (95% CI)	0.731 (0.647, 0.815)	0.819 (0.740, 0.899)	0.730 (0.670, 0.789)	0.762 (0.703, 0.820)
Calibration (p-values)[b]				
Hosmer-Lemeshow	0.32	0.30	0.37	0.47
Gronnesby and Borgan (GB) test	0.76	0.59	0.47	0.88
Included Predictors	**Model 1**	**Model 2**	**Model 1**	**Model 2**
Female	1.11 (0.64, 1.94)	1.34 (0.54, 3.34)	1.25 (0.82, 1.90)	1.35 (0.82, 2.21)
Current alcohol use	1.25 (0.73, 2.12)	0.90 (0.46, 1.75)	1.31 (0.89, 1.92)	1.30 (0.85, 2.00)
HIV viral load [i]	1.43 (0.84, 2.44)	1.51 (0.78, 2.95)	1.17 (0.79, 1.74)	1.21 (0.80, 1.85)
Log Baseline APRI [ii]	3.43 (1.92, 6.12)	2.91 (1.54, 5.50)	3.19 (2.05, 4.96)	2.71 (1.72, 4.26)
Age	0.99 (0.92, 1.06)	1.00 (0.91, 1.11)	1.00 (0.95, 1.05)	0.99 (0.95, 1.04)
Age*	0.98 (0.91, 1.05)	0.94 (0.86, 1.04)	0.99 (0.93, 1.04)	0.99 (0.93, 1.05)
HCV genotype 3[iii]	1.37 (0.73, 2.57)	1.04 (0.44, 2.48)	1.34 (0.80, 2.25)	1.36 (0.79, 2.36)
rs8099917 TT [iv]	--	2.12 (1.01, 4.46)	--	1.39 (0.90, 2.16)
IL-8 [v]	--	2.09 (1.44, 3.04)	--	1.48 (1.08, 2.02)
sICAM-1 [vi]	--	3.85 (1.70, 8.75)	--	2.04 (1.05, 3.97)
RANTES [vii]	--	0.58 (0.38, 0.88)	--	0.83 (0.64, 1.07)
hsCRP [viii]	--	0.95 (0.74, 1.23)	--	0.95 (0.78, 1.16)
sCD14 [ix]	--	0.36 (0.11, 1.19)	--	0.56 (0.24, 1.30)

a. Using Analysis 1, regression coefficients and standard errors combined with Rubin's rules

b. Using quintiles of risk. Results similar with tertiles.

i. Missing in 3% cases and noncases

ii. Missing in 6% of cases and 5% of noncases

iii. Missing in 14% of cases and 22% of noncases

iv. Missing in 1% of cases and 13% of noncases

v. Missing in 24% of cases and 28% of noncases

vi. Missing in 24% of cases and 28% of noncases

vii. Missing in 24% of cases and 28% of noncases

viii. Missing in 24% of cases and 30% of noncases

ix. Missing in 24% of cases and 28% of noncases

Abbreviations: HR, hazard ratio; CI, confidence interval; APRI, aspartate aminotransferase (AST) to platelet ratio index, calculated as follows: [(AST/upper limit of normal)/platelet count (109 /L)] x 100; HCV, Hepatitis C virus; IL-8, interleukin-8; sICAM-1, soluble intercellular adhesion molecule 1; RANTES, Regulated upon Activation, Normal T cell Expressed and Secreted protein; hsCRP high-sensitivity C-reactive protein; sCD14, soluble CD14.

Table 4. Net Reclassification Improvement (NRI) index for 3-year Risk of Significant Liver Fibrosis Using Model 1 (Clinical Predictors Only) and Model 2 (Model 1+ Selected Markers)

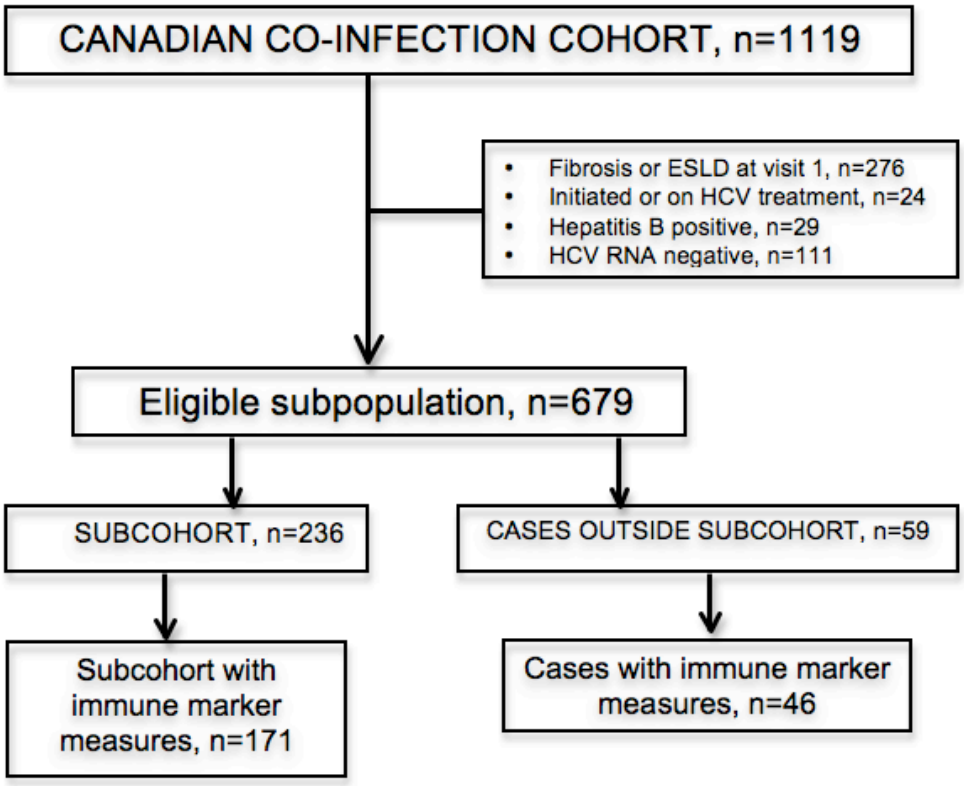
		MODEL 2				Reclassified				
	MODEL 1	<=10	>10-25	>25	Total	N	Higher	Lower	Net	**NRI**
Developed fibrosis**	<=10	2	0	1	3					
	>10-25	4	6	11	21		12 20.3%	9 15.2%	5.1	
	>25	1	4	30	35					
	Total	7	10	42	**59****	150				**26**
No fibrosis	<=10	16	5	2	23					
	>10-25	24	10	11	45		18 19.8%	37 40.6%	20.9	
	>25	4	9	10	23					
	Total	44	24	23	**91**					

Purple=movement to a lower risk category; Green=movement to a higher risk category

**Without multiple imputation, only 59 cases had event times within 3 years and had complete information to be used in the analysis. Results were similar if multiple imputation was used to complete information for 94 cases (not shown).

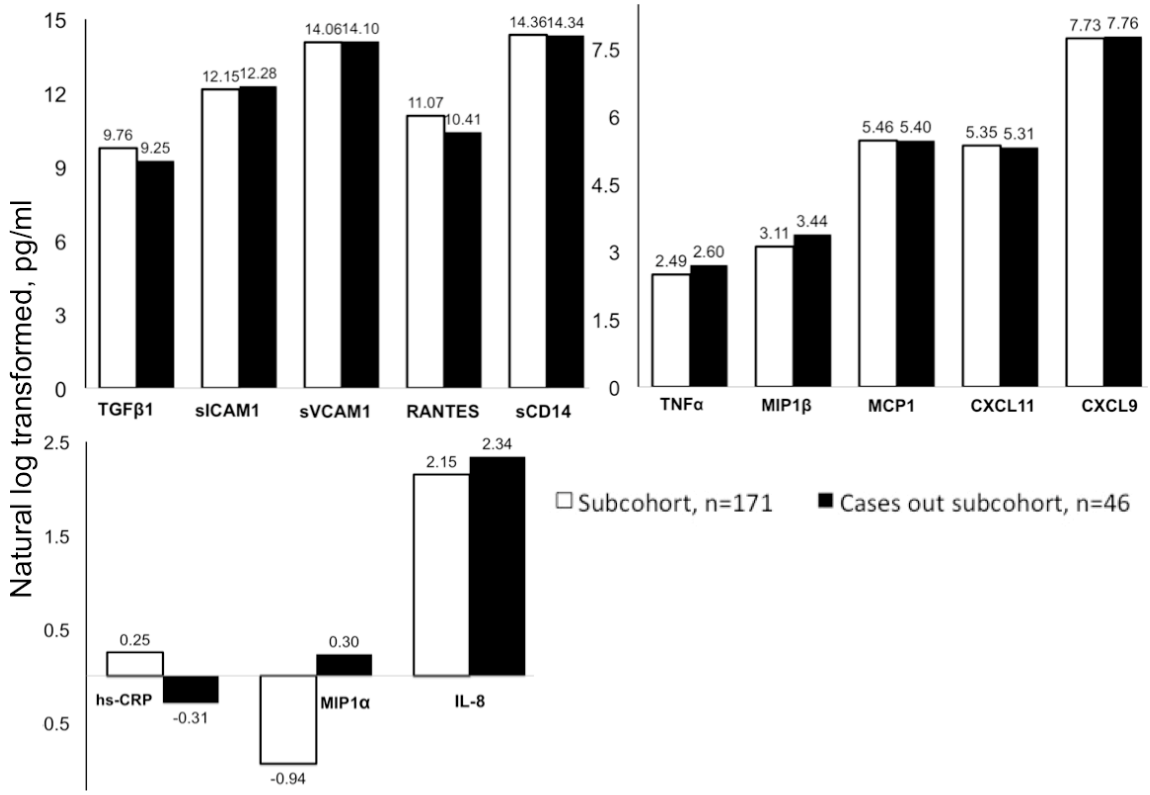
	Category-based NRI (p-value)	**Continuous NRI (p-value)**
Events	0.051 (0.51)	0.356 (0.006)
Nonevents	0.209 (0.010)	0.187 (0.075)
Overall	0.26 (0.02)	0.543 (0.001)

Figure 1. Source and Study Population for Developing a Prognostic Model for Significant Liver Fibrosis



Abbreviations: ESLD, end-stage liver disease; HCV, Hepatitis C virus.

Figure 2. Medians of Log-Transformed Immune Markers: Subcohort vs. Cases Outside Subcohort



Abbreviations: TGF-β1, transforming growth factor beta 1; ; sICAM-1, soluble intercellular adhesion molecule 1; sVCAM-1, soluble vascular cell adhesion molecule 1; RANTES, Regulated upon Activation, Normal T cell Expressed and Secreted protein; sCD14, soluble CD14; TNF-α, tumor necrosis factor alpha; MIP1β, macrophage inflammatory protein 1 beta; MCP-1, monocyte chemotactic protein-1; CXCL11, chemokine (C-X-C motif) ligand 11; CXCL9, chemokine (C-X-C motif) ligand 9; hsCRP high-sensitivity C-reactive protein; MIP1α, macrophage inflammatory protein 1 alpha; IL-8, interleukin-8.

APPENDIX

Supplementary Table 1. Summary of Immune Markers of Interest

		Description	Ref
Cytokines	**TGF-β1**	• Triggered directly by HIV proteins • Called most pro-fibrotic cytokine: directly stimulates molecules that are imbalanced in fibrosis	58,77,79,196
	TNF-α	• Elevated in livers of co-infected individuals • Produced by liver macrophages in response to viruses, alcohol, LPS • Upregulates HIV replication	197,198,200,201
Chemokines	**IL-8**	• Serum levels linked with HCV disease progression and interferon unresponsiveness • In cell cultures, HCV and HIV proteins triggered IL-8 in a dose-dependent manner	207-210
	MCP-1	• Not expressed in healthy livers but triggered by HCV infection • Stimulates fibrogenesis	211,212
	MIP1α **MIP1β**	• Binds CCR5, receptor needed for HIV entry • CCR5-MIP1α pairs dominate during HCV infection • Levels higher in co-infected livers than mono-infected	178 212
	RANTES	• Binds CCR5 • Stimulates fibrogenesis • Higher in co-infected livers vs. monoinfected	178,211
	CXCL9	• Correlated to histologic liver disease activity after HCV infection • Significantly elevated in plasma in those with advanced liver necroinflammation	214,307,308
	CXCL11	• Has chemoattractant properties that initiate and perpetuate liver inflammation-guides T cells through inflamed liver • Correlated to histologic disease activity in the liver after HCV infection	214,309
Endothelial activation markers	**sICAM-1** **sVCAM1**	• Higher in HIV-infected individuals • Associated with higher HIV viral load and faster HIV disease progression • Higher in cirrhotic patients	215,216
	hsCRP	• Produced in the liver • Released during the acute phase of an infection • Reproducible, dynamic reflection of ongoing tissue injury • In HIV+ve individuals, higher levels linked with higher risk of opportunistic infections and all-cause mortality • Trends unclear with HIV-HCV co-infection	202-206
	sCD14	• Marker of microbial translocation • Co-receptor to LPS, a component of bacterial cell walls • Release in soluble form triggers immune activation and cytokine production • In HIV+ve, highest sCD14 levels linked with 6-fold higher risk of mortality vs. lowest levels of sCD14 • In HCV mono and co-infection, higher levels linked with lower treatment response and cirrhosis • In co-infected individuals, highest quartile linked with 8-fold higher odds of cirrhosis vs. lowest quartile	184,186,188,189,217,218

Abbreviations: TGF-β1, transforming growth factor beta 1; TNF-α, tumor necrosis factor alpha; IL-8, interleukin-8; LPS, lipopolysaccharide; MCP-1, monocyte chemotactic protein-1; MIP1α, macrophage inflammatory protein 1 alpha; MIP1β, macrophage inflammatory protein 1 beta; RANTES, Regulated upon Activation, Normal T cell Expressed and Secreted protein; CXCL9, chemokine (C-X-C motif) ligand 9; CXCL11, chemokine (C-X-C motif) ligand 11; sICAM-1, soluble intercellular adhesion molecule 1; sVCAM-1, soluble vascular cell adhesion molecule 1; hsCRP high-sensitivity C-reactive protein; sCD14, soluble CD14.

Supplementary Table 2. Distribution of Immune Markers in Random Subcohort and Cases outside Subcohort

	Subcohort, n=171	Cases outside subcohort, n=46
IL-8		
Average (SD)	11.87 (12.54)	14.47 (12.7)
Median (IQR)	8.58 (5.20-13.94)	10.37 (7.51-16.92)
sICAM-1		
Average (SD)	214,405 (92,171)	231,504 (89,213)
Median (IQR)	189,927 (158,884-250,298)	216,092 (182,731- 258,416)
RANTES		
Average (SD)	68,786.33 (52,971)	57,956.3 (65,551)
Median (IQR)	63,970 (25,527-97,741)	33,058 (18,120-73,487)
sCD14		
Average (SD)	1.79 $x10^6$ (571,015)	1.76 $x10^6$ (520,672)
Median (IQR)	1.72 $x10^6$ (1.44 $x10^6$-2.03 $x10^6$)	1.69 $x10^6$ (1.45 $x10^6$- 2.11 $x10^6$)
hsCRP		
Average (SD)	3.14 $x10^6$ (5.77$x10^6$)	1.59 $x10^6$ (1.84 $x10^6$)
Median (IQR)	1.29 $x10^6$ (0.49 $x10^6$-2.74 $x10^6$)	0.74 $x10^6$ (0.29 $x10^6$-2.10 $x10^6$)
MCP-1		
Average (SD)	278.97 (179)	267.80 (157)
Median (IQR)	235 (150-344)	223 (157-357)
MIP1α		
Average (SD)	2.78 (5.40)	3.08 (3.62)
Median (IQR)	0.39 (0.28-3.46)	1.37 (0.28-5.41)
MIP1β		
Average (SD)	29.71 (25.25)	38.13 (33.07)
Median (IQR)	22.33 (13.46- 37.02)	31.35(17.16-47.55)
TNFα		
Average (SD)	14.02 (9.15)	15.30 (9.07)
Median (IQR)	12.07 (7.63-17.73)	20.19 (13.44- 24.53)
sVCAM-1		
Average (SD)	1,310,723 (434,816.8)	1,382,531 (393,647)
Median (IQR)	1.28 $x10^6$ (1.02 $x10^6$-1.51$x10^6$)	1.34 $x10^6$ (1.10$x10^6$-1.71$x10^6$)
CXCL9		
Average (SD)	2875.94 (2303.19)	3123 (2216.94)
Median (IQR)	2265 (1444-3619)	2337 (1567-4156)
CXCL11		
Average (SD)	274.50 (231)	316.30 (279)
Median (IQR)	210 (116-360)	202 (134-425)
TGF-β1		
Average (SD)	21194.81 (15866)	19,707.1 (20,503)
Median (IQR)	17385 (8731-29232)	10,406 (5668-32,814)

Units: all in pg/ml

Abbreviations: SD, standard deviation; IQR, interquartile range; IL-8, interleukin-8; sICAM-1, soluble intercellular adhesion molecule 1; RANTES, Regulated upon Activation, Normal T cell Expressed and Secreted protein; sCD14, soluble CD14; hsCRP high-sensitivity C-reactive protein; MCP-1, monocyte chemotactic protein-1; MIP1α, macrophage inflammatory protein 1 alpha; MIP1β, macrophage inflammatory protein 1 beta; TNF-α, tumor necrosis factor alpha; sVCAM-1, soluble vascular cell adhesion molecule 1; CXCL9, chemokine (C-X-C motif) ligand 9; CXCL11, chemokine (C-X-C motif) ligand 11; TGF-β1, transforming growth factor beta 1.

Supplementary Table 3. Univariate Results of the Association of Significant Liver Fibrosis with Log-transformed Immune markers [HR (95% CI)]

	Not imputed	After Multiple Imputation
IL-8	1.59 (1.20, 2.11)	1.37 (1.08, 1.74)
MCP-1	1.01 (0.71, 1.46)	1.06 (0.75, 1.49)
MIP1α	1.04 (0.91, 1.20)	1.09 (0.93, 1.20)
MIP1β	1.21 (0.88, 1.65)	1.16 (0.89, 1.51)
TNF-α	1.21 (0.84, 1.75)	1.07 (0.76, 1.51)
sICAM-1	3.22 (2.00, 5.18)	2.66 (1.53, 4.61)
sVCAM-1	2.38 (1.35, 4.20)	1.88 (1.03, 3.41)
RANTES	0.79 (0.63, 0.98)	0.93 (0.76, 1.14)
CXCL9	1.43 (1.08, 1.90)	1.22 (0.91,1.64)
CXCL11	1.30 (1.00, 1.70)	1.26 (1.00, 1.59)
TGF-β1	0.83 (0.64, 1.08)	0.96 (0.76, 1.20)
hsCRP	0.81 (0.68, 0.96)	0.89 (0.76, 1.05)
sCD14	1.14 (0.62, 2.08)	1.08 (0.62, 1.88)

Abbreviations: HR, hazard ratio; CI, confidence interval; IL-8, interleukin-8; MCP-1, monocyte chemotactic protein-1; MIP1α, macrophage inflammatory protein 1 alpha; MIP1β, macrophage inflammatory protein 1 beta; TNF-α, tumor necrosis factor alpha; sICAM-1, soluble intercellular adhesion molecule 1; sVCAM-1, soluble vascular cell adhesion molecule 1; RANTES, Regulated upon Activation, Normal T cell Expressed and Secreted protein; CXCL9, chemokine (C-X-C motif) ligand 9; CXCL11, chemokine (C-X-C motif) ligand 11; TGF-β1, transforming growth factor beta 1; hsCRP high-sensitivity C-reactive protein; sCD14, soluble CD14.

Supplementary Table 4. Estimated Regression Coefficients and Standard Errors for Calculating Risk Score [Estimated beta (SE)] in final Models 1 and 2

	Before imputation		After imputation[a]	
	Model 1	**Model 2**	**Model 1**	**Model 2**
Female	0.11 (0.28)	0.29 (0.47)	0.22 (0.21)	0.30 (0.25)
Current alcohol use	0.22 (0.27)	-0.10 (0.34)	0.27 (0.20)	0.26 (0.22)
HIV viral load	0.36 (0.27)	0.41 (0.34)	0.16 (0.20)	0.19 (0.21)
Log Baseline APRI	1.23 (0.30)	1.07 (0.32)	1.16 (0.22)	0.997 (0.23)
Age	-0.012 (0.04)	0.003 (0.05)	-0.004 (0.02)	-0.007 (0.03)
Age*	-0.018 (0.04)	-0.058 (0.05)	-0.013 (0.03)	-0.01 (0.03)
HCV genotype 3	0.32 (0.32)	0.04 (0.44)	0.29 (0.26)	0.31 (0.28)
rs8099917 TT	--	0.75 (0.38)	--	0.33 (0.22)
IL-8	--	0.74 (0.19)	--	0.39 (0.16)
sICAM-1	--	1.35 (0.42)	--	0.71 (0.34)
RANTES	--	-0.55 (0.21)	--	-0.19 (0.13)
hsCRP	--	-0.05 (0.13)	--	-0.05 (0.11)
sCD14	--	-1.02 (0.61)	--	-0.58 (0.43)

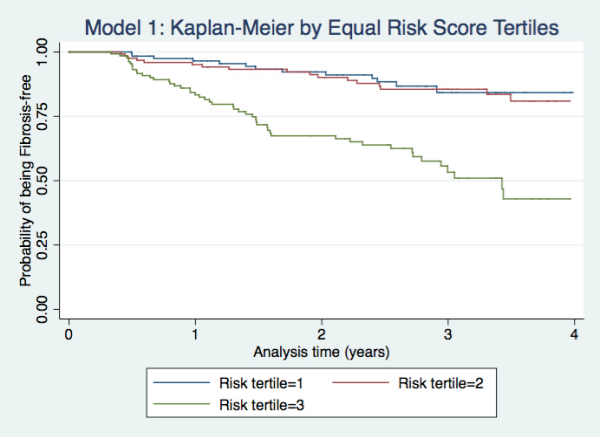
Abbreviations: APRI, aspartate aminotransferase (AST) to platelet ratio index, calculated as follows: [(AST/upper limit of normal)/platelet count (109 /L)] x 100; HCV, Hepatitis C virus; IL-8, interleukin-8; sICAM-1, soluble intercellular adhesion molecule 1; RANTES, Regulated upon Activation, Normal T cell Expressed and Secreted protein; hsCRP high-sensitivity C-reactive protein; sCD14, soluble CD14.

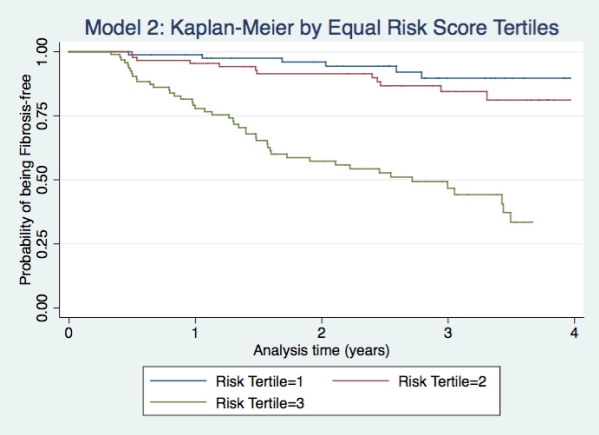
a. Using Analysis 1, combined with Rubin's rules

In the final models, the natural log transformations of the continuous variables [immune markers (IL-8, sICAM-1, RANTES, hsCRP, and sCD14) and the baseline APRI] and centering at the means are used. The form of the other included predictors are detailed in "Predictor selection and functional form" in the Methods section.

The risk score is constructed from the linear predictors of the Cox model. The linear predictor is a weighted sum of the variables in the final model, where the weights are the regression coefficients. High values indicate a higher risk of significant liver fibrosis. The risk score for an individual is then the log relative hazard compared with a hypothetical individual whose risk score is zero.[290] In our dataset, an individual with a risk score zero is a 45-year-old male who is not currently drinking alcohol, with a detectable HIV viral load and host IFNL rs8099917 non-TT genotype, infected with HCV genotype 1, 2 or 4 and with mean values for APRI as well as the immune markers (IL-8, sICAM-1, RANTES, hsCRP, sCD14) at baseline.

Supplementary Figure 1. Tertiles of Risk Score Against Kaplan-Meier Estimates of Risk of Significant Liver Fibrosis

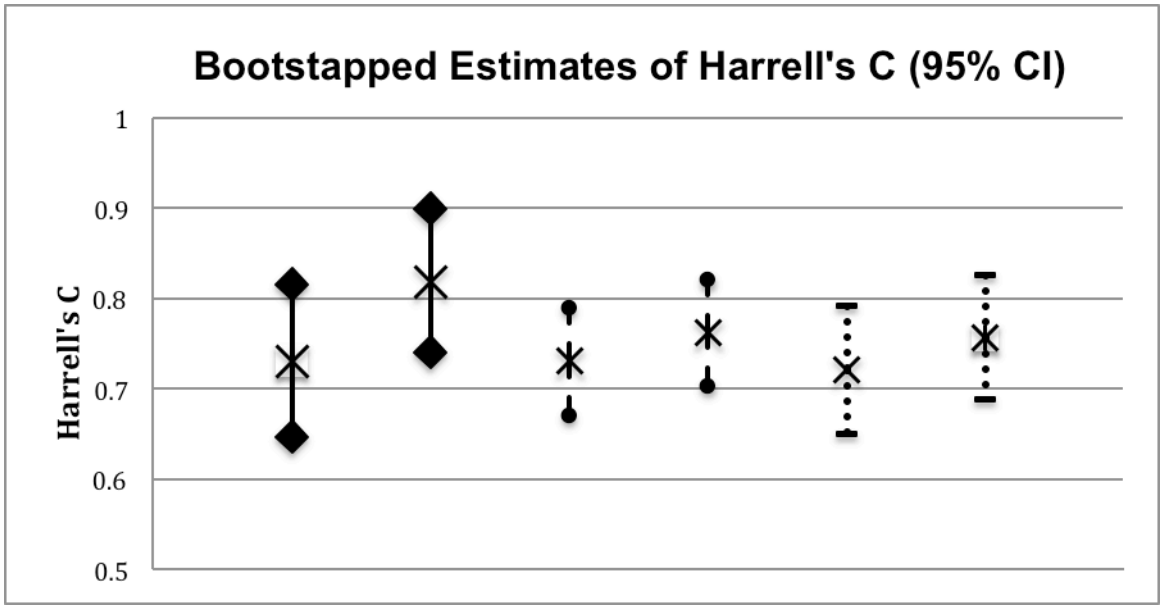




Blue=score tertile 1 (lowest risk); Red=score tertile 2; Green=score tertile 3 (highest risk)

The wider separation of the Kaplan-Meier curves in Model 2 provides a visual representation of the improvement in discrimination with the addition of the markers compared to Model 1.

Supplementary Figure 2. Bootstrapped Estimates of weighted Harrell's C (95% CI) of Prognostic Models 1 and 2



Lines represent point estimates of bootstrapped weighted Harrell's C and 95% Confidence Intervals.

No multiple imputation, solid line: Models 1 and 2--0.731 (0.647, 0.815) and 0.819 (0.740, 0.899)

Post multiple imputation,

- --- -- Dashed line: **Analysis 1, Models 1 and 2**--0.730 (0.670, 0.789) and 0.762 (0.703, 0.820)

..... Dotted line: **Analysis 2, Models 1 and 2**-- 0.720 (0.649, 0.791) and 0.756 (0.688, 0.825)

a) **Model 1**= clinical predictors only [sex, current alcohol use, HIV RNA, baseline APRI, HCV genotype 3, and age as a restricted cubic spline with 3 knots]

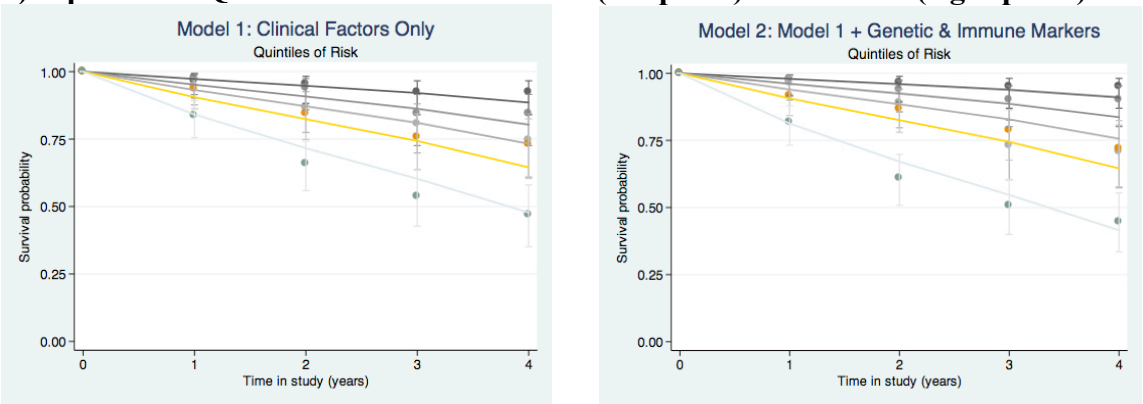
b) **Model 2**= Model 1 + *IFNL3* SNP rs8099917+ 5 selected immune markers [IL-8, sICAM-1, RANTES, hs-CRP, and sCD14]

c) **Analysis 1:** Multiple imputation analysis for full cohort (n=679), no weighting

d) **Analysis 2:** Case cohort generated in each imputed set, analyzed with Barlow weights, bootstrapped, combined using Rubin's rules

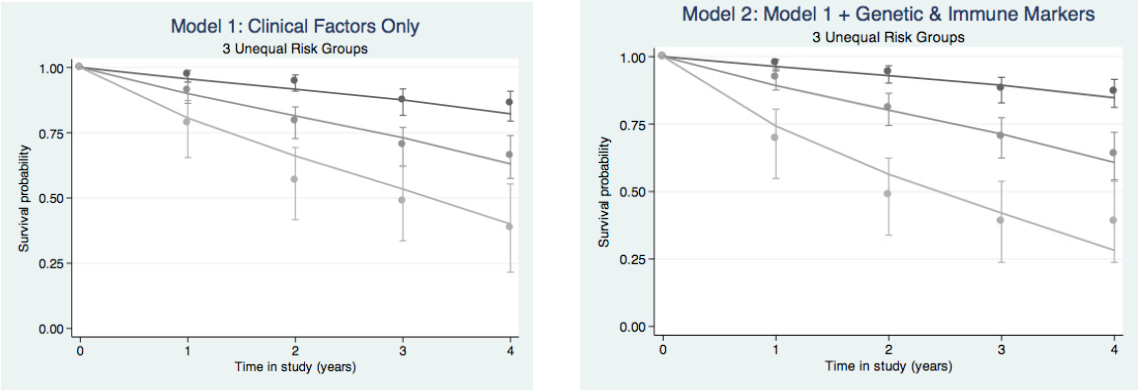
Supplementary Figure 3. Calibration with Predicted Survival Curves and Kaplan-Meier Estimates in Model 1 (Clinical Predictors Only) and Model 2 (Model 1 plus rs809917 and 5 Selected Immune Markers)

a) Equal-sized Quintiles of Risk in Model 1 (left panel) vs. Model 2 (right panel)



Smooth lines represent predicted survival probabilities, and vertical capped lines denote Kaplan–Meier estimates with 95% confidence intervals. Five prognosis groups are plotted: the "Good" group (darkest lines) and the "Poor" group (palest lines) at the highest and lowest risk categories, respectively, with the other 3 in between.

b) 3 Unequal Risk Groups (Cut at the 25th and the 75th Percentiles of the Failure Times)



Smooth lines represent predicted survival probabilities, and vertical capped lines denote Kaplan–Meier estimates with 95% confidence intervals. Three prognosis groups are plotted: the "Good" group (darkest lines), the "Intermediate" group (medium-dark lines), and the "Poor" group (paler lines).

6A. APPENDIX FOR CHAPTER 6 (Manuscript 3)

Appendix Table 6.1 Correlation Matrix of Clinical Predictors from Model 1 at Baseline in the Subcohort

	Female	Alcohol use	HIV RNA	Log APRI	Age
Alcohol use	-0.11	1.00			
HIV RNA	0.02	-0.14*	1.00		
Log APRI	-0.03	0.06	-0.07	1.00	
Age	-0.17*	0.12	0.20*	-0.12	1.00
HCV genotype 3	-0.12	0.05	0.03	0.00	0.07

Abbreviations: APRI, aspartate aminotransferase (AST) to platelet ratio index, calculated as follows: [(AST/upper limit of normal)/platelet count (109 /L)] x 100; HCV, hepatitis C virus.

*** yellow box= Statistical significance (p-value<0.05)**

Appendix Table 6.2 Correlation Matrix of Prospective Log-Transformed Markers for Model 2 and Clinical Predictors from Model 1 at Baseline in the Subcohort

	Female	Alcohol use	HIV RNA	Log APRI	Age	HCV genotype 3
rs8099917 TT	0.02	0.00	-0.05	0.06	-0.03	-0.05
IL8	0.03	-0.02	-0.03	-0.08	0.10	-0.05
sICAM1	0.05	0.07	-0.06	0.29*	-0.04	0.00
RANTES	-0.13	0.11	-0.06	-0.16*	-0.02	-0.04
TGF-β1	-0.13	0.14	-0.06	-0.17*	0.08	-0.03
hsCRP	-0.05	-0.09	0.23*	-0.15	0.19*	0.09
sCD14	0.34*	-0.02	0.04	0.01	0.02	0.09
MCP1	-0.02	0.09	0.04	-0.03	0.04	0.02
MIP1α	-0.07	0.13	0.07	0.08	0.04	0.03
MIP1β	0.06	0.06	0.18*	0.00	0.09	0.00
TNFα	0.05	0.12	-0.13	0.03	-0.08	-0.11
sVCAM1	0.06	0.02	-0.12	0.31*	0.01	-0.08
CXCL9	0.12	0.17*	-0.33*	0.15	-0.14	-0.12
CXCL11	0.02	0.08	-0.23*	0.09	-0.15	-0.10

Abbreviations: APRI, aspartate aminotransferase (AST) to platelet ratio index, calculated as follows: [(AST/upper limit of normal)/platelet count (109 /L)] x 100; HCV, hepatitis C virus; IL-8, interleukin-8; sICAM-1, soluble intercellular adhesion molecule 1; RANTES, Regulated upon Activation, Normal T cell Expressed and Secreted protein; TGF-β1, transforming growth factor beta 1; hsCRP high-sensitivity C-reactive protein; sCD14, soluble CD14; MCP-1, monocyte chemotactic protein-1; MIP1α, macrophage inflammatory protein 1 alpha; MIP1β, macrophage inflammatory protein 1 beta; TNF-α, tumor necrosis factor alpha; sVCAM-1, soluble vascular cell adhesion molecule 1; CXCL9, chemokine (C-X-C motif) ligand 9; CXCL11, chemokine (C-X-C motif) ligand 11.

*** yellow box= Statistical significance (p-value<0.05)**

Appendix Table 6.3 Correlation Matrix of Prospective Log-Transformed Markers for Model 2 at Baseline in the Subcohort

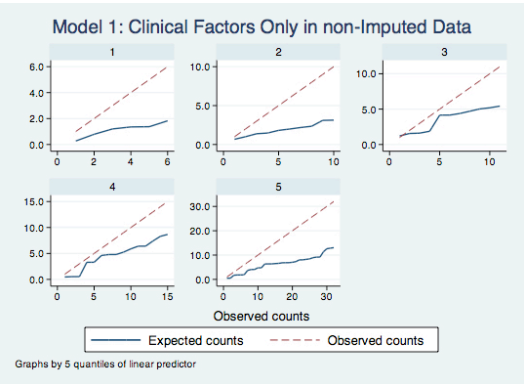
	rs8099917 TT	IL8	sICAM1	RANTES	TGFβ	hsCRP	sCD14	MCP1	MIP1α	MIP1β	TNF-α	sVCAM1	CXCL9
IL8	0.03	1.00											
sICAM1	-0.04	0.27*	1.00										
RANTES	0.01	0.38*	0.10	1.00									
TGFβ	0.06	0.48*	0.05	0.77*	1.00								
hsCRP	0.06	0.06	-0.02	-0.02	-0.03	1.00							
sCD14	-0.10	0.20*	0.35*	0.02	0.03	0.13	1.00						
MCP1	0.02	0.43*	0.23*	0.25*	0.36*	-0.01	0.14	1.00					
MIP1α	-0.13	0.26*	0.19*	0.19*	0.17*	0.00	0.12	0.20*	1.00				
MIP1β	0.04	0.38*	0.15	0.18*	0.30*	0.04	0.11	0.35*	0.38*	1.00			
TNFα	0.05	0.42*	0.29*	0.23*	0.22*	0.04	0.09	0.50*	0.23*	0.39*	1.00		
sVCAM1	-0.09	0.14	0.54*	0.06	-0.05	-0.10	0.23*	0.24*	0.18*	0.07	0.34*	1.00	
CXCL9	0.04	0.25*	0.40*	0.18*	0.14	-0.05	0.15	0.16*	0.18*	0.10	0.45*	0.40*	1.00
CXCL11	-0.05	0.47*	0.26*	0.62*	0.58*	-0.11	0.04	0.22*	0.24*	0.15*	0.33*	0.21*	0.44*

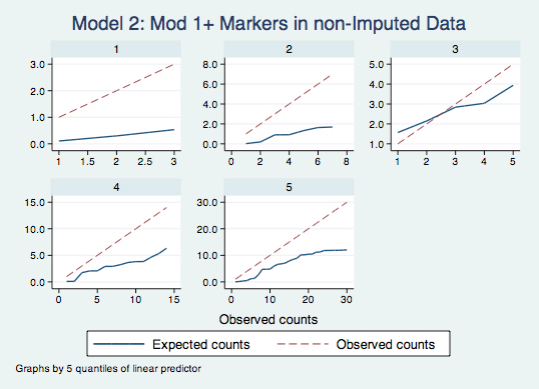
Abbreviations: IL-8, interleukin-8; sICAM-1, soluble intercellular adhesion molecule 1; RANTES, Regulated upon Activation, Normal T cell Expressed and Secreted protein; TGF-β, transforming growth factor beta; hsCRP high-sensitivity C-reactive protein; sCD14, soluble CD14; MCP-1, monocyte chemotactic protein-1; MIP1α, macrophage inflammatory protein 1 alpha; MIP1β, macrophage inflammatory protein 1 beta; TNF-α, tumor necrosis factor alpha; sVCAM-1, soluble vascular cell adhesion molecule 1; CXCL9, chemokine (C-X-C motif) ligand 9; CXCL11, chemokine (C-X-C motif) ligand 11.

*** yellow box= Statistical significance (p-value<0.05)**

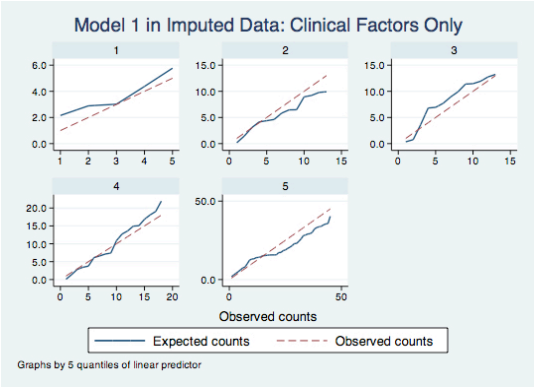
Appendix Figure 6.1 Calibration with Arjas-like Plots of Predicted and Observed Events in Model 1 (Clinical Predictors Only) and Model 2 (Model 1 plus rs809917 and 5 Selected Immune Markers) from the GB Test

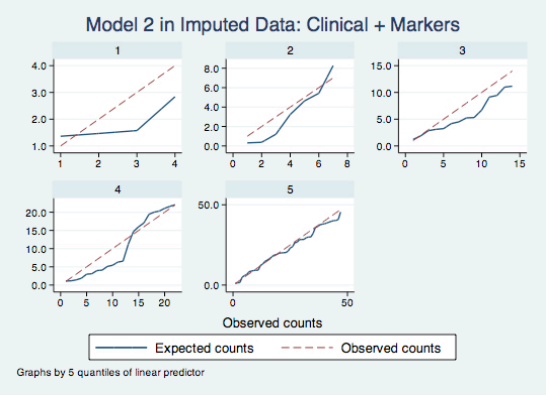
a) In Non-Imputed Data: Equal-sized Quintiles of Risk in Model 1 (left panel) vs. Model 2 (right panel)





b) In imputed Data: Equal-sized Quintiles of Risk in Model 1 (left panel) vs. Model 2 (right panel)





CHAPTER 7. Discussion and Conclusion

7.1 SUMMARY OF FINDINGS

Adverse outcomes like lower HCV spontaneous clearance and faster liver fibrosis progression are more common in HIV-HCV co-infected individuals than their HCV mono-infected counterparts. Two reasons for this are immune dysregulation and heightened inflammation, some of which can be alleviated with viral cure achieved with effective HCV treatment. However, newer, more efficacious HCV DAAs are highly expensive and low treatment uptake and HCV re-infection also remain problems. Given these challenges, our findings could aid in better clinical decision-making and also improve understanding of HCV etiology in co-infected individuals. In this dissertation, we investigated the role of host genetic factors in various HCV outcomes, characterized their distribution in Aboriginal populations and developed a prognostic model with immune markers to predict 3-year risk of significant liver fibrosis in HIV-HCV co-infected Canadians.

In Manuscript 1, our goal was two-fold: 1) To test the association of HCV spontaneous clearance and three single nucleotide polymorphisms (SNPs) near the *Interferon Lambda 3 (IFNL3)* gene (rs12979860, rs8099917, functional variant rs8103142) and 2) to compare the SNP frequencies between HIV-HCV co-infected whites and Aboriginal peoples, as the latter are reported to have higher HCV spontaneous clearance rates[8,37] and HCV treatment responses.[44] We wanted to know if these SNPs, linked with higher clearance, were more frequent in Aboriginal peoples as IFNL allele frequencies vary by ethnicity, being almost universal among East Asians and rarer in those of African descent.[125,136] We analyzed the individual effect of each SNP using Cox proportional

hazards adjusted for sex and self-reported ethnicity and inheritance of the SNPs as a unit (haplotype), using haplotype analysis. Our results show that carriers of the favourable genotypes (rs12979860CC, rs8099917TT, and rs8103142TT) as well as the haplotype TCT (containing the favourable allele at each SNP) were over three times more likely to clear HCV spontaneously than those lacking the genotypes or haplotype. We also found that the allelic, genotypic and haplotypic frequencies differed between Aboriginal people and Canadian whites, with the beneficial variants being more common in the Aboriginal population. This could be a partial explanation for why Aboriginal people cleared more frequently than whites, but it did not fully account for the higher clearance rate.

Some studies reported that as the mechanism of IFNL SNPs is pro-inflammatory, the responder IFNL genotypes (that is, the genotype linked with spontaneous clearance and improved treatment response) leads to a more rapid progression to fibrosis and cirrhosis in those with chronic HCV.[1,11] This is biologically plausible as liver disease is caused by inflammation rather than HCV replication, and IFNL genotypes might be serving as markers for a strong immune response, which could, on the one hand, help clear HCV (as in Manuscript 1) but, on the other, cause liver damage if viral persistence and replication set in. In Manuscript 2, we thus examined the association of responder IFNL genotypes and haplotypes with significant liver fibrosis as measured by APRI $\geq$ 1.5 in HIV-HCV co-infected Canadians. As in Manuscript 1, we also used a Cox proportional hazards model with adjustment for other known risk factors for liver disease such as age, sex, ethnicity, alcohol use, CD4 count, HCV genotype, GGT and baseline APRI. We also did a haplotype analysis after adjusting for ethnicity. Our results show genotypes rs12979860CC, rs8099917 TT and rs81013142 TT were each individually linked to a higher rate

of significant liver fibrosis, even after accounting for other known risk factors. The association of rs8099917-TT with fibrosis was strongest, indicating an almost doubling of risk. As with the individual SNPs, the results from the haplotype analysis indicated that those with haplotype TCT had a higher risk of fibrosis than those lacking the haplotype, regardless of the mode of inheritance. Based on the dominant model, those with one or two copies of TCT had a 14% higher risk of fibrosis compared to those with no copies (HR 1.14, 95% CI: 0.73, 1.77).

Building on the results from Manuscript 2, we studied the prognostic ability of direct markers of the underlying inflammatory process of fibrogenesis for Manuscript 3. We wanted to know if measuring specific markers at baseline would improve ability to predict three-year risk of significant liver fibrosis beyond traditional clinical predictors alone. As these markers were not collected routinely as part of regular study visits, we nested a case cohort study in the eligible study population from the Canadian Coinfection Cohort. A randomly selected unstratified subcohort served as comparison group (APRI<1.5 when cases occurred), while cases developed an APRI≥1.5 (significant fibrosis) over study follow-up. Pro-fibrogenic markers (IL-8, MIP-1α, MIP-1β, MCP-1,TNF-α, RANTES, sICAM-1, sVCAM-1, CXCL9, CXCL11, TGF- β1, hsCRP, sCD14) were measured from first available plasma or serum in the subcohort and cases only. Results were analyzed using Cox proportional hazards with robust variance and Barlow weights. We compared the following prediction metrics between Model 1 (selected clinical predictors only) and Model 2 (clinical predictors from Model 1 plus selected markers): discrimination, calibration and net reclassification improvement. In our results, Model 1 included sex, current alcohol use, HIV viral load, baseline APRI, HCV genotype, and age as a restricted cubic spline with 3 knots. Model 1 was nested in Model 2, which also included IFNL rs8099917 genotype

and 5 immune markers: IL-8, sICAM-1, RANTES, hsCRP, and sCD14. Both the models were well-calibrated. The improvement in discrimination with model 2 was small, but the model with the markers fit and classified risk better. Further cost-benefit analyses is needed to assess whether this improvement justifies the additional cost of measuring these markers in the face of highly expensive HCV treatment.

7.2 STRENGTHS AND LIMITATIONS

The strengths of our study include data quality, generalizability of results, wide selection criteria, use of novel study design for developing a prognostic model and numerous sensitivity analyses to confirm the robustness of our findings.

The biggest strength was the quality of the data from the Canadian Coinfection Cohort, which provided regular longitudinal follow-up of a large representative sample of the Canadian HIV-HCV co-infected population under care. This allowed us to capture outcomes and time-varying covariates and also included marginalized groups such as PWID and Aboriginal peoples, making our results directly applicable to clinicians and researchers working with these populations. The findings from Manuscript 1 are especially relevant to co-infected Aboriginal peoples, as it is one of the only published reports on the frequency and possible effect of IFNL genotypes in Aboriginal participants. Possible implications of this are detailed in the next section "**7.3 Implications for Clinical Care and Research**."

Most of the individuals in the CCC had acquired HCV many years before study entry. By including those who acquired HCV remotely, we were able to avoid the "referral bias" of

previous retrospective studies. Those studies overestimated the severe outcomes of HCV infection (cirrhosis, hepatocellular carcinoma, and death) and did not allow for the examination of those who spontaneously recover from their infection or have milder forms of the disease.[268] Our study thus has characteristics of a prospective cohort but allows long-term follow-up that rarely can be achieved in prospective studies.

Another strength of the CCC data was the eligibility criteria, which did not exclude based on prior treatment history or severity of pre-existing liver disease, thus enabling us to avoid the selection bias that affected previous studies, as detailed in Manuscript 2. In those studies, mainly nested in clinical trials, recruitment depended on failing previous HCV treatment or having advanced liver disease, thus possibly enriching by the non-responder IFNL genotypes that was the independent variable in Manuscripts 1 and 2.[144]

In Manuscript 3, we used a case cohort study to design a prognostic model. This lesser-used study design was an economical, practical way to gather data that was not collected during regular follow-up. It also allowed us a more direct way to make absolute risk predictions by estimating the baseline survival function using the Barlow method, which mimics the proportions observed in the full cohort sample and so can be used for the estimation of absolute risk without any further rescaling.[271]

Several study limitations have been described in the Discussion sections of Manuscripts 1 to 3 in Chapters 4 to 6. This section highlights some of the main thesis limitations, which are often

common to research in PWID and assessment of liver disease, such as use of noninvasive APRI to measure liver fibrosis.

Left truncation was a problem with our data because most of the CCC participants were already decades into their HCV infection. Left truncation occurs when time at risk does not coincide with time in the study, and may reduce power or cause selection bias.[225] We addressed this problem using delayed entry as well as sensitivity analysis using inverse probability weights (IPW). More details are given in Chapters 3-5.

Interval censoring was also a problem because data from the CCC was collected at visits every six months. We addressed this issue using discrete time hazards methods such as pooled logistic regression with indicator variables for visits which are detailed in Chapters 3-5.

Another limitation was the uncertainty in the estimate of the date of HCV infection, which was approximate in most instances and used as the origin in Manuscripts 1 and 2. HCV duration was estimated based on date of HCV seroconversion, if known, or on the year of first injection drug use or blood product exposure. To address this uncertainty, a sensitivity analysis was conducted using age as the time axis with similar results. See Chapters 4-5 for more details.

For Manuscripts 1 and 2, we lacked data on the putative causal variant rs368234815 (*IFNL4*),[167] which was first reported in 2013. We instead examined the association of another potential causal SNP rs8103142, which marks physical changes in the IFNL3 protein, with HCV

outcomes. Implications of this are detailed in the Discussion sections of Manuscripts 1 and 2 in Chapters 4 and 5.

Another potential limitation was the lack of power to detect interaction between IFNL genotypes with sex, age, or HCV genotypes.[1,152] This is also detailed in the Discussion sections of Manuscripts 1 and 2 in Chapters 4 and 5.

A final limitation is the use of APRI score to measure the outcome (significant liver fibrosis) in Manuscripts 2 and 3. Liver biopsy, the gold standard for measuring liver fibrosis, is impractical for longitudinal research purposes. A cutoff of 1.5 has been validated in our study and others to detect significant liver fibrosis (METAVIR stage F2 or higher).[93] This cutoff has higher specificity than sensitivity for diagnosing significant liver fibrosis, but has been associated with clinically relevant endpoints in our study cohort,[94] as well as in others.[95,96] Further details are given in Chapter 2, section 2.9. The lack of biopsies also prevented us from investigating liver necroinflammation as the underlying mechanism driven by IFNL genotypes in Manuscript 2.

7.3 IMPLICATIONS FOR CLINICAL CARE AND RESEARCH

As the research aims in this dissertation all utilize data from the Canadian Coinfection Cohort, our results are directly generalizable to HIV-HCV co-infected Canadians under care and relevant to treating clinicians. Our findings provide insight into HCV disease etiology, whether it progresses to spontaneous HCV clearance (Manuscript 1) or liver fibrosis (Manuscript 2). This can inform future research on the pathogenesis of HCV, especially in the context of HIV-HCV co-infection and in marginalized groups such as Canadian Aboriginal peoples. Results from

Manuscript 3 are especially pertinent to clinical decision-making, and are especially timely in an era of newly approved expensive HCV treatments, when reimbursement by public and private payers has been restricted to those with advanced fibrosis (METAVIR stage F2 and higher).

Manuscript 1 was the only published report on the frequency of IFNL genotypes in the Canadian co-infected population, which has a unique genetic mix due to an overrepresentation of Aboriginal peoples. While these genetic markers are less important with the newer DAAs, they can still help identify individuals for shorter treatment durations or interferon-free regimens.[133-135] Our study was also one of the only published reports that documented that unlike Caucasians, Aboriginal people might be clearing HCV independent of an IFNL pathway. Differences in killer-cell immunoglobulin-like receptors (KIR) or IL-10 variants may explain higher HCV resolution among Aboriginal peoples.[38,39] Whether or how these factors interact with IFNL could reveal novel mechanisms of spontaneous clearance, which is of special interest to vaccine researchers. Finally, as Aboriginal people are disproportionately affected by both HIV and HCV and experience significant challenges accessing treatment, our findings could be relevant to researchers and clinicians working in those communities.

Findings from Manuscript 2 could be especially useful for making clinical decisions in a population with many hepatotoxic behaviors, such as injecting drug use. Our results support findings from another study,[1] which reported higher risk of fibrosis associated with specific IFNL genotypes, regardless of cause of liver disease. Taken together, these results would suggest that HCV treatment be prioritized for those co-infected individuals carrying the IFNL genotypes of interest before liver disease reaches a point of irreversible damage, even after viral cure. The

results could also suggest closer post-treatment follow-up of co-infected individuals with the at-risk alleles. This is also relevant because liver fibrosis is a precursor for advanced liver disease, which could affect HCV treatment efficacy. Finally, functional studies that investigate the role of IFNL in inflammation or fibrogenesis could reveal potential intervention points for designing effective anti-fibrotic therapies.

The impact of Manuscript 3 is almost entirely in clinical decision-making with fewer applications to etiological research. Our goal was to design a prognostic model with data that healthcare providers would routinely have access to. As such, our findings are most applicable to co-infected patients under care, especially those whose risk profiles are most like the participants in the Canadian Coinfection Cohort—individuals chronically infected with HCV with a history of injection drug use, which is representative of many co-infected people in Western countries. Our choice to use a case cohort study contributes to the growing body of research on novel genetic, metabolic, and immune markers that support the use of this design as an economical way to assess the clinical potential of markers not routinely collected during follow-up. Before clinical application, our model would benefit from a formal cost-benefit analysis of the markers and external validation.

7.4 CONCLUSION

Our results reveal the role of IFNL genetic markers on both spontaneous clearance and significant liver fibrosis in the distinct genetic mix of the Canadian co-infected population. Carrying the major IFNL alleles has paradoxical effects-- more frequent spontaneous clearance and if that fails, higher risk of liver fibrosis. These variants are also common among Aboriginal

peoples, who naturally clear HCV more often, but appear to have other immunological mechanisms that are likely driving their HCV response. We also developed a prognostic model with immune markers which could be directly applicable to treatment decisions involving co-infected patients.

In conclusion, even as HCV cure becomes more common with effective DAAs, liver disease will remain a strong concern in HIV-HCV co-infected individuals, where hepatotoxic exposures are common and liver fibrosis is accelerated. Furthermore, as cost of DAAs remain high and uptake remains low, there is a need to better understand the etiology, progression, and clinical management of HCV in co-infected patients, who face many other competing challenges. Findings from this dissertation will therefore be relevant to the care and research involving this population.

REFERENCES

1. Eslam M, Hashem AM, Leung R, et al. Interferon-lambda rs12979860 genotype and liver fibrosis in viral and non-viral chronic liver disease. *Nature communications.* 2015;6:6422.
2. Platt L, Easterbrook P, Gower E, et al. Prevalence and burden of HCV co-infection in people living with HIV: a global systematic review and meta-analysis. *The Lancet infectious diseases.*
3. Roe B, Hall WW. Cellular and molecular interactions in coinfection with hepatitis C virus and human immunodeficiency virus. *Expert reviews in molecular medicine.* 2008;10:e30.
4. Sulkowski MS. Hepatitis C virus-human immunodeficiency virus coinfection. *Liver international : official journal of the International Association for the Study of the Liver.* 2012;32 Suppl 1:129-134.
5. Pellicoro A, Ramachandran P, Iredale JP, Fallowfield JA. Liver fibrosis and repair: immune regulation of wound healing in a solid organ. *Nat Rev Immunol.* 2014;14(3):181-194.
6. Marshall A, Saeed S, Barrett L, Cooper CL, Treloar C, Bruneau J, Feld JJ, Gallagher L, Klein MB, Krajden M, Shoukry NH, Taylor LE, and Grebely J on behalf of the Canadian Network on Hepatitis C (CanHepC). A review of restrictions for reimbursement of direct-acting antiviral treatment for hepatitis C virus infection in Canada. *CMAJ Open.* 2016.
7. Grady BP, Schinkel J, Thomas XV, Dalgard O. Hepatitis C virus reinfection following treatment among people who use drugs. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America.* 2013;57 Suppl 2:S105-110.
8. Grebely J, Raffa JD, Lai C, Krajden M, Conway B, Tyndall MW. Factors associated with spontaneous clearance of hepatitis C virus among illicit drug users. *Canadian journal of gastroenterology = Journal canadien de gastroenterologie.* 2007;21(7):447-451.
9. Larcombe L, Rempel JD, Dembinski I, Tinckam K, Rigatto C, Nickerson P. Differential cytokine genotype frequencies among Canadian Aboriginal and Caucasian populations. *Genes and immunity.* 2005;6(2):140-144.
10. Rempel JD, Aborsangaya KB, Alphonse MP, Minuk GY. The influence of North American Aboriginal ethnicity on pro-inflammatory and anti-inflammatory cytokine responses to IFN-alpha. *Journal of viral hepatitis.* 2009;16(4):292-297.
11. Guzman-Fulgencio M, Berenguer J, Garcia-Alvarez M, et al. IL28B polymorphisms are associated with severity of liver disease in human immunodeficiency virus (HIV) patients coinfected with hepatitis C virus. *The Journal of infection.* 2013;66(2):170-178.
12. Bochud P, Bibert S, Kutalik Z, et al. IL28B polymorphisms linked to poor response to treatment are associated with low necroinflammatory activity and slow fibrosis progression in HCV genotype non1-infected patients. *Journal of hepatology.* 2011;54(54):S539.
13. Greenland P, O'Malley PG. When is a new prediction marker useful? A consideration of lipoprotein-associated phospholipase A2 and C-reactive protein for stroke risk. *Archives of internal medicine.* 2005;165(21):2454-2456.
14. UNAIDS. *Fact Sheet: 2014 Global Statistics.* 2014.

15. Maartens G, Celum C, Lewin SR. HIV infection: epidemiology, pathogenesis, treatment, and prevention. *Lancet.* 2014;384(9939):258-271.
16. PHAC. *Summary: Estimates of HIV Prevalence and Incidence in Canada, 2011.* 2012.
17. PHAC. *PHAC Population-Specific HIV/AIDS Status Report: Aboriginal Peoples* Public Health Agency of Canada;2012.
18. PHAC. *Public Health Agency of Canada. HIV and AIDS in Canada: Surveillance Report to December 31, 2013.* 2014.
19. (WHO) WHO. *Hepatitis C Fact Sheet.* 2015.
20. Xia G-L, Liu C-B, Cao H-L, et al. Prevalence of hepatitis B and C virus infections in the general Chinese population. Results from a nationwide cross-sectional seroepidemiologic study of hepatitis A, B, C, D, and E virus infections in China, 1992. *International Hepatology Communications.* 1996;5(1):62-73.
21. Qin Q, Smith MK, Wang L, et al. Hepatitis C virus infection in China: an emerging public health issue. *Journal of viral hepatitis.* 2015;22(3):238-244.
22. Frank C, Mohamed MK, Strickland GT, et al. The role of parenteral antischistosomal therapy in the spread of hepatitis C virus in Egypt. *Lancet.* 2000;355(9207):887-891.
23. Alter MJ. Epidemiology of viral hepatitis and HIV co-infection. *Journal of hepatology.* 2006;44(1 Suppl):S6-9.
24. Gower E, Estes C, Blach S, Razavi-Shearer K, Razavi H. Global epidemiology and genotype distribution of the hepatitis C virus infection. *Journal of hepatology.* 2014;61(1 Suppl):S45-57.
25. Alter MJ. Prevention of spread of hepatitis C. *Hepatology.* 2002;36(5 Suppl 1):S93-98.
26. PHAC. *Epidemiology of acute hepatitis C infection in Canada: results from the Enhanced Hepatitis Strain Surveillance System (EHSSS)* Ottawa: Public Health Agency of Canada;2009.
27. PHAC. *Hepatitis C in Canada: 2005-2010 Surveillance Report.* 2012.
28. Saskatchewan MoH. *Saskatchewan's HIV Strategy 2010 – 2014.* 2010.
29. Myers RP, Krajden M, Bilodeau M, et al. Burden of disease and cost of chronic hepatitis C infection in Canada. *Canadian journal of gastroenterology & hepatology.* 2014;28(5):243-250.
30. Remis R. *Modelling The Incidence And Prevalence Of Hepatitis C Infection And Its Sequelae In Canada*
. Public Health Agency of Canada 2007.
31. Wasley A, Alter MJ. Epidemiology of hepatitis C: geographic differences and temporal trends. *Seminars in liver disease.* 2000;20(1):1-16.
32. Hernandez MD, Sherman KE. HIV/hepatitis C coinfection natural history and disease progression. *Current opinion in HIV and AIDS.* 2011;6(6):478-482.
33. Statistics C. *Aboriginal peoples in Canada in 2006: Inuit, Métis and First Nations, 2006–findings.* Ottawa2008.
34. Wu HX, Wu J, Wong T, et al. Incidence and risk factors for newly acquired hepatitis C virus infection among Aboriginal versus non-Aboriginal Canadians in six regions, 1999-2004. *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology.* 2007;26(3):167-174.
35. Rempel JD, Uhanova J. Hepatitis C virus in American Indian/Alaskan Native and Aboriginal peoples of North America. *Viruses.* 2012;4(12):3912-3931.

36. Lelutiu-Weinberger C, Pouget ER, Des Jarlais DD, et al. A meta-analysis of the hepatitis C virus distribution in diverse racial/ethnic drug injector groups. *Social science & medicine.* 2009;68(3):579-590.
37. Minuk GY, Uhanova J. Viral hepatitis in the Canadian Inuit and First Nations populations. *Canadian journal of gastroenterology = Journal canadien de gastroenterologie.* 2003;17(12):707-712.
38. Aborsangaya KB, Dembinski I, Khatkar S, Alphonse MP, Nickerson P, Rempel JD. Impact of aboriginal ethnicity on HCV core-induced IL-10 synthesis: interaction with IL-10 gene polymorphisms. *Hepatology.* 2007;45(3):623-630.
39. Rempel JD, Hawkins K, Lande E, Nickerson P. The potential influence of KIR cluster profiles on disease patterns of Canadian Aboriginals and other indigenous peoples of the Americas. *European journal of human genetics : EJHG.* 2011;19(12):1276-1280.
40. Bowker SL, Soskolne CL, Houston SC, Newman SC, Jhangri GS. Human immunodeficiency virus (HIV) and hepatitis C virus (HCV) in a Northern Alberta population. *Canadian journal of public health = Revue canadienne de sante publique.* 2004;95(3):188-192.
41. Rourke SB, Sobota M, Tucker R, et al. Social determinants of health associated with hepatitis C co-infection among people living with HIV: results from the Positive Spaces, Healthy Places study. *Open Med.* 2011;5(3):e120-131.
42. Braitstein P, Montessori V, Chan K, et al. Quality of life, depression and fatigue among persons co-infected with HIV and hepatitis C: outcomes from a population-based cohort. *AIDS Care.* 2005;17(4):505-515.
43. Klein MB, Saeed S, Yang H, et al. Cohort profile: the Canadian HIV-hepatitis C co-infection cohort study. *International journal of epidemiology.* 2010;39(5):1162-1169.
44. Cooper CL, Bailey RJ, Bain VG, et al. Outcomes of peginterferon alpha-2a and ribavirin hepatitis C therapy in Aboriginal Canadians. *Canadian journal of gastroenterology = Journal canadien de gastroenterologie.* 2008;22(8):677-680.
45. Antiretroviral Therapy Cohort C. Causes of death in HIV-1-infected patients treated with antiretroviral therapy, 1996-2006: collaborative analysis of 13 HIV cohort studies. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America.* 2010;50(10):1387-1396.
46. Smith CJ, Ryom L, Weber R, et al. Trends in underlying causes of death in people with HIV from 1999 to 2011 (D:A:D): a multicohort collaboration. *Lancet.* 2014;384(9939):241-248.
47. Hua L, Andersen JW, Daar ES, Glesby MJ, Hollabaugh K, Tierney C. Hepatitis C virus/HIV coinfection and responses to initial antiretroviral treatment. *AIDS.* 2013;27(17):2725-2734.
48. Operskalski EA, Kovacs A. HIV/HCV co-infection: pathogenesis, clinical complications, treatment, and new therapeutic technologies. *Current HIV/AIDS reports.* 2011;8(1):12-22.
49. Hajarizadeh B, Grebely J, Dore GJ. Epidemiology and natural history of HCV infection. *Nature reviews. Gastroenterology & hepatology.* 2013;10(9):553-562.
50. Alberti A, Chemello L, Benvegnu L. Natural history of hepatitis C. *Journal of hepatology.* 1999;31 Suppl 1:17-24.

51. Micallef JM, Kaldor JM, Dore GJ. Spontaneous viral clearance following acute hepatitis C infection: a systematic review of longitudinal studies. *Journal of viral hepatitis.* 2006;13(1):34-41.
52. Grebely J, Page K, Sacks-Davis R, et al. The effects of female sex, viral genotype, and IL28B genotype on spontaneous clearance of acute hepatitis C virus infection. *Hepatology.* 2014;59(1):109-120.
53. Bataller R, Brenner DA. Liver fibrosis. *The Journal of clinical investigation.* 2005;115(2):209-218.
54. Romero-Gomez M, Eslam M, Ruiz A, Maraver M. Genes and hepatitis C: susceptibility, fibrosis progression and response to treatment. *Liver international : official journal of the International Association for the Study of the Liver.* 2011;31(4):443-460.
55. Ghany MG, Kleiner DE, Alter H, et al. Progression of fibrosis in chronic hepatitis C. *Gastroenterology.* 2003;124(1):97-104.
56. Teixeira R, Marcos LA, Friedman SL. Immunopathogenesis of hepatitis C virus infection and hepatic fibrosis: New insights into antifibrotic therapy in chronic hepatitis C. *Hepatology research : the official journal of the Japan Society of Hepatology.* 2007;37(8):579-595.
57. Arthur MJ. Fibrogenesis II. Metalloproteinases and their inhibitors in liver fibrosis. *American journal of physiology. Gastrointestinal and liver physiology.* 2000;279(2):G245-249.
58. Bataller R, Paik YH, Lindquist JN, Lemasters JJ, Brenner DA. Hepatitis C virus core and nonstructural proteins induce fibrogenic effects in hepatic stellate cells. *Gastroenterology.* 2004;126(2):529-540.
59. Marra F. Hepatic stellate cells and the regulation of liver inflammation. *Journal of hepatology.* 1999;31(6):1120-1130.
60. Gabele E, Brenner DA, Rippe RA. Liver fibrosis: signals leading to the amplification of the fibrogenic hepatic stellate cell. *Frontiers in bioscience : a journal and virtual library.* 2003;8:d69-77.
61. Ellis EL, Mann DA. Clinical evidence for the regression of liver fibrosis. *Journal of hepatology.* 2012;56(5):1171-1180.
62. D'Ambrosio R, Aghemo A, Rumi MG, et al. A morphometric and immunohistochemical study to assess the benefit of a sustained virological response in hepatitis C virus patients with cirrhosis. *Hepatology.* 2012;56(2):532-543.
63. Ramachandran P, Iredale JP, Fallowfield JA. Resolution of liver fibrosis: basic mechanisms and clinical relevance. *Seminars in liver disease.* 2015;35(2):119-131.
64. Resino S, Sanchez-Conde M, Berenguer J. Coinfection by human immunodeficiency virus and hepatitis C virus: noninvasive assessment and staging of fibrosis. *Current opinion in infectious diseases.* 2012;25(5):564-569.
65. Pinzani M. Liver Fibrosis in the Post-HCV Era. *Seminars in liver disease.* 2015;35(2):157-165.
66. Lee YA, Wallace MC, Friedman SL. Pathobiology of liver fibrosis: a translational success story. *Gut.* 2015;64(5):830-841.
67. Soto B, Sanchez-Quijano A, Rodrigo L, et al. Human immunodeficiency virus infection modifies the natural history of chronic parenterally-acquired hepatitis C with an unusually rapid progression to cirrhosis. *Journal of hepatology.* 1997;26(1):1-5.

68. Benhamou Y, Bochet M, Di Martino V, et al. Liver fibrosis progression in human immunodeficiency virus and hepatitis C virus coinfected patients. The Multivirc Group. *Hepatology.* 1999;30(4):1054-1058.
69. Thein HH, Yi Q, Dore GJ, Krahn MD. Natural history of hepatitis C virus infection in HIV-infected individuals and the impact of HIV in the era of highly active antiretroviral therapy: a meta-analysis. *AIDS.* 2008;22(15):1979-1991.
70. Kirk GD, Mehta SH, Astemborski J, et al. HIV, age, and the severity of hepatitis C virus-related liver disease: a cohort study. *Annals of internal medicine.* 2013;158(9):658-666.
71. Salmon-Ceron D, Lewden C, Morlat P, et al. Liver disease as a major cause of death among HIV infected patients: role of hepatitis C and B viruses and alcohol. *Journal of hepatology.* 2005;42(6):799-805.
72. Sulkowski MS, Mehta SH, Torbenson MS, et al. Rapid fibrosis progression among HIV/hepatitis C virus-co-infected adults. *AIDS.* 2007;21(16):2209-2216.
73. Macias J, Berenguer J, Japon MA, et al. Fast fibrosis progression between repeated liver biopsies in patients coinfected with human immunodeficiency virus/hepatitis C virus. *Hepatology.* 2009;50(4):1056-1063.
74. Graham CS, Baden LR, Yu E, et al. Influence of human immunodeficiency virus infection on the course of hepatitis C virus infection: a meta-analysis. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America.* 2001;33(4):562-569.
75. Benhamou Y, Di Martino V, Bochet M, et al. Factors affecting liver fibrosis in human immunodeficiency virus-and hepatitis C virus-coinfected patients: impact of protease inhibitor therapy. *Hepatology.* 2001;34(2):283-287.
76. Villano SA, Vlahov D, Nelson KE, Cohn S, Thomas DL. Persistence of viremia and the importance of long-term follow-up after acute hepatitis C infection. *Hepatology.* 1999;29(3):908-914.
77. Shek FW, Benyon RC. How can transforming growth factor beta be targeted usefully to combat liver fibrosis? *European journal of gastroenterology & hepatology.* 2004;16(2):123-126.
78. Kim AY, Chung RT. Coinfection with HIV-1 and HCV--a one-two punch. *Gastroenterology.* 2009;137(3):795-814.
79. Munshi N, Balasubramanian A, Koziel M, Ganju RK, Groopman JE. Hepatitis C and human immunodeficiency virus envelope proteins cooperatively induce hepatocytic apoptosis via an innocent bystander mechanism. *The Journal of infectious diseases.* 2003;188(8):1192-1204.
80. Malta M, Magnanini MM, Strathdee SA, Bastos FI. Adherence to antiretroviral therapy among HIV-infected drug users: a meta-analysis. *AIDS and behavior.* 2010;14(4):731-747.
81. Tedaldi E, Peters L, Neuhaus J, et al. Opportunistic disease and mortality in patients coinfected with hepatitis B or C virus in the strategic management of antiretroviral therapy (SMART) study. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America.* 2008;47(11):1468-1475.
82. El-Sadr WM, Lundgren JD, Neaton JD, et al. CD4+ count-guided interruption of antiretroviral treatment. *The New England journal of medicine.* 2006;355(22):2283-2296.

83. Calmy A, Gayet-Ageron A, Montecucco F, et al. HIV increases markers of cardiovascular risk: results from a randomized, treatment interruption trial. *AIDS.* 2009;23(8):929-939.
84. Marin B, Thiebaut R, Bucher HC, et al. Non-AIDS-defining deaths and immunodeficiency in the era of combination antiretroviral therapy. *AIDS.* 2009;23(13):1743-1753.
85. Thorpe J, Saeed S, Moodie EE, Klein MB, Canadian Co-infection Cohort S. Antiretroviral treatment interruption leads to progression of liver fibrosis in HIV-hepatitis C virus co-infection. *AIDS.* 2011;25(7):967-975.
86. Peters L, Neuhaus J, Duprez D, et al. Biomarkers of Inflammation and Coagulation and Risk of Non-AIDS Death (NAD) in HIV/Hepatitis Co-infected Patients in the SMART Study. Paper presented at: 17th Conference on Retroviruses & Opportunistic Infections (CROI 2010); February 16-19, 2010, 2010; San Francisco, California, USA.
87. Macias J, Castellano V, Merchante N, et al. Effect of antiretroviral drugs on liver fibrosis in HIV-infected patients with chronic hepatitis C: harmful impact of nevirapine. *AIDS.* 2004;18(5):767-774.
88. Pineda JA, Macias J. Progression of liver fibrosis in patients coinfected with hepatitis C virus and human immunodeficiency virus undergoing antiretroviral therapy. *The Journal of antimicrobial chemotherapy.* 2005;55(4):417-419.
89. Afdhal NH, Nunes D. Evaluation of liver fibrosis: a concise review. *The American journal of gastroenterology.* 2004;99(6):1160-1174.
90. Everhart JE, Wright EC, Goodman ZD, et al. Prognostic Value of Ishak Fibrosis Stage: Findings from the HALT-C Trial. *Hepatology (Baltimore, Md.).* 2010;51(2):10.1002/hep.23315.
91. Kelleher TB, Mehta SH, Bhaskar R, et al. Prediction of hepatic fibrosis in HIV/HCV co-infected patients using serum fibrosis markers: the SHASTA index. *Journal of hepatology.* 2005;43(1):78-84.
92. Lin ZH, Xin YN, Dong QJ, et al. Performance of the aspartate aminotransferase-to-platelet ratio index for the staging of hepatitis C-related fibrosis: an updated meta-analysis. *Hepatology.* 2011;53(3):726-736.
93. Al-Mohri H, Cooper C, Murphy T, Klein MB. Validation of a simple model for predicting liver fibrosis in HIV/hepatitis C virus-coinfected patients. *HIV medicine.* 2005;6(6):375-378.
94. Al-Mohri H, Murphy T, Lu Y, Lalonde RG, Klein MB. Evaluating liver fibrosis progression and the impact of antiretroviral therapy in HIV and hepatitis C coinfection using a noninvasive marker. *J Acquir Immune Defic Syndr.* 2007;44(4):463-469.
95. Macias J, Giron-Gonzalez JA, Gonzalez-Serrano M, et al. Prediction of liver fibrosis in human immunodeficiency virus/hepatitis C virus coinfected patients by simple non-invasive indexes. *Gut.* 2006;55(3):409-414.
96. Macias J, Mira J, Gilabert I, et al. Combined use of aspartate aminotransferase, platelet count and matrix metalloproteinase 2 measurements to predict liver fibrosis in HIV/hepatitis C virus-coinfected patients. *HIV medicine.* 2011;12(1):14-21.

97. Sterling RK, Lissen E, Clumeck N, et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology.* 2006;43(6):1317-1325.
98. Vallet-Pichard A, Mallet V, Nalpas B, et al. FIB-4: an inexpensive and accurate marker of fibrosis in HCV infection. comparison with liver biopsy and fibrotest. *Hepatology.* 2007;46(1):32-36.
99. Jiang D, Liang J, Noble PW. Hyaluronan as an immune regulator in human diseases. *Physiol Rev.* 2011;91(1):221-264.
100. Patel K, Lajoie A, Heaton S, et al. Clinical use of hyaluronic acid as a predictor of fibrosis change in hepatitis C. *Journal of gastroenterology and hepatology.* 2003;18(3):253-257.
101. Castera L, Forns X, Alberti A. Non-invasive evaluation of liver fibrosis using transient elastography. *Journal of hepatology.* 2008;48(5):835-847.
102. Foucher J, Chanteloup E, Vergniol J, et al. Diagnosis of cirrhosis by transient elastography (FibroScan): a prospective study. *Gut.* 2006;55(3):403-408.
103. Chazouilleres O, Wright TL. Hepatitis C and liver transplantation. *Journal of gastroenterology and hepatology.* 1995;10(4):471-480.
104. Charlton M. Natural history of hepatitis C and outcomes following liver transplantation. *Clinics in liver disease.* 2003;7(3):585-602.
105. Grebely J, Raffa JD, Lai C, et al. Low uptake of treatment for hepatitis C virus infection in a large community-based study of inner city residents. *Journal of viral hepatitis.* 2009;16(5):352-358.
106. Simmons B, Saleem J, Heath K, Cooke GS, Hill A. Long-Term Treatment Outcomes of Patients Infected With Hepatitis C Virus: A Systematic Review and Meta-analysis of the Survival Benefit of Achieving a Sustained Virological Response. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America.* 2015;61(5):730-740.
107. Marcellin P, Boyer N, Gervais A, et al. Long-term histologic improvement and loss of detectable intrahepatic HCV RNA in patients with chronic hepatitis C and sustained response to interferon-alpha therapy. *Annals of internal medicine.* 1997;127(10):875-881.
108. Poynard T, McHutchison J, Manns M, et al. Impact of pegylated interferon alfa-2b and ribavirin on liver fibrosis in patients with chronic hepatitis C. *Gastroenterology.* 2002;122(5):1303-1313.
109. Morgan RL, Baack B, Smith BD, Yartel A, Pitasi M, Falck-Ytter Y. Eradication of hepatitis C virus infection and the development of hepatocellular carcinoma: a meta-analysis of observational studies. *Annals of internal medicine.* 2013;158(5 Pt 1):329-337.
110. van der Meer AJ, Veldt BJ, Feld JJ, et al. Association between sustained virological response and all-cause mortality among patients with chronic hepatitis C and advanced hepatic fibrosis. *JAMA : the journal of the American Medical Association.* 2012;308(24):2584-2593.
111. Fabrizi F, Dixit V, Messa P. Antiviral therapy of symptomatic HCV-associated mixed cryoglobulinemia: meta-analysis of clinical studies. *Journal of medical virology.* 2013;85(6):1019-1027.

112. Martin NK, Vickerman P, Grebely J, et al. Hepatitis C virus treatment for prevention among people who inject drugs: Modeling treatment scale-up in the age of direct-acting antivirals. *Hepatology.* 2013;58(5):1598-1609.
113. Fried MW, Shiffman ML, Reddy KR, et al. Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. *The New England journal of medicine.* 2002;347(13):975-982.
114. Pearlman BL, Traub N. Sustained Virologic Response to Antiviral Therapy for Chronic Hepatitis C Virus Infection: A Cure and So Much More. *Clinical Infectious Diseases.* 2011;52(7):889-900.
115. Sherman M, Shafran S, Burak K, et al. Management of chronic hepatitis C: consensus guidelines. *Canadian journal of gastroenterology = Journal canadien de gastroenterologie.* 2007;21 Suppl C:25C-34C.
116. Ghany MG, Strader DB, Thomas DL, Seeff LB. Diagnosis, management, and treatment of hepatitis C: an update. *Hepatology.* 2009;49(4):1335-1374.
117. Hadziyannis SJ, Sette H, Jr., Morgan TR, et al. Peginterferon-alpha2a and ribavirin combination therapy in chronic hepatitis C: a randomized study of treatment duration and ribavirin dose. *Annals of internal medicine.* 2004;140(5):346-355.
118. Shafran SD. HIV Coinfected Have Similar SVR Rates as HCV Monoinfected With DAAs: It's Time to End Segregation and Integrate HIV Patients Into HCV Trials. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America.* 2015.
119. Hull M, Shafran S, Tseng A, Giguere P, Klein MB, Cooper C. CIHR Canadian HIV Trials Network Co-Infection and Concurrent Diseases Core: Updated Canadian guidelines for the treatment of hepatitis C infection in HIV-hepatitis C coinfected adults. *The Canadian journal of infectious diseases & medical microbiology = Journal canadien des maladies infectieuses et de la microbiologie medicale / AMMI Canada.* 2014;25(6):311-320.
120. Peltekian KM. With all the new treatment regimens, complete elimination of hepatitis C virus in Canada is a possibility! But when will Canadians have access these drugs? *Canadian journal of gastroenterology & hepatology.* 2014;28(8):417-418.
121. Duggal P, Thio CL, Wojcik GL, et al. Genome-wide association study of spontaneous resolution of hepatitis C virus infection: data from multiple cohorts. *Annals of internal medicine.* 2013;158(4):235-245.
122. Alric L, Fort M, Izopet J, et al. Genes of the major histocompatibility complex class II influence the outcome of hepatitis C virus infection. *Gastroenterology.* 1997;113(5):1675-1681.
123. Balagopal A, Thomas DL, Thio CL. IL28B and the control of hepatitis C virus infection. *Gastroenterology.* 2010;139(6):1865-1876.
124. Ge D, Fellay J, Thompson AJ, et al. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature.* 2009;461(7262):399-401.
125. Tanaka Y, Nishida N, Sugiyama M, et al. Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. *Nature genetics.* 2009;41(10):1105-1109.

126. Rauch A, Kutalik Z, Descombes P, et al. Genetic variation in IL28B is associated with chronic hepatitis C and treatment failure: a genome-wide association study. *Gastroenterology.* 2010;138(4):1338-1345, 1345 e1331-1337.
127. Suppiah V, Moldovan M, Ahlenstiel G, et al. IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. *Nature genetics.* 2009;41(10):1100-1104.
128. Aparicio E, Parera M, Franco S, et al. IL28B SNP rs8099917 is strongly associated with pegylated interferon-alpha and ribavirin therapy treatment failure in HCV/HIV-1 coinfected patients. *PloS one.* 2010;5(10):e13771.
129. de Araujo ES, Dahari H, Cotler SJ, et al. Pharmacodynamics of PEG-IFN-[alpha]-2a and HCV response as a function of IL28B polymorphism in HIV/HCV-coinfected patients. *J Acquir Immune Defic Syndr.* 2011;56(2):95-99.
130. Rallon NI, Naggie S, Benito JM, et al. Association of a single nucleotide polymorphism near the interleukin-28B gene with response to hepatitis C therapy in HIV/hepatitis C virus-coinfected patients. *AIDS.* 2010;24(8):F23-29.
131. Rallon NI, Soriano V, Naggie S, et al. IL28B gene polymorphisms and viral kinetics in HIV/hepatitis C virus-coinfected patients treated with pegylated interferon and ribavirin. *AIDS.* 2011;25(8):1025-1033.
132. Bellanti F, Vendemiale G, Altomare E, Serviddio G. The impact of interferon lambda 3 gene polymorphism on natural course and treatment of hepatitis C. *Clinical & developmental immunology.* 2012;2012:849373.
133. Poordad F, Bronowicki JP, Gordon SC, et al. Factors that predict response of patients with hepatitis C virus infection to boceprevir. *Gastroenterology.* 2012;143(3):608-618 e601-605.
134. Akuta N, Suzuki F, Hirakawa M, et al. Amino acid substitution in hepatitis C virus core region and genetic variation near the interleukin 28B gene predict viral response to telaprevir with peginterferon and ribavirin. *Hepatology.* 2010;52(2):421-429.
135. Chu TW, Kulkarni R, Gane EJ, et al. Effect of IL28B genotype on early viral kinetics during interferon-free treatment of patients with chronic hepatitis C. *Gastroenterology.* 2012;142(4):790-795.
136. Thomas DL, Thio CL, Martin MP, et al. Genetic variation in IL28B and spontaneous clearance of hepatitis C virus. *Nature.* 2009;461(7265):798-801.
137. Grebely J, Petoumenos K, Hellard M, et al. Potential role for interleukin-28B genotype in treatment decision-making in recent hepatitis C virus infection. *Hepatology.* 2010;52(4):1216-1224.
138. Montes-Cano MA, Garcia-Lozano JR, Abad-Molina C, et al. Interleukin-28B genetic variants and hepatitis virus infection by different viral genotypes. *Hepatology.* 2010;52(1):33-37.
139. Mizukoshi E, Eisenbach C, Edlin BR, et al. Hepatitis C virus (HCV)-specific immune responses of long-term injection drug users frequently exposed to HCV. *The Journal of infectious diseases.* 2008;198(2):203-212.
140. Tillmann HL, Thompson AJ, Patel K, et al. A polymorphism near IL28B is associated with spontaneous clearance of acute hepatitis C virus and jaundice. *Gastroenterology.* 2010;139(5):1586-1592, 1592 e1581.

141. Ruiz-Extremera A, Munoz-Gamez JA, Salmeron-Ruiz MA, et al. Genetic variation in interleukin 28B with respect to vertical transmission of hepatitis C virus and spontaneous clearance in HCV-infected children. *Hepatology.* 2011;53(6):1830-1838.
142. Clark PJ, Thompson AJ. Host genomics and HCV treatment response. *Journal of gastroenterology and hepatology.* 2011.
143. Marabita F, Aghemo A, De Nicola S, et al. Genetic variation in the interleukin-28B gene is not associated with fibrosis progression in patients with chronic hepatitis C and known date of infection. *Hepatology.* 2011;54(4):1127-1134.
144. Noureddin M, Wright EC, Alter HJ, et al. Association of IL28B genotype with fibrosis progression and clinical outcomes in patients with chronic hepatitis C: a longitudinal analysis. *Hepatology.* 2013;58(5):1548-1557.
145. Bruno S, Thompson AJ, Critelli R, et al. Interferon lambda-3 is not associated with clinical outcome in patients with HCV-induced compensated cirrhosis: a long-term cohort study. *Antiviral research.* 2015;113:27-32.
146. Agundez JA, Garcia-Martin E, Maestro ML, et al. Relation of IL28B gene polymorphism with biochemical and histological features in hepatitis C virus-induced liver disease. *PloS one.* 2012;7(5):e37998.
147. Abe H, Ochi H, Maekawa T, et al. Common variation of IL28 affects gamma-GTP levels and inflammation of the liver in chronically infected hepatitis C virus patients. *Journal of hepatology.* 2010;53(3):439-443.
148. Rembeck K, Alsio A, Christensen PB, et al. Impact of IL28B-related single nucleotide polymorphisms on liver histopathology in chronic hepatitis C genotype 2 and 3. *PloS one.* 2012;7(1):e29370.
149. Barreiro P, Pineda JA, Rallon N, et al. Influence of interleukin-28B single-nucleotide polymorphisms on progression to liver cirrhosis in human immunodeficiency virus-hepatitis C virus-coinfected patients receiving antiretroviral therapy. *The Journal of infectious diseases.* 2011;203(11):1629-1636.
150. Ydreborg M, Westin J, Rembeck K, et al. Impact of Il28b-related single nucleotide polymorphisms on liver transient elastography in chronic hepatitis C infection. *PloS one.* 2013;8(11):e80172.
151. Layden JE, Tayo BO, Cotler SJ, et al. Association of genetic variants with rapid fibrosis: progression after liver transplantation for hepatitis C. *Transplantation.* 2014;97(10):1072-1078.
152. Bochud PY, Bibert S, Kutalik Z, et al. IL28B alleles associated with poor hepatitis C virus (HCV) clearance protect against inflammation and fibrosis in patients infected with non-1 HCV genotypes. *Hepatology.* 2012;55(2):384-394.
153. Fabris C, Falleti E, Cussigh A, et al. The interleukin 28B rs12979860 C/T polymorphism and serum cholesterol as predictors of fibrosis progression in patients with chronic hepatitis C and persistently normal transaminases. *Journal of medical virology.* 2012;84(5):747-755.
154. Di Marco V, Bronte F, Calvaruso V, et al. IL28B polymorphisms influence stage of fibrosis and spontaneous or interferon-induced viral clearance in thalassemia patients with hepatitis C virus infection. *Haematologica.* 2012;97(5):679-686.

155. Falleti E, Bitetto D, Fabris C, et al. Role of interleukin 28B rs12979860 C/T polymorphism on the histological outcome of chronic hepatitis C: relationship with gender and viral genotype. *Journal of clinical immunology.* 2011;31(5):891-899.
156. Fabris C, Falleti E, Cussigh A, et al. IL-28B rs12979860 C/T allele distribution in patients with liver cirrhosis: role in the course of chronic viral hepatitis and the development of HCC. *Journal of hepatology.* 2011;54(4):716-722.
157. Tillmann HL, Patel K, Muir AJ, et al. Beneficial IL28B genotype associated with lower frequency of hepatic steatosis in patients with chronic hepatitis C. *Journal of hepatology.* 2011;55(6):1195-1200.
158. Chen JY, Feeney ER, Chung RT. HCV and HIV co-infection: mechanisms and management. *Nature reviews. Gastroenterology & hepatology.* 2014;11(6):362-371.
159. Sheahan T, Imanaka N, Marukian S, et al. Interferon Lambda Alleles Predict Innate Antiviral Immune Responses and Hepatitis C Virus Permissiveness. *Cell Host & Microbe.* 2014;15(2):190-202.
160. Ank N, West H, Bartholdy C, Eriksson K, Thomsen AR, Paludan SR. Lambda interferon (IFN-lambda), a type III IFN, is induced by viruses and IFNs and displays potent antiviral activity against select virus infections in vivo. *Journal of virology.* 2006;80(9):4501-4509.
161. Brand S, Beigel F, Olszak T, et al. IL-28A and IL-29 mediate antiproliferative and antiviral signals in intestinal epithelial cells and murine CMV infection increases colonic IL-28A expression. *American journal of physiology. Gastrointestinal and liver physiology.* 2005;289(5):G960-968.
162. Robek MD, Boyd BS, Chisari FV. Lambda interferon inhibits hepatitis B and C virus replication. *Journal of virology.* 2005;79(6):3851-3854.
163. Hou W, Wang X, Ye L, et al. Lambda interferon inhibits human immunodeficiency virus type 1 infection of macrophages. *Journal of virology.* 2009;83(8):3834-3842.
164. Zeuzem S, Arora S, Bacon B, et al. Pegylated interferon-lambda (pegIFN-[lambda]) shows superior viral response with improved safety and tolerability versus pegIFN-[alpha]-2a in HCV patients (G1/2/3/4): EMERGE phase IIb through week 12. Paper presented at: 46th Annual Meeting of the European Association for the Study of the Liver (EASL 2011). ; March 30-April 3 2011, 2011; Berlin.
165. Honda M, Sakai A, Yamashita T, et al. Hepatic ISG expression is associated with genetic variation in interleukin 28B and the outcome of IFN therapy for chronic hepatitis C. *Gastroenterology.* 2010;139(2):499-509.
166. Urban TJ, Thompson AJ, Bradrick SS, et al. IL28B genotype is associated with differential expression of intrahepatic interferon-stimulated genes in patients with chronic hepatitis C. *Hepatology.* 2010;52(6):1888-1896.
167. Prokunina-Olsson L, Muchmore B, Tang W, et al. A variant upstream of IFNL3 (IL28B) creating a new interferon gene IFNL4 is associated with impaired clearance of hepatitis C virus. *Nature genetics.* 2013;45(2):164-171.
168. Aka PV, Kuniholm MH, Pfeiffer RM, et al. Association of the IFNL4-DeltaG Allele With Impaired Spontaneous Clearance of Hepatitis C Virus. *The Journal of infectious diseases.* 2014;209(3):350-354.
169. O'Brien TR, Prokunina-Olsson L, Donnelly RP. IFN-lambda4: The Paradoxical New Member of the Interferon Lambda Family. *Journal of interferon & cytokine research :*

the official journal of the International Society for Interferon and Cytokine Research. 2014.
170. Bibert S, Roger T, Calandra T, et al. IL28B expression depends on a novel TT/-G polymorphism which improves HCV clearance prediction. *The Journal of experimental medicine.* 2013;210(6):1109-1116.
171. McFarland AP, Horner SM, Jarret A, et al. The favorable IFNL3 genotype escapes mRNA decay mediated by AU-rich elements and hepatitis C virus-induced microRNAs. *Nature immunology.* 2014;15(1):72-79.
172. Kotenko SV, Gallagher G, Baurin VV, et al. IFN-lambdas mediate antiviral protection through a distinct class II cytokine receptor complex. *Nature immunology.* 2003;4(1):69-77.
173. Marcello T, Grakoui A, Barba-Spaeth G, et al. Interferons alpha and lambda inhibit hepatitis C virus replication with distinct signal transduction and gene regulation kinetics. *Gastroenterology.* 2006;131(6):1887-1898.
174. Morgan TR, O'Brien TR. IL28B-genotype testing now and in the era of direct-acting antiviral agents. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association.* 2011;9(4):293-294.
175. Raglow Z, Thoma-Perry C, Gilroy R, Wan Y-JY. IL28B genotype and the expression of ISGs in normal liver. *Liver International.* 2013;33(7):991-998.
176. Lalor PF, Shields P, Grant A, Adams DH. Recruitment of lymphocytes to the human liver. *Immunology and cell biology.* 2002;80(1):52-64.
177. Sitia G, De Bona A, Bagaglio S, et al. Naive HIV/HCV-coinfected patients have higher intrahepatic pro-inflammatory cytokines than coinfected patients treated with antiretroviral therapy. *Antiviral therapy.* 2006;11(3):385-389.
178. Kuntzen T, Tural C, Li B, et al. Intrahepatic mRNA expression in hepatitis C virus and HIV/hepatitis C virus co-infection: infiltrating cells, cytokines, and influence of HAART. *AIDS.* 2008;22(2):203-210.
179. Jordan WJ, Eskdale J, Srinivas S, et al. Human interferon lambda-1 (IFN-lambda1/IL-29) modulates the Th1/Th2 response. *Genes and immunity.* 2007;8(3):254-261.
180. Morrow MP, Yan J, Pankhong P, et al. IL-28B/IFN-lambda 3 drives granzyme B loading and significantly increases CTL killing activity in macaques. *Molecular therapy : the journal of the American Society of Gene Therapy.* 2010;18(9):1714-1723.
181. Thurairajah PH, Hegazy D, Chokshi S, et al. Hepatitis C virus (HCV)--specific T cell responses in injection drug users with apparent resistance to HCV infection. *The Journal of infectious diseases.* 2008;198(12):1749-1755.
182. Kamal SM, Fehr J, Roesler B, Peters T, Rasenack JW. Peginterferon alone or with ribavirin enhances HCV-specific CD4 T-helper 1 responses in patients with chronic hepatitis C. *Gastroenterology.* 2002;123(4):1070-1083.
183. Lechner F, Wong DK, Dunbar PR, et al. Analysis of successful immune responses in persons infected with hepatitis C virus. *The Journal of experimental medicine.* 2000;191(9):1499-1512.
184. Marchetti G, Tincati C, Silvestri G. Microbial translocation in the pathogenesis of HIV infection and AIDS. *Clinical microbiology reviews.* 2013;26(1):2-18.
185. Giorgi JV, Hultin LE, McKeating JA, et al. Shorter survival in advanced human immunodeficiency virus type 1 infection is more closely associated with T

lymphocyte activation than with plasma virus burden or virus chemokine coreceptor usage. *The Journal of infectious diseases.* 1999;179(4):859-870.
186. Caradonna L, Mastronardi ML, Magrone T, et al. Biological and clinical significance of endotoxemia in the course of hepatitis C virus infection. *Current pharmaceutical design.* 2002;8(11):995-1005.
187. Garcia-Alvarez M, Berenguer J, Guzman-Fulgencio M, et al. Bacterial DNA translocation and liver disease severity among HIV-infected patients with chronic hepatitis C. *J Acquir Immune Defic Syndr.* 2012;61(5):552-556.
188. Marchetti G, Cozzi-Lepri A, Tincati C, et al. Immune activation and microbial translocation in liver disease progression in HIV/hepatitis co-infected patients: results from the Icona Foundation study. *BMC infectious diseases.* 2014;14:79.
189. Balagopal A, Philp FH, Astemborski J, et al. Human immunodeficiency virus-related microbial translocation and progression of hepatitis C. *Gastroenterology.* 2008;135(1):226-233.
190. Neaton JD, Neuhaus J, Emery S. Soluble biomarkers and morbidity and mortality among people infected with HIV: summary of published reports from 1997 to 2010. *Current opinion in HIV and AIDS.* 2010;5(6):480-490.
191. Park SH, Kim CH, Kim DJ, et al. Diagnostic value of multiple biomarker panel for prediction of significant fibrosis in chronic hepatitis C. *Clinical biochemistry.* 2011;44(17-18):1396-1399.
192. Pepe MS, Janes H, Longton G, Leisenring W, Newcomb P. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *American journal of epidemiology.* 2004;159(9):882-890.
193. Ware JH. The limitations of risk factors as prognostic tools. *The New England journal of medicine.* 2006;355(25):2615-2617.
194. Gressner AM, Weiskirchen R, Breitkopf K, Dooley S. Roles of TGF-beta in hepatic fibrosis. *Frontiers in bioscience : a journal and virtual library.* 2002;7:d793-807.
195. Annoni G, Weiner FR, Zern MA. Increased transforming growth factor-beta 1 gene expression in human liver disease. *Journal of hepatology.* 1992;14(2-3):259-264.
196. Poggi A, Zocchi MR. HIV-1 Tat triggers TGF-beta production and NK cell apoptosis that is prevented by pertussis toxin B. *Clinical & developmental immunology.* 2006;13(2-4):369-372.
197. Poli G, Kinter A, Justement JS, et al. Tumor necrosis factor alpha functions in an autocrine manner in the induction of human immunodeficiency virus expression. *Proceedings of the National Academy of Sciences of the United States of America.* 1990;87(2):782-785.
198. Poli G, Bressler P, Kinter A, et al. Interleukin 6 induces human immunodeficiency virus expression in infected monocytic cells alone and in synergy with tumor necrosis factor alpha by transcriptional and post-transcriptional mechanisms. *The Journal of experimental medicine.* 1990;172(1):151-158.
199. Breen EC, Rezai AR, Nakajima K, et al. Infection with HIV is associated with elevated IL-6 levels and production. *J Immunol.* 1990;144(2):480-484.
200. Folks TM, Clouse KA, Justement J, et al. Tumor necrosis factor alpha induces expression of human immunodeficiency virus in a chronically infected T-cell clone. *Proceedings of the National Academy of Sciences of the United States of America.* 1989;86(7):2365-2368.

201. Decrion AZ, Dichamp I, Varin A, Herbein G. HIV and inflammation. *Current HIV research.* 2005;3(3):243-259.
202. Lau B, Sharrett AR, Kingsley LA, et al. C-reactive protein is a marker for human immunodeficiency virus disease progression. *Archives of internal medicine.* 2006;166(1):64-70.
203. Pinato DJ, Bains J, Irkulla S, et al. Advanced age influences the dynamic changes in circulating C-reactive protein following injury. *Journal of clinical pathology.* 2013;66(8):695-699.
204. Rodger AJ, Fox Z, Lundgren JD, et al. Activation and coagulation biomarkers are independent predictors of the development of opportunistic disease in patients with HIV infection. *The Journal of infectious diseases.* 2009;200(6):973-983.
205. Kuller LH, Tracy R, Belloso W, et al. Inflammatory and coagulation biomarkers and mortality in patients with HIV infection. *PLoS medicine.* 2008;5(10):e203.
206. Reingold J, Wanke C, Kotler D, et al. Association of HIV infection and HIV/HCV coinfection with C-reactive protein levels: the fat redistribution and metabolic change in HIV infection (FRAM) study. *J Acquir Immune Defic Syndr.* 2008;48(2):142-148.
207. Polyak SJ, Khabar KS, Rezeiq M, Gretch DR. Elevated levels of interleukin-8 in serum are associated with hepatitis C virus infection and resistance to interferon therapy. *Journal of virology.* 2001;75(13):6209-6211.
208. Zimmermann HW, Seidler S, Gassler N, et al. Interleukin-8 is activated in patients with chronic liver diseases and associated with hepatic macrophage accumulation in human liver fibrosis. *PloS one.* 2011;6(6):e21381.
209. Clement S, Pascarella S, Conzelmann S, Gonelle-Gispert C, Guilloux K, Negro F. The hepatitis C virus core protein indirectly induces alpha-smooth muscle actin expression in hepatic stellate cells via interleukin-8. *Journal of hepatology.* 2010;52(5):635-643.
210. Balasubramanian A, Ganju RK, Groopman JE. Hepatitis C virus and HIV envelope proteins collaboratively mediate interleukin-8 secretion through activation of p38 MAP kinase and SHP2 in hepatocytes. *The Journal of biological chemistry.* 2003;278(37):35755-35766.
211. Schwabe RF, Bataller R, Brenner DA. Human hepatic stellate cells express CCR5 and RANTES to induce proliferation and migration. *American journal of physiology. Gastrointestinal and liver physiology.* 2003;285(5):G949-958.
212. Reeves HL, Friedman SL. Activation of hepatic stellate cells--a key issue in liver fibrosis. *Frontiers in bioscience : a journal and virtual library.* 2002;7:d808-826.
213. Bruno R, Galastri S, Sacchi P, et al. gp120 modulates the biology of human hepatic stellate cells: a link between HIV infection and liver fibrogenesis. *Gut.* 2010;59(4):513-520.
214. Moura AS, Carmo RA, Teixeira AL, Leite VH, Rocha MO. Soluble inflammatory markers as predictors of liver histological changes in patients with chronic hepatitis C virus infection. *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology.* 2010;29(9):1153-1161.
215. Wolf K, Tsakiris DA, Weber R, Erb P, Battegay M. Antiretroviral therapy reduces markers of endothelial and coagulation activation in patients infected with human

immunodeficiency virus type 1. *The Journal of infectious diseases.* 2002;185(4):456-462.
216. Giron-Gonzalez J. Adhesion molecules as a prognostic marker of liver cirrhosis. *Scand J Gastroenterol* 2005;40:217-224.
217. Shive CL, Jiang W, Anthony DD, Lederman MM. Soluble CD14 is a nonspecific marker of monocyte activation. *AIDS.* 2015;29(10):1263-1265.
218. Sandler NG, Wand H, Roque A, et al. Plasma levels of soluble CD14 independently predict mortality in HIV infection. *The Journal of infectious diseases.* 2011;203(6):780-790.
219. Marchetti G, Nasta P, Bai F, et al. Circulating sCD14 is associated with virological response to pegylated-interferon-alpha/ribavirin treatment in HIV/HCV co-infected patients. *PloS one.* 2012;7(2):e32028.
220. Hardy GH. Mendelian proportions in a mixed population. 1908. *The Yale Journal of Biology and Medicine.* 2003;76(2):79-80.
221. Little J, Higgins JP, Ioannidis JP, et al. Strengthening the reporting of genetic association studies (STREGA): an extension of the STROBE statement. *European journal of epidemiology.* 2009;24(1):37-55.
222. Salanti G, Amountza G, Ntzani EE, Ioannidis JP. Hardy-Weinberg equilibrium in genetic association studies: an empirical evaluation of reporting, deviations, and power. *European journal of human genetics : EJHG.* 2005;13(7):840-848.
223. Tregouet DA, Tiret L. Cox proportional hazards survival regression in haplotype-based association analysis using the Stochastic-EM algorithm. *European journal of human genetics : EJHG.* 2004;12(11):971-974.
224. Lin DY, Zeng D, Millikan R. Maximum likelihood estimation of haplotype effects and haplotype-environment interactions in association studies. *Genetic epidemiology.* 2005;29(4):299-312.
225. Schisterman EF, Cole SR, Ye A, Platt RW. Accuracy loss due to selection bias in cohort studies with left truncation. *Paediatric and perinatal epidemiology.* 2013;27(5):491-502.
226. Bacchetti P, Tien PC, Seaberg EC, et al. Estimating past hepatitis C infection risk from reported risk factor histories: implications for imputing age of infection and modeling fibrosis progression. *BMC infectious diseases.* 2007;7:145.
227. Mansournia MA, Altman DG. Inverse probability weighting. *BMJ.* 2016;352:i189.
228. Statistics C. *2011 National Household Survey: Aboriginal Peoples in Canada: First Nations People, Métis and Inuit.* 2011.
229. Pritchard JK, Przeworski M. Linkage disequilibrium in humans: models and data. *American journal of human genetics.* 2001;69(1):1-14.
230. Slatkin M. Linkage disequilibrium--understanding the evolutionary past and mapping the medical future. *Nat Rev Genet.* 2008;9(6):477-485.
231. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics.* 2005;21(2):263-265.
232. Gabriel SB, Schaffner SF, Nguyen H, et al. The structure of haplotype blocks in the human genome. *Science.* 2002;296(5576):2225-2229.
233. Stephens M, Donnelly P. A comparison of bayesian methods for haplotype reconstruction from population genotype data. *American journal of human genetics.* 2003;73(5):1162-1169.

234. Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. *American journal of human genetics.* 2001;68(4):978-989.
235. Fallin D, Schork NJ. Accuracy of haplotype frequency estimation for biallelic loci, via the expectation-maximization algorithm for unphased diploid genotype data. *American journal of human genetics.* 2000;67(4):947-959.
236. Lin DY, Zeng D, Sabatti C, et al. Likelihood-Based Inference on Haplotype Effects in Genetic Association Studies [with Comment]. *Journal of the American Statistical Association.* 2006;101(473):89-118.
237. Wang CC, Krantz E, Klarquist J, et al. Acute hepatitis C in a contemporary US cohort: modes of acquisition and factors influencing viral clearance. *The Journal of infectious diseases.* 2007;196(10):1474-1482.
238. Kamal SM, Kassim SK, Ahmed AI, et al. Host and viral determinants of the outcome of exposure to HCV infection genotype 4: a large longitudinal study. *The American journal of gastroenterology.* 2014;109(2):199-211.
239. Rao HY, Sun DG, Jiang D, et al. IL28B genetic variants and gender are associated with spontaneous clearance of hepatitis C virus infection. *Journal of viral hepatitis.* 2012;19(3):173-181.
240. Dawood M, Smart G, Wood M, Wu HX, Paton S, Wu J. Hepatitis C virus infection among First Nation and non-First Nation people in Manitoba, Canada: a public health laboratory study. *Canadian journal of microbiology.* 2006;52(10):999-1005.
241. Mosbruger TL, Duggal P, Goedert JJ, et al. Large-scale candidate gene analysis of spontaneous clearance of hepatitis C virus. *The Journal of infectious diseases.* 2010;201(9):1371-1380.
242. Indolfi G, Mangone G, Calvo PL, et al. Interleukin 28B rs12979860 single-nucleotide polymorphism predicts spontaneous clearance of hepatitis C virus in children. *Journal of pediatric gastroenterology and nutrition.* 2014;58(5):666-668.
243. Indolfi G, Sambrotta M, Moriondo M, Azzari C, Resti M. Genetic variation in interleukin-28B locus is associated with spontaneous clearance of HCV in children with non-1 viral genotype infection. *Hepatology.* 2011;54(4):1490-1491.
244. de Castellarnau M, Aparicio E, Parera M, et al. Deciphering the interleukin 28B variants that better predict response to pegylated interferon-alpha and ribavirin therapy in HCV/HIV-1 coinfected patients. *PloS one.* 2012;7(2):e31016.
245. Minuk GY, Zhang M, Wong SG, et al. Viral hepatitis in a Canadian First Nations community. *Canadian journal of gastroenterology = Journal canadien de gastroenterologie.* 2003;17(10):593-596.
246. Minuk GY, O'Brien M, Hawkins K, et al. Treatment of chronic hepatitis C in a Canadian Aboriginal population: results from the PRAIRIE study. *Canadian journal of gastroenterology = Journal canadien de gastroenterologie.* 2013;27(12):707-710.
247. Wang S, Lewis CM, Jakobsson M, et al. Genetic variation and population structure in native Americans. *PLoS genetics.* 2007;3(11):e185.
248. Reich D, Patterson N, Campbell D, et al. Reconstructing Native American population history. *Nature.* 2012;488(7411):370-374.
249. van den Berg CH, Grady BP, Schinkel J, et al. Female Sex and IL28B, a Synergism for Spontaneous Viral Clearance in Hepatitis C Virus (HCV) Seroconverters from a Community-Based Cohort. *PloS one.* 2011;6(11):e27555.

250. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in medicine.* 2011;30(4):377-399.
251. Clausen LN, Weis N, Astvad K, et al. Interleukin-28B polymorphisms are associated with hepatitis C virus clearance and viral load in a HIV-1-infected cohort. *Journal of viral hepatitis.* 2011;18(4):e66-74.
252. Clausen LN, Lundbo LF, Benfield T. Hepatitis C virus infection in the human immunodeficiency virus infected patient. *World journal of gastroenterology : WJG.* 2014;20(34):12132-12143.
253. di Iulio J, Ciuffi A, Fitzmaurice K, et al. Estimating the net contribution of interleukin-28B variation to spontaneous hepatitis C virus clearance. *Hepatology.* 2011;53(5):1446-1454.
254. Pedergnana V, Abdel-Hamid M, Guergnon J, et al. Analysis of IL28B variants in an Egyptian population defines the 20 kilobases minimal region involved in spontaneous clearance of hepatitis C virus. *PloS one.* 2012;7(6):e38578.
255. Suppiah V, Gaudieri S, Armstrong NJ, et al. IL28B,HLA-C, and KIR Variants Additively Predict Response to Therapy in Chronic Hepatitis C Virus Infection in a European Cohort: A Cross-Sectional Study. *PLoS medicine.* 2011;8(9):e1001092.
256. The International HapMap Project. *Nature.* 2003;426(6968):789-796.
257. Uhanova J, Tate RB, Tataryn DJ, Minuk GY. The epidemiology of hepatitis C in a Canadian Indigenous population. *Canadian journal of gastroenterology = Journal canadien de gastroenterologie.* 2013;27(6):336-340.
258. Sugiyama M, Tanaka Y, Wakita T, Nakanishi M, Mizokami M. Genetic variation of the IL-28B promoter affecting gene expression. *PloS one.* 2011;6(10):e26620.
259. Urban T, Charlton MR, Goldstein DB. Introduction to the genetics and biology of interleukin-28B. *Hepatology.* 2012;56(1):361-366.
260. Chinnaswamy S. Genetic Variants at the IFNL3 Locus and Their Association with Hepatitis C Virus Infections Reveal Novel Insights into Host-Virus Interactions. *Journal of interferon & cytokine research : the official journal of the International Society for Interferon and Cytokine Research.* 2014.
261. Lin DY. Haplotype-based association analysis in cohort studies of unrelated individuals. *Genetic epidemiology.* 2004;26(4):255-264.
262. Lin DY, Hu Y, Huang BE. Simple and efficient analysis of disease association with missing genotype data. *American journal of human genetics.* 2008;82(2):444-452.
263. Moqueet N, Infante-Rivard C, Platt RW, et al. Favourable IFNL3 genotypes are associated with spontaneous clearance and are differentially distributed in Aboriginals in Canadian HIV-hepatitis C co-infected individuals. *International journal of molecular sciences.* 2015;16(3):6496-6512.
264. Probst A, Dang T, Bochud M, Egger M, Negro F, Bochud PY. Role of hepatitis C virus genotype 3 in liver fibrosis progression--a systematic review and meta-analysis. *Journal of viral hepatitis.* 2011;18(11):745-759.
265. *Stata Statistical Software: Release 12* [computer program]. Version 12. College Station, TX2011.
266. D'Ambrosio R, Aghemo A, De Francesco R, et al. The association of IL28B genotype with the histological features of chronic hepatitis C is HCV genotype dependent. *International journal of molecular sciences.* 2014;15(5):7213-7224.

267. Naggie S, Osinusi A, Katsounas A, et al. Dysregulation of innate immunity in hepatitis C virus genotype 1 IL28B-unfavorable genotype patients: impaired viral kinetics and therapeutic response. *Hepatology.* 2012;56(2):444-454.
268. Alter HJ, Seeff LB. Recovery, persistence, and sequelae in hepatitis C virus infection: a perspective on long-term outcome. *Seminars in liver disease.* 2000;20(1):17-35.
269. D'Agostino RB, Lee ML, Belanger AJ, Cupples LA, Anderson K, Kannel WB. Relation of pooled logistic regression to time dependent Cox regression analysis: the Framingham Heart Study. *Statistics in medicine.* 1990;9(12):1501-1515.
270. Cook NR. Statistical evaluation of prognostic versus diagnostic models: beyond the ROC curve. *Clinical chemistry.* 2008;54(1):17-23.
271. Sanderson J, Thompson SG, White IR, Aspelund T, Pennells L. Derivation and assessment of risk prediction models using case-cohort data. *BMC medical research methodology.* 2013;13:113.
272. Prentice RL. A case-cohort design for epidemiologic cohort studies and disease prevention trials. *Biometrika.* 1986;73(1):1-11.
273. Barlow WE, Ichikawa L, Rosner D, Izumi S. Analysis of case-cohort designs. *Journal of clinical epidemiology.* 1999;52(12):1165-1172.
274. Rembeck K, Westin J, Lindh M, et al. Association between interleukin-28B-related genetic variants and liver histopathology differs between hepatitis C virus genotypes. *Hepatology.* 2012;56(1):394.
275. Marti H, Chavance M. Multiple imputation analysis of case-cohort studies. *Statistics in medicine.* 2011;30(13):1595-1607.
276. Janes H, Pepe MS, Gu W. Assessing the value of risk predictions by using risk stratification tables. *Annals of internal medicine.* 2008;149(10):751-760.
277. Kattan MW. Judging new markers by their ability to improve predictive accuracy. *Journal of the National Cancer Institute.* 2003;95(9):634-635.
278. Carney PA, Miglioretti DL, Yankaskas BC, et al. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. *Annals of internal medicine.* 2003;138(3):168-175.
279. Pepe MS, Feng Z, Huang Y, et al. Integrating the predictiveness of a marker with its performance as a classifier. *American journal of epidemiology.* 2008;167(3):362-368.
280. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation.* 2007;115(7):928-935.
281. Harrell F. Regression modeling strategies. With applications to linear models, logistic regression and survival. 2001.
282. Lo Re V, Kallan MJ, Tate JP, et al. Predicting Risk of End-Stage Liver Disease in Antiretroviral-Treated Human Immunodeficiency Virus/Hepatitis C Virus-Coinfected Patients. *Open Forum Infectious Diseases.* 2015;2(3):ofv109.
283. Tate JP, Justice AC, Hughes MD, et al. An internationally generalizable risk index for mortality after one year of antiretroviral therapy. *AIDS (London, England).* 2013;27(4):563-572.
284. Kamath PS, Kim WR, Advanced Liver Disease Study G. The model for end-stage liver disease (MELD). *Hepatology.* 2007;45(3):797-805.
285. Yourman LC, Lee SJ, Schonberg MA, Widera EW, Smith AK. Prognostic indices for older adults: a systematic review. *JAMA : the journal of the American Medical Association.* 2012;307(2):182-192.

286. Ridker PM, Rifai N, Rose L, Buring JE, Cook NR. Comparison of C-Reactive Protein and Low-Density Lipoprotein Cholesterol Levels in the Prediction of First Cardiovascular Events. *New England Journal of Medicine.* 2002;347(20):1557-1565.
287. Wilson PF, Nam B, Pencina M, et al. C-reactive protein and risk of cardiovascular disease in men and women from the framingham heart study. *Archives of internal medicine.* 2005;165(21):2473-2478.
288. Gronnesby JK, Borgan O. A method for checking regression models in survival analysis based on the risk score. *Lifetime data analysis.* 1996;2(4):315-328.
289. Royston P. Tools for checking calibration of a Cox model in external validation: Prediction of population-averaged survival curves based on risk groups. *Stata Journal.* 2015;15(1):275-291.
290. Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. *BMC medical research methodology.* 2013;13:33.
291. Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Statistics in medicine.* 2008;27(2):157-172; discussion 207-112.
292. Vergniol J, Boursier J, Coutzac C, et al. Evolution of noninvasive tests of liver fibrosis is associated with prognosis in patients with chronic hepatitis C. *Hepatology.* 2014;60(1):65-76.
293. Price JC, Thio CL. Liver Disease in the HIV-Infected Individual. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association.* 2010;8(12):1002-1012.
294. Lo Re V, 3rd, Kallan MJ, Tate JP, et al. Hepatic decompensation in antiretroviral-treated patients co-infected with HIV and hepatitis C virus compared with hepatitis C virus-monoinfected patients: a cohort study. *Annals of internal medicine.* 2014;160(6):369-379.
295. American Association for the Study of Liver Diseases and the Infectious Diseases Society of America. *HCV Guidance: Recommendations for Testing, Managing, and Treating Hepatitis C.* 2015.
296. European Association for the Study of the LiverAsociacion Latinoamericana para el Estudio del H. EASL-ALEH Clinical Practice Guidelines: Non-invasive tests for evaluation of liver disease severity and prognosis. *Journal of hepatology.* 2015;63(1):237-264.
297. *British Columbia PharmaCare Information for Prescribers: Determining fibrosis stage for the treatment of chronic hepatitis C.* British Columbia PharmaCare;October 2014.
298. Imbert-Bismut F, Ratziu V, Pieroni L, Charlotte F, Benhamou Y, Poynard T. Biochemical markers of liver fibrosis in patients with hepatitis C virus infection: a prospective study. *Lancet.* 2001;357(9262):1069-1075.
299. Adams LA, Bulsara M, Rossi E, et al. Hepascore: an accurate validated predictor of liver fibrosis in chronic hepatitis C infection. *Clinical chemistry.* 2005;51(10):1867-1873.
300. Forns X, Ampurdanès S, Llovet JM, et al. Identification of chronic hepatitis C patients without hepatic fibrosis by a simple predictive model. *Hepatology.* 2002;36(4):986-992.
301. Royston P, Moons KG, Altman DG, Vergouwe Y. Prognosis and prognostic research: Developing a prognostic model. *BMJ.* 2009;338:b604.

302. Newson RB. Comparing the predictive powers of survival models using Harrell's C or Somers' D. *Stata Journal.* 2010;10(3):339-358.
303. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in medicine.* 1996;15(4):361-387.
304. Marti H, Carcaillon L, Chavance M. Multiple imputation for estimating hazard ratios and predictive abilities in case-cohort surveys. *BMC medical research methodology.* 2012;12:24.
305. Leening MJ, Vedder MM, Witteman JC, Pencina MJ, Steyerberg EW. Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician's guide. *Annals of internal medicine.* 2014;160(2):122-131.
306. Leening MJG, Vedder MM, Witteman JCM, Pencina MJ, Steyerberg EW. Net Reclassification Improvement: Computation, Interpretation, and ControversiesA Literature Review and Clinician's Guide. *Annals of internal medicine.* 2014;160(2):122-131.
307. Zeremski M, Dimova R, Brown Q, Jacobson IM, Markatou M, Talal AH. Peripheral CXCR3-Associated Chemokines as Biomarkers of Fibrosis in Chronic Hepatitis C Virus Infection. *Journal of Infectious Diseases.* 2009;200(11):1774-1780.
308. Zeremski M, Petrovic LM, Talal AH. The role of chemokines as inflammatory mediators in chronic hepatitis C virus infection. *Journal of viral hepatitis.* 2007;14(10):675-687.
309. Helbig KJ, Ruszkiewicz A, Semendric L, Harley HA, McColl SR, Beard MR. Expression of the CXCR3 ligand I‐TAC by hepatocytes in chronic hepatitis C and its correlation with hepatic inflammation. *Hepatology.* 2004;39(5):1220-1229.