Section 2: English Abstract

Full title of research project

Non-alcoholic Fatty Liver and Liver Fibrosis Progression Predictive Analytics: Risk Prediction and Machine Learning Techniques for Improved Preventive Medicine

Investigators: full name and affiliations

Ofir Ben-Assuli, Ono Academic College, Business Administration
Shani Shenhar-Tsarfaty, Tel-Aviv University, Sackler Faculty of Medicine
Shira Zelber-Sagi, University of Haifa, School of Public Health, Faculty of Social Welfare and Health Sciences

Key Words

Machine Learning, Risk Prediction, Non-alcoholic fatty liver disease, Predictive Analytics

Abstract

Scientific Background
Non-alcoholic fatty liver disease (NAFLD) is the most common liver disease worldwide, with a prevalence of 20-30% in the general population. While most **NAFLD** patients are asymptomatic**,** NAFLD may progress to cirrhosis and cardiovascular disease and Type-2 diabetes.
Objectives
To stratify patients' risks for NAFLD and advanced fibrosis over time and suggest preventive medical decisions.
Methodology
The data were obtained from a prospective cohort of apparently healthy volunteers from the Tel-Aviv medical center inflammation survey (TAMCIS), admitted for a routine annual health check-up. We stratified the patients' risks for fibrosis from 2007 to 2017 by modeling the risk in 5,579 individuals in 5 main stages: 1. Compilation of TAMCIS database. 2. Data preparation and computation of time series variables 3. Application of time-series models (Hidden Markov Models and Group-Based Trajectory Models) to generate distinct fibrosis risk trajectories. 4. Predictive analytics & Sensitivity Analysis incorporating trajectory group membership to predict individual patient risk. 5. Profiling the patients and Recommendations.
Findings
Time-series machine learning models (Hidden Markov Models and Group-Based Trajectory Models) that profiled fibrosis risk by modeling patients' latent medical status generated three groups. The high-risk group had abnormal lab test values and a higher prevalence of chronic conditions. Classification models with group information showed that they significantly improved risk fibrosis prediction.
Conclusions
Longitudinal risk assessment (optionally accessed via EMRs/EHRs) may enable early identification of specific individual groups exhibiting distinct medical trajectories based on their routine visits.
Policy Recommendations
This process can be used to make population-specific medical recommendations targeting early detection of high risk patients, to avoid the progression of fibrosis disease and its complications as well as reduce healthcare costs. The findings contribute to risk stratification for disease progression and planning future preventive strategies based on lifestyle modifications and medical treatment implemented to reduce disease burden.

Section 3: Hebrew Abstract

כותרת מלאה של המחקר

חיזוי התקדמות מחלת כבד שומני לא אלכוהולי ופיברוזיס באמצעות כריית נתונים מתקדמת לצרכי רפואה מונעת

חוקרים: שם מלא ושיוך מוסדי

אופיר בן אסולי, הקריה האקדמית אונו, הפקולטה למנהל עסקים
שני שנהר-צרפתי, אוניברסיטת תל אביב, הפקולטה לרפואה על שם סוראסקי
שירה זלבר-שגיא, אוניברסיטת חיפה, בית הספר לבריאות הציבור

מילות מפתח

לימוד מכונה, ניתוח סיכונים לאורך זמן, מחלת כבד שומני לא אלכוהולי ופיברוזיס, כריית נתונים

תקציר

רקע מדעי
כבד שומני לא-אלכוהולי הפך לנטל משמעותי בבריאות הציבור עקב הארעותו הגבוהה והסיכון שלו להתקדם לשחמת כבד, סכרת מסוג 2 ומחלות לב וכלי דם. זיהוי מוקדם של חולים עם כבד שומני שנמצאים בסיכון להתקדמות המחלה הינו קריטי ליישום התערבויות מניעתיות וטיפוליות.
מטרות
בניית מודלי חיזוי לסיכון למחלת כבד והתדרדרות לאורך זמן על מנת לבנות המלצות מחויטות אישית לאיתור מוקדם ולרפואה מונעת של כבד שומני לא-אלכוהולי או סיבוכיו כולל פיברוזיס.
שיטה
הנתונים הגיעו ממרכז תל-אביב לחקר הדלקת (TAMCIS) בין 2007-2017 של 5,579 נבדקים. השלבים המרכזיים: 1) שליפת נתונים והכנתם. 2) ניתוחי סדרות עיתיות שבוחנות את השינויים לאורך זמן והקבצה לאשכולות (לפי רמות סיכון שונות) של הנבדקים לפי מודלים מתאימים (HMM ו- GBTM). 3) ניתוחי ניבוי וניתוחי רגישות. 4) אפיון נבדקים והמלצות.
ממצאים עיקריים
המודלים שהשתמשנו בהם אפיינו את הנבדקים לאורך זמן וזיהו שלוש רמות סיכון לפיברוזיס. לקבוצה עם רמת סיכון גבוהה לפיתוח פיברוזיס היו ערכי בדיקות מעבדה חריגים ושכיחות אבחנות כרוניות גבוהה. סיווג על בסיס השתייכות לקבוצת סיכון שיפר את הניבוי.
המלצות למקבלי ההחלטות
חיזוי התדרדרות במדדי כבד שומני ואיתור צמתי התערבות פוטנציאלים, עשוי בו-זמנית להקטין סיבוכים ונטל כספי הנובע ממנה. המחקר יכול לתרום להבנת השפעת כריית נתונים לאורך זמן על חיזוי מחלות כדי למנוע התדרדרות ולטובת רפואה מונעת.

Section 4: Executive Summary- Hebrew

רקע מדעי

כבד שומני לא-אלכוהולי הפך לנטל משמעותי בבריאות הציבור עקב הארעותו הגבוהה (30 אחוזים מכלל האוכלוסיה) והסיכון שלו להתקדם לשחמת כבד, סכרת מסוג 2 ומחלות לב וכלי דם (Zelber-Sagi et al., 2006). זיהוי מוקדם של חולים עם כבד שומני שנמצאים בסיכון להתקדמות המחלה הינו קריטי ליישום התערבויות מניעתיות וטיפוליות. אמנם גם כיום, ביופסיה לכבד היא עדיין הסטנדרט להעריך מחלות כבד. אולם, בדיקות חוזרות מסוג ביופסיה כדי לזהות מגמות במחלות כבד הן קשות מאוד מבחינה קלינית וגם אתית. גם בפרקטיקה הרבה מאוד אנשים עם כבד שומני לא-אלכוהולי לא נבדקים וכלל לא מודעים לכך ולסיכון להתפתחות עתידית של מחלות כבד כמו פיברוזיס שאפשר למנוע לפי זיהוי מוקדם. אנו הצענו להשתמש במודלים מתקדמים לאורך זמן ללא כל בדיקה פולשנית אלא פשוט ע"י נתונים קיימים ולנתח את הסיכונים לאורך זמן למחלות כבד ע"י נתונים שגם כך זמינים ממערכות מידע רפואיות (כגון EMR או EHR).

מטרות

בניית מודלי חיזוי לסיכון למחלת כבד והתדרדרות לאורך זמן על מנת לבנות המלצות מחויטות אישית לאיתור מוקדם ולרפואה מונעת של כבד שומני לא-אלכוהולי או סיבוכיו כולל פיברוזיס מתקדמת. המטרה הייתה לפתח תהליך שלם שכולל בתוכו שלבים שמיועדים לנתח רמות סיכון למחלה לכל הנבדקים וזיהוי רמות סיכון שונות שקיימות בתוך כלל קבוצת הנבדקים.

שיטה

ניתוח הנתונים התבסס על אנליזה של מודל חיזוי על בסיס מאגר מידע של מרכז תל אביב לחקר הדלקת (TAMCIS). בסיס הנתונים כולל מידע מעמיק לאורך זמן בין 2007-2017 של כ- 5,579 נבדקים עם ביקורים חוזרים, המגיעים לבדיקות סקר רפואיות. אלו השלבים המרכזיים: 1) שליפת נתונים. 2) הכנת הנתונים כולל חישוב משתנים נדרשים לשלבי הניתוח המתקדמים. 3) ניתוחי סדרות עיתיות שבוחנות את השינויים לאורך זמן והקבצה לאשכולות (לפי רמות סיכון שונות) של החולים לפי מודלים שמנתחים שינויים לאורך זמן (HMM ו- GBTM). 4) ניתוחי כריית נתונים כולל ניתוחי רגישות לחיזוי סיבוכים קליניים עתידיים. 5) אפיון פרופיל החולים לפי השתייכות לאשכולות והמלצות להתערבות רפואית לפני התקדמות מחלות.

המשתנה התלוי להערכת המצב של מחלה בכבד היה Fibrosis-4 שמחושב על בסיס ארבעה משתנים כולל גיל, AST, ALT, platelet (Sterling et al., 2006). מדד זה בשימוש ונמצא תקף בקרב אגודות מקצועיות רבות באירופה ובאמריקה.

המשתנים הבלתי תלויים כללו: מאפיינים סוציו-אקונומיים, דמוגרפיה, הרגלי ספורט, הרגלי תזונה, בדיקות מעבדה שונות כולל כימיה, ביולוגיה, מדדים מטבוליים, אבחנות כרוניות, הערכת ביצועי הלב והכבד, רשימת תרופות בשימוש ועוד.

הממצאים

לאחר קבלת המידע על 51,078 ביקורים סוננו אלו שמכילים מידע על Fib-4 ונשארו 41,423 ביקורים. היה צורך בניתוחים לאורך זמן להשאיר את כל הנתונים פרט לביקור האחרון שלא תורם לניבוי של סיכון עתידי ולפיכך נשארו 21,526 ביקורים. מעבר מרמת ביקור לרמת מטופל הניב 9,609 נבדקים. היה צורך בניתוחים לאורך זמן להשאיר רק כאלו שיש להם לפחות שני ביקורים ולפיכך נשארו 5,579 נבדקים. הסטטיסטיקה התיאורית הראתה שתוצאות בדיקות המעבדה, הרגלי הספורט והתזונה מייצגים אוכלוסייה כללית (לא בריאה יותר ולא חולה יותר) והגיל הממוצע היה 45.8 שנים עם 71.8 אחוז מהנבדקים שהיו גברים. רוב המדגם לא עישנו אף פעם (62.8 אחוזים). בדיקות הדם היו בטווחי הנורמה ולא חריגות.

שני המודלים שהשתמשנו בהם (HMM ו- GBTM) אפיינו את הנבדקים לאורך זמן וזיהו שלוש רמות סיכון לפיברוזיס. השימוש בשני מודלים שונים שהגיעו לתוצאות דומות וזאת לאחר מבחני goodness-of-fits שונים מחזקות משמעותית את ממצאי המחקר.

בין השאר הממצאים הראו שהשיטות מפלחות טוב מאוד לשלוש רמות סיכון החל מהביקור השני (GBTM) או השלישי (HMM) וזה הולך ומשתפר עד סוף הביקורים שניתחנו (שש ביקורים). כך שלמעשה ככל שיש יותר אינפורמציה הפילוח לרמות הסיכון טוב יותר אולם כבר מן ההתחלה בביקור הראשון שהכנסנו הקבוצה עם רמת הסיכון הגבוהה ביותר לפיברוזיס זוהתה היטב ע"י המודלים.

הוספנו ניתוחי רגישות. ניתוח אחד הרבה שסיווג על בסיס השתייכות הנבדקים לקבוצות סיכון מסוימות בשיטות classification וזה שיפר משמעותית את רמת הניבוי גם ברמת הנבדק. ניתוח שני מסוג Propensity Score Matching הראה שגם אם מאזנים את הנבדקים מבחינת מאיינים

קליניים שלהם מקבלים תוצאות דומות ותומכות בתוצאות הכלליות של המחקר. תוצאות נוספות אפשר לראות בהרחבה בשני המאמרים שהתקבלו תודות למענק מחקר זה (Ben-Assuli et al., 2022; Goldman et al., 2021).

מסקנות

בקבוצת עם רמת סיכון גבוהה לפיתוח פיברוזיס היה המספר הקטן ביותר של נבדקים החל מהביקור השני (GBTM) או השלישי (HMM) והחל מזמן זה כמעט לא נרשמו מעברים של נבדקים בין קבוצות סיכון מביקור לביקור וחזינו התייצבות גבוהה של נבדקים בקבוצות שלהן. לקבוצה זו היו ערכי בדיקות מעבדה לא בטווח הנורמה ושכיחות אבחנות כרוניות גבוהה (למשל Hypertension and Diabetes). גם הקבוצה עם רמת סיכון בינונית אופיינה באבחנות כרוניות מסוימות ששוות מעקב וטיפול בפני עצמו כדי להימנע ממעבר של נבדקים אלו לרמת סיכון גבוהה לפיברוזיס.

המלצות אפשריות למקבלי ההחלטות

חיזוי התדרדרות במדדי כבד שומני ואיתור צמתי התערבות פוטנציאלים, עשוי בו-זמנית להקטין סיבוכים ונטל כספי הנובע ממנה. המחקר תורם להבנת השפעת כריית נתונים לאורך זמן על חיזוי המחלה. הטכניקות שהשתמשנו בהן יכולות לשמש לעולמות תוכן דומים ברפואה מונעת כדי לבצע סיווג, אפיון, וניבוי סיכון התדרדרות לטובת שיפור רפואה מונעת.

מקורות

Ben-Assuli, O., Jacobi, A., Goldman, O., Shenhar-Tsarfaty, S., Rogowski, O., Zeltser, D., . . . Zelber-Sagi, S. (2022). Stratifying Individuals into Non-Alcoholic Fatty Liver Disease Risk Levels using Time Series Machine Learning Models. Journal of biomedical informatics, 103986: 103981-10391

Goldman, O., Ben-Assuli, O., Rogowski, O., Zeltser, D., Shapira, I., Berliner, S., . . . Shenhar-Tsarfaty, S. (2021). Non-alcoholic fatty liver and liver fibrosis predictive analytics: risk prediction and machine learning techniques for improved preventive medicine. *Journal of medical systems*, *45*(2), 1-12 .

Sterling, R. K., Lissen, E., Clumeck, N., Sola, R., Correa, M. C., Montaner, J., . . . Thomas, D. L. (2006). Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology*, *43*(6), 1317-1325.

Zelber-Sagi, S., Nitzan-Kaluski, D., Halpern, Z., & Oren, R. (2006). Prevalence of primary non-alcoholic fatty liver disease in a population-based study and its association with biochemical and anthropometric measures. *Liver Int*, *26*(7), 856-863.

Section 5: Comprehensive scientific report - English[1]

1. Scientific background

Non-alcoholic Fatty liver disease (NAFLD) is becoming a major global health burden in both developed and developing countries (Loomba & Sanyal, 2013). NAFLD is the most prevalent liver disease (30% in the general population) and may progress to liver cirrhosis, liver cancer and can increase the risk for cardiovascular disease and type-2 diabetes (Younossi et al., 2016; S. Zelber-Sagi, Nitzan-Kaluski, Halpern, & Oren, 2006). NAFLD covers a wide pathological spectrum ranging from simple steatosis to steatohepatitis (NASH) with variable degrees of fibrosis and cirrhosis (European Association for the Study of the, European Association for the Study of, & European Association for the Study of, 2016). The presence of significant fibrosis in NASH is known to be associated with mortality (Ekstedt et al., 2015; Hagstrom et al., 2017; Younossi et al., 2016). Since NASH is rapidly becoming one of the most frequent causes of cirrhosis and liver transplantation worldwide, it is crucial to identify NAFLD patients at risk of progression, in order to implement therapeutic interventions that can help prevent/reverse the deleterious consequences of advanced NASH.

Liver biopsy remains the gold standard for histological evaluation of liver disease; however, obtaining repeated liver biopsies to identify the histological progression is difficult clinically and ethically. Therefore, non-invasive methods to identify steatosis and especially advanced fibrosis are warranted to pre-empt the need for liver biopsies. Fibrosis serum markers have shown acceptable diagnostic accuracy as defined by the area under the receiver operating characteristic curve (AUROC) >0.8 (European Association for Study of & Asociacion Latinoamericana para el Estudio del, 2015). The NAFLD fibrosis score (NFS) and fibrosis 4 calculator (FIB-4) have been externally validated in ethnically different NAFLD populations, with consistent results. NFS and FIB-4 predict overall mortality, cardiovascular mortality and liver-related mortality (European Association for the Study of the et al., 2016). The

[1] Many parts of this summery were taken from our forthcoming manuscript "Stratifying Individuals into Non-Alcoholic Fatty Liver Disease Risk Levels using Time Series Machine Learning Models" (Ben-Assuli et al., 2022) to be published by *Journal of Biomedical Informatics* and some parts were taken from our published manuscript "Non-alcoholic fatty liver and liver fibrosis predictive analytics: risk prediction and machine learning techniques for improved preventive medicine" (Goldman et al., 2021) by *Journal of medical systems.* Those two papers are results of this grant and acknowledged this grant.

diagnosis and detection of risk factors for steatosis are also important. Besides the hepatic damage, in recent years NAFLD has emerged as an independent risk factor for Type- 2 diabetes and cardiovascular disease (Stefan, Kantartzis, & Haring, 2008). Therefore, efforts to prevent NAFLD progression and extra-hepatic manifestations must include liver fat screening and surveillance strategies. The occult nature of the disease has led to increased efforts to develop simple and cost-effective diagnostic methods, preferably quantitative, that would be useful for screening, follow-up and evaluation of response to treatment in both clinical practice and research. The Fatty Liver Index (FLI) is a simple validated calculating tool for the prediction of fatty liver in the general population, and is based on body mass index (BMI), waist circumference, levels of the liver enzyme Gamma glutamyl Transferase (GGT) and plasma triglycerides (G. Bedogni et al., 2006; Shira Zelber-Sagi et al., 2013) .

NAFLD is a multifactorial disease and has many potential reasons for its progression, some less well-established than others, including genetic background, weight and weight change, the presence of Type-2 diabetes, medications and lifestyle factors such as diet, physical activity, coffee and tea drinking and smoking habits (Romero-Gomez, Zelber-Sagi, & Trenell, 2017) . For example, weight gain by itself, even if as small as 3-5 Kg predicts the development of NAFLD, regardless of baseline BMI (S. Zelber-Sagi et al., 2012). Identifying new risk factors and the elaboration of known ones is extremely important in a preventable and treatable up to complete remission disease like NAFLD.

The healthcare sector has invested heavily in health technologies that integrate clinical, demographic and financial data from multiple sources to support medical decision-making, reduce healthcare costs (Hillestad et al., 2005) and is used nowadays to collect data for analysis through sophisticated predictive analytics methods.

However, there are few prediction methods for medical outcomes after a patient's first visit. Previous works have mostly been based on narrow data, rather than on patient-level big data derived from rich datasets as in the case of CHF (Mitchell et al., 2016; Sommerfeld, Althouse, Prince, & Hickey, 2016). Few works have dealt with predicting NAFLD for FLI patients (Koehler et al., 2013; Lee et al., 2010) or predicting FLI (Ruhl & Everhart, 2015; Shira Zelber-Sagi et al., 2013). New models focusing on predicting mortality and readmission risks for chronic patients based on

data mining tools require large, broad datasets (Krumholz et al., 2009; Tsugawa et al., 2017; van Walraven et al., 2010; Wang et al., 2014).

Crucially, these predictive analytic tools can help physicians and hospital management develop early interventions to identify and treat patients at high risk of frequent readmissions (Bardhan, Oh, Zheng, & Kirksey, 2014). **This study applied a novel machine learning approach for risk stratification by tracking high risk groups for disease incidence and progression and assessed the potential protective factors.** Our approach is based on a chain of predictive analytics techniques combined to optimize prediction performance by clustering patient trajectories over time as the preliminary technique for the advanced classification methods. Our techniques were applied to the case of FLI accepted thresholds (Giorgio Bedogni et al., 2006; Cuthbertson et al., 2014; Koehler et al., 2013), and FIB-4 accepted thresholds (Kim et al., 2010; Shah et al., 2009; Sterling et al., 2006; Tapper & Lok, 2017).

2. Research objectives

The main goal of this project was to enable physicians and clinicians to stratify patients' risks for fibrosis disease progression and to suggest possible medical remedies optimal for preventive medical decisions.

More specifically, we developed an innovative chain of predictive analytic techniques (including clustering, feature selection and classification) utilizing unique data of patients having few (or multiple) chronic conditions that manifest over extended time periods. Our models were developed to identify possible variables associated with disease scores as well as the progression from low to intermediate and high disease score categories over follow-up time as well as extra-hepatic outcomes.

3. Methodology

3.1 Sample and Data

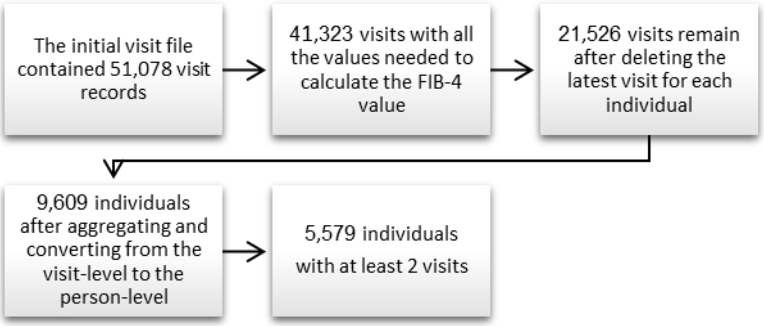

Figure 1. Flow Chart of the Data (Ben-Assuli et al., 2022; Goldman et al., 2021)

Participants - 5,579 individuals presenting at Tel Aviv Medical Center for a repetitive health checkup from July 2007 to July 2017 during routine medical tests took part in the TAMCIS. This database includes over six hundred features for each individual's visit with at least two annual health assessment visits. We constructed a FIB-4 panel over time at the individual level. The dataset contained a series of visits for all individuals where the fibrosis diagnosis appeared or not (Figure 1).

Dependent Variable - The dependent variables were NAFLD baseline markers of FIB-4. This measurement is a predictor of liver-related mortality (EASL, EASD, EASO 2016) based on four variables (So-Armah et al., 2020): age (in years), aspartate aminotransferase (AST), alanine aminotransferase (ALT) in units per liter, and total platelet sum, measured as 10^9 per liter (Zhuang et al., 2017). The FIB-4 index was computed (Sterling et al., 2006) as:

$$\frac{Age_{years} \cdot AST_{\frac{U}{L}}}{Plateletcount_{\frac{10^9}{L}} \cdot \sqrt{ALT_{\frac{U}{L}}}} \quad (1)$$

The FIB-4 score was dichotomized into two fibrosis levels (low vs. intermediate-advanced) using the accepted threshold value of 1.3 (Anstee et al., 2019) (FIB-4 < 1.3 as 'NO' and FIB-4 ≥ 1.3 as 'YES'). The individual-level cohorts were constructed from the annual individual visits, and the dependent variables were measured based on the previous visit values. For example, if a patient had five visits without an abnormal FIB-4, all four dependent variable values (previous visits) were labelled 'NO'.

The Independent Variables – social determinants, demography, exercise, lifestyle, metabolic status (e.g. metabolic syndrome, BMI and HbA1c), routine blood count results, chemistry blood test results, lipid profile, hypertension status, history of

diabetes, cardiovascular diagnoses, previous diagnoses of the liver, dyslipidemia and medication consumption.

3.2 Schematic Process Description

This study implemented a novel non-invasive assessment process for identifying fibrosis risk levels in terms of time. Figure 2 shows the different stages and the flow of the process: 1) Retrieve, merge, and integration of the individual-visit data to the visit level data. Data preparation techniques were implemented to obtain all visits and features and impute missing values or outliers. 2) Transform the visit level data into an individual level panel. 3) Product clusters of individuals based on their FIB-4 risk level. This process included two time-series models: HMM and GBTM, to profile fibrosis risk by modeling patients' medical status in groups representing the risk level of having fibrosis. 4) Sensitivity analysis which involved a propensity score matching (PSM) technique to generate a balanced dataset after which the methods were re-run. Figure 2 shows the use of HMM and GBTM with (represented by red lines), and without (represented by broken green lines) the PSM. The PSM robustness check also supports the validity and accuracy of the process results. 5) Validity analysis using classification methods (CHAID and XGBoost) based on latent class modeling to test whether the risk states (obtained from HMM/GBTM). 6) Profile the groups, analyzing the findings to evaluate the models for each latent state/trajectory and generating practical medical care recommendations.

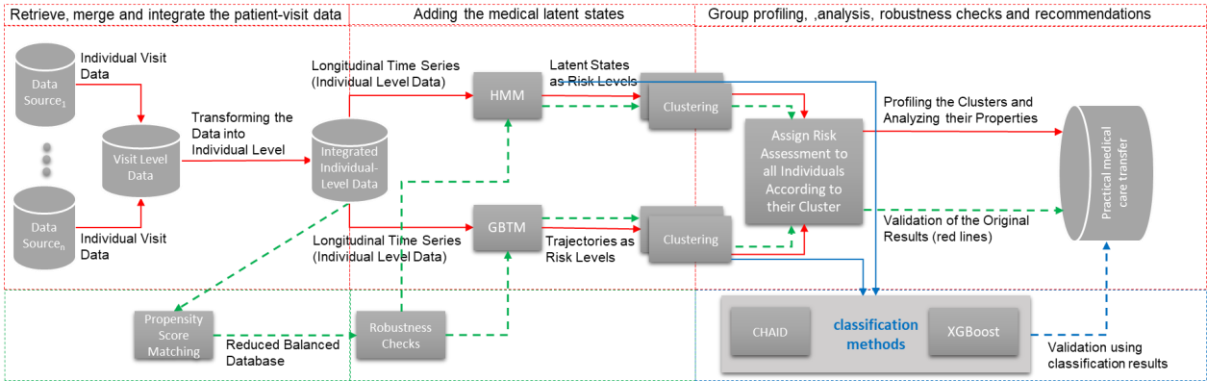

Figure 2. Schematic Process Diagram (Ben-Assuli et al., 2022)

3.3 Methods

3.3.1 Studying the longitudinal progression of diseases

GBTM (D. S. Nagin, 2005) and HMM (Rabiner & Juang, 1986) are important time-series models and are described below in detail to model the progression of NAFLD. These models can take the heterogeneity in the entire population into account as well, so that interventions can be oriented correctly to the latent subgroups to enhance desirable outcomes (Ben-Assuli & Padman, 2020). Combining these two models in our process has several advantages in that they can augment each other's results and each model supports and validates the other.

3.3.1.1 Latent class modeling using a Hidden Markov Model (HMM)

A HMM expresses the exchanges between observed and hidden stochastic processes. These processes progress over time via a set of hidden states, a sequence of observation states, a hidden state transition probability distribution, an observation state probability distribution, and an initial state distribution (Netzer, Lattin, & Srinivasan, 2008; Rabiner & Juang, 1986; Zhang & Padman, 2015). In our case, the latent health status shaped the hidden states using a first-order Markov process. The upcoming medical state's probability depended solely on the current state and not on past states (Rabiner & Juang, 1986). HMM, by decoding, calculates the series of hidden medical states with the uppermost probability given the series of observations and the probability distribution of the model. Figure 3 depicts the possible hidden or observed medical states in which a patient can be and illustrates the HMM schematically. Observation refers to data that are known and can be observed. As shown in Figure 3, the observation variable FIB-4 was the medical information available about a patient on each visit (0 for a normal FIB-4 result and 1 for an abnormal FIB-4 result). The latent medical state representing the three levels of risk for having fibrosis; i.e., high, medium and low (S_{high}, S_{medium}, or S_{low}, respectively) was hidden. Figure 3 illustrates the relationships between the set of observation states (FIB-4 index values), the set of hidden states (levels of risk for fibrosis), the transition probabilities among the hidden states, and the emission probabilities from the hidden states to the observable states.

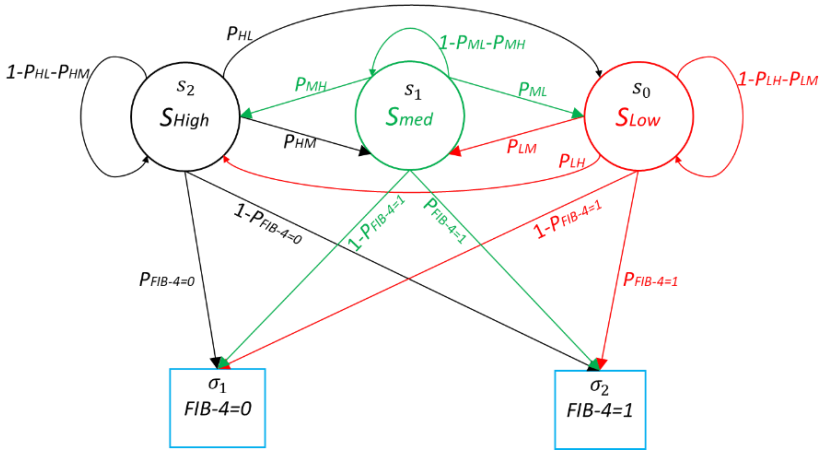

Figure 3. Representation of the HMM Process (Ben-Assuli et al., 2022)
Note: The hidden states (circles) show the patient's fibrosis risk level, whereas the observable symbols (rectangles) represent an individual's FIB-4 values over time

3.3.3.2 Group-Based Trajectory Modeling (GBTM)

GBTM is designed to identify clusters of individuals following possibly distinctive trajectories toward a certain outcome (D. S. Nagin, 2005). It is based on finite mixture modeling, to study the developmental course of patients' multiple visits. The two primary outputs of the model are the shape of the trajectory for each group, which may differ not only in its level but also in the direction and rate of its movement, and the size of the group as measured by the proportion of the population in the given trajectory. This method has been applied to study disease progression (Burckhardt, Nagin, & Padman, 2016; Daniel S Nagin & Odgers, 2010; Padman, Nagin, & Xie, 2014). The fundamental concept is the distribution of outcomes conditioned on a time-related metric such as time from the onset of treatment; that is, the distribution of outcome trajectories is denoted by $P(\mathbf{Y}_i/\mathbf{Z}_i)$, where the random vector Y_i represents individual *i*'s longitudinal sequence of behavioral outcomes and vector Z_i represents the characteristics of *i* measured at baseline or otherwise invariant. The GBTM assumes that the population distribution of trajectories arises from a finite mixture of unknown order *J*. The likelihood for each individual *i*, conditional on the number of groups *J*, can be written as $P(\mathbf{Y}_i \mid \mathbf{Z}_i) = \sum_{j=1}^{J} \pi_i P(\mathbf{Y}_i \mid \mathbf{Z}_i, j; \boldsymbol{\beta}_j)$, where π_j is the probability of membership in group *j*, and the conditional distribution of Y_i given membership in *j* is indexed by the unknown parameter vector β_j, which among other things determines the shape of the group-specific trajectory. Typically, the trajectory is modeled using a polynomial function of the time-related metric, such as the sequence of visits; i.e.,

$y_{it}^{j} = \beta_0^j + \beta_1^j \cdot FIB4_{it} + \beta_0^j \cdot FIB4_{it}^2 + \beta_0^{j \cdot} FIB4_{it}^3 + \varepsilon_t$, where for given *j*, conditional independence is assumed for the sequential realizations of the elements of Y_i, y_{it}, over the *T* periods of measurement. Thus, we can write $P(\mathbf{Y}_i \mid \mathbf{Z}_i, j; \boldsymbol{\beta}_j) = \prod_{t=1}^{T} p(y_{it} \mid z_{it}, j; \boldsymbol{\beta}_j)$, where *p*(.) is the distribution of y_{it} conditional on membership in group *j* and the z_{it} (D. S. Nagin, 2005).

See more technical details in the published paper (Ben-Assuli et al., 2022).

3.4 Classification Methods

We used two classification methods (CHAID (Kass, 1980) and XGBoost Tree (Semenov et al., 2016)) to test whether our risk states (obtained from the clustering) were helpful and improved individual-level prediction, as done in other works (Ben-Assuli & Padman, 2020; Gorunescu & Belciug, 2014). We used these classification methods to develop three classification models for the FIB-4 dependent dichotomous variables in each classification method: 1. Using all the independent variables plus HMM risk states (84 variables); 2. Using all the independent variables plus the GBTM risk states (84 variables) and 3. Using all the independent variables (83 variables) without the additional variables obtained from the clustering (identification of risk state/trajectory). The independent variables were defined for the penultimate visit in each model, and the dependent variable was from the last visit. We developed the models utilizing 5-fold cross-validation via IBM-Modeler (Modeler & Guide, 2010). For more technical details, see (Ben-Assuli et al., 2022; Goldman et al., 2021).

4. Findings

4.1 Descriptive statistics

The descriptive statistics (Table 1) included 5579 individuals who had at least two visits. The mean age was 45.8 years. The majority of the cohort were men (71.8%). The majority of the cohort had never smoked (62.3%), and the rest were current smokers or smoked in the past. The mean years between the last and the first visits was 4.4 years. The average blood test results were overall in the normal ranges.

Table 1. Sample Descriptive Statistics at Baseline (Ben-Assuli et al., 2022)

Demographics		
Age, years		45.78±10.02
Gender, n (%) Male		4505 (71.79%)
Waist (CM)		91.47±12.09
BMI (Kg/m^2)		26.39±3.98
Smoking status, n (%)	Ever	2104 (37.7%)
	Never	3471 (62.3%)
Years between the last and the first visits		4.40±2.56
Blood Tests		
AST (U/L)		23.81±7.71
ALT (U/L)		26.11±13.18
GGT (U/L)		19.72±17.02
Total Cholesterol (mg/dL)		198.76±36.47
HDL Cholesterol (mg/Dl)		54.09±13.52
Triglycerides (mg/dL)		118.43±70.30
Glucose (mg/dL)		92.53±14.87
hemA1C, %		5.36±0.57
High sensitive CRP (mg/dL)		2.55±4.32
Platelets (X109/L)		246.83±57.69
Uric Acid (mg/dL)		5.61±1.31
Albumin		45.25±2.35
Bilirubin		0.79±0.35
Alkaline Phosphatase (IU/L)		59.70±17.85
Diagnoses*		
Hypertension, n (%)		802 (14.38%)
Diabetes, n (%)		385 (6.9%)
Metabolic Syndrome, n (%)		301 (5.4%)
Hypertriglyceridemia, n (%)		1248 (22.38%)
HDL Cholesterol <LCL, n (%)		815 (14.63%)

Note: *__Metabolic syndrome__ was defined based on statement of major Federations (Alberti et al., 2009).
Hypertriglyceridemia - high triglycerides. **HDL Cholesterol <LCL** - low high-density lipoprotein-cholesterol (HDL).

4.2 HMM Results

After model fitting, the three latent HMM states were chosen for maximum consistency across all visits.

Table 3. FIB-4 Prevalence Distributed according to States (Ben-Assuli et al., 2022)

	Visit_3		**Visit_4**		**Visit _5**		**Visit _6**	
HMM	**Individuals (%)**	**FIB4>1.3 (%)**	**Individuals (%)**	**FIB4>1.3 (%)**	**Individuals (%)**	**FIB4>1.3 (%)**	**Individuals (%)**	**FIB4>1.3 (%)**
Low	2457 (77.97%)	10.012%	1253 (75.48%)	12.69%	550 (65.32%)	16.36%	293 (70.77%)	27.98%
Medium	-	-	155 (9.34%)	54.2%	178 (21.14%)	59.55%	70 (16.91%)	81.43%
High	694 (22.03%)	71.47%	252 (15.18%)	82.94%	114 (13.54)	92.98%	51 (12.32%)	96.08%
Total	3151	23.55%	1660	27.23%	842	35.87%	414	45.41%

Table 3 presents the distribution of individuals and the FIB-4 prevalence across the six visits for each of the latent states and expresses individuals' transitions across states over time. The low-risk state was the most common latent state for each period from the third visit. High-risk individuals were the smallest group in each period beyond visit 2. This group had the highest consistent FIB-4 prevalence across all visits (shown in the second column in each visit). Individuals could

change their states from visit to visit depending on their relative risk changes for FIB-4 as expressed, captured, and evaluated using the HMM. As can be seen in Table 3, generally, the number of individuals defined as high-risk decreased from visit to visit (visits: three (22%), four (15.2%), five (13.5%), and six (12.3%)). This result shows that the HMM became more discriminating as more information was loaded into the model.

Table 4. Transitions of Individuals across Visits (Ben-Assuli et al., 2022)

	Visit 2 to Visit 3	Visit 3 to Visit 4	Visit 4 to Visit 5	Visit 5 to Visit 6
Stayed the same	41.80% (1317)	87.23% (1448)	88.60% (746)	89.86% (372)
Transitioned to a lower state	56.52% (1781)	5.36% (89)	2.02% (17)	10.14% (42)
Transitioned to a higher state	1.68% (53)	7.41% (123)	9.38% (79)	0
Total	3151	1660	842	414

Table 4 displays that most of the patients stayed in the same risk state across all visits in order of increasing percentages from each visit to the subsequent one.

The profiling of individuals on the fifth visit (reported in (Ben-Assuli et al., 2022)) showed that individuals in the high-risk group had more extreme laboratory test values, demographics and diagnoses. Hypertension and Diabetes had the worst values for the high-risk group followed by the medium risk group. BMI, waist-hip measurement and Metabolic Syndrome were higher and worse, albeit not significantly, for the high-risk group, probably due to the small sample size. AST, Total Cholesterol, Glucose, Platelets, Albumin and Bilirubin were significantly different.

4.3 GBTM Results

After model fitting, the 3-group model was chosen.

Table 6. FIB-4 Prevalence Distributed according to Groups (Ben-Assuli et al., 2022)

	Visit_3		Visit_4		Visit _5		Visit _6	
GBTM	**Individuals (%)**	**FIB4>1.3 (%)**	**Individuals (%)**	**FIB4>1.3 (%)**	**Individuals (%)**	**FIB4>1.3 (%)**	**Individuals (%)**	**FIB4>1.3 (%)**
Low	2457 (77.98%)	10.01%	1196 (72.05%)	11.62%	658 (78.15%)	21.88%	291 (70.29%)	27.49%
Medium	347 (11.01%)	53.31%	322 (19.4%)	56.52%	113 (13.42%)	78.76%	83 (20.05%)	84.34%
High	347 (11.01%)	89.62%	142 (8.55%)	92.25%	71 (8.43%)	97.18%	40 (9.66%)	95%
Total	3151	23.55%	1660	27.23%	842	35.87%	414	45.41%

Table 6 shows that the high-risk group was the smallest in each time period, and had the highest consistent FIB-4 prevalence across all visits. The low-risk state was the most prevalent latent state for each time period across all visits. There was a good differentiation between all three risk levels. The number of individuals defined as high-risk decreased from visit to visit, starting at 14.5% in the second visit and

ending with 8.4% in the fifth visit and 9.7% in the sixth visit. Compared to the HMM results (Table 3), the risk level for fibrosis was similar to the HMM for the high-risk group, but the number of individuals in the high-risk group here was smaller (Table 6). The other levels were more similar (in the HMM and GBTM) in their overall fibrosis risk as well as the number of individuals in each group.

Table 7. Transitions of Individuals across Visits (Ben-Assuli et al., 2022)

	Visit 2 to Visit 3	Visit 3 to Visit 4	Visit 4 to Visit 5	Visit 5 to Visit 6
Stayed the same	88.99% (2804)	91.39% (1517)	93.47% (787)	94.20% (390)
Transitioned to low	4.00% (126)	1.20% (20)	6.53% (55)	0
Transitioned to high	7.01% (221)	7.41% (123)	0	5.80% (24)
Total	3151	1660	842	414

Table 7 displays that most patients remained in the same state across all visits in increasing percentages from each visit to the subsequent one (89%, 91%, 93%, 94%) in a very similar pattern to HMM (Table 4). Still, unlike the HMM, these trends started from the first visit.

As shown in (Ben-Assuli et al., 2022), GBTM and HMM had the highest values for the high-risk group for all demographic characteristics. The proportion of males increased with the rise in the risk level of fibrosis. The significant diagnosis characteristics were hypertension and diabetes, similar to the HMM results; both had the highest values for the high-risk trajectory. The blood test, including AST, Total Cholesterol, Glucose, and Platelets, were significantly different and quite similar to the HMM findings.

4.4 Prediction Results Based on Latent Class Modeling

The results of the classification models show the significant contribution of the risk states (obtained from HMM or GBTM), that significantly (DeLong, DeLong, & Clarke-Pearson, 1988) improved the individual-level prediction of FIB-4. For example, in the CHAID model, the AUC without Clustering was 0.665 as compared to 0.803 for the GBTM risk group identifiers (0.717 for the HMM risk group). In the XGBoost model, the AUC without Clustering was 0.73 compared to 0.819 in the GBTM risk group identifiers (0.77 for the HMM risk group). There were higher results across all evaluation indices when adding the risk state identifiers to the predictors. Improvements in the AUC were significant based on DeLong et al. (1988) at $p<0.001$ for all comparisons. These results illustrate the contribution of latent class modeling as a pre-stage for probable disease classification-based predictions.

4.5 Robustness Check

As in many other works, we faced a lopsided distribution of our main dependent variables (FIB-4). To account for the distribution of the dependent variable, we applied the Propensity Score Matching (PSM) econometric procedure (Rosenbaum & Rubin, 1983). PSM has rarely been used on a NAFLD population. We applied the PSM econometric procedure to account for the distribution of the dependent variable. Afterwards, we repeated the whole analytic process for both the HMM and the GBTM and obtained results that supported and echoed our main results(Ben-Assuli et al., 2022).

5. Discussion and Conclusion

In today's era of precision medicine, the role of data science applied to large volumes of healthcare data is destined to grow exponentially. Interventions to improve patient uptake of chronic disease management services can substantially reduce medical complications. Unlike previous works that have employed classification (e.g. feature variables, symptoms, diagnoses and laboratory tests) or clustering (grouping a set of patient data with inner group similarity and external heterogeneity) here these techniques were combined to track the same patients over multiple visits. There is substantial value in finding novel risk factors for NAFLD and more specifically for fibrosis progression. This was a prospective study with repeated measures, with a relatively long-term follow-up in contrast to many studies that have been cross-sectional and lack the kinetics of disease progression and risk factors over time.

Our results underscore the need to integrate stratification models over time to assess progressive disease risk by leveraging longitudinal healthcare delivery procedures.

We used HMM and GBTM independently to develop a longitudinal time series for fibrosis risk. We described them along with their technical specifications adjusted to our specific NAFLD context. We verified the goodness of fit measurements for these methods to assess several HMM states/GBTM trajectories by employing all known measurements. In addition, we controlled for all visits and time points (compared to a previous study that only ran the models on the last visit).

6. Policy Implications and Recommendations

The findings can help in planning future preventive strategies based on lifestyle modifications and medical treatment which can be implemented to reduce the NAFLD disease burden.

The implementation of these methods contributes by promoting the functional integration of machine learning techniques at various points of care. These methods can improve and promote the field of data science. They can supply forms of feedback and control that can provide the developers of expert systems with an additional set of key success factors.

This study contributes to activity and policy decision making in the healthcare sector. The results can assist policy decision makers in Israel in terms of the adoption of point of care recommendation systems.

Specifically, this study can help determine policy regarding early detection of NAFLD and/or advanced fibrosis in primary care settings for individuals with various medical problems (e.g. obesity or metabolic syndrome or in the setting of diabetes clinics), which represent high-risk populations for the more advanced forms of NAFLD. This study enables further refinement of high-risk groups, by identifying numerous parameters and the interaction between then and detection of potential low cost markers available in a regular clinical setting (stored in EMRs and EHRs). In terms of public health, primary prevention policies and actions can be initiated in patients defined as high risk according to the prediction models, and secondary prevention can be implemented by early detection. The advantage of such tools is that they are cost-free and performed with readily available measurements, and thus can be automatically calculated in computerized medical records (EMRs and EHRs) and increase the awareness of general practitioners and the public to this important and prevalent liver disease and promote early treatment. Even though most of the treatment focuses on lifestyle modification, awareness of liver disease can increase motivation among patients and greater monitoring on the part of the physicians and dietitians following these patients. This increased monitoring may also help in early detection of compensated cirrhosis and hepatocellular carcinoma. In addition, new medications are on the way, and low cost readily available prediction models can be used to refer suitable patients for treatment.

The findings can also contribute to medical technology designers (by promoting the algorithms for the design of new medical recommendation systems and improving existing ones), physicians (by enhancing the efficient use of predictive analytics),

patients (by improving healthcare services) and mainly policy makers in the healthcare sector (the advisability of investing in these advanced decision support systems).

7. References

Alberti, K., Eckel, R. H., Grundy, S. M., Zimmet, P. Z., Cleeman, J. I., Donato, K. A., Fruchart, J.-C., James, W. P. T., Loria, C. M., & Smith Jr, S. C. (2009). Harmonizing the metabolic syndrome: a joint interim statement of the international diabetes federation task force on epidemiology and prevention; national heart, lung, and blood institute; American heart association; world heart federation; international atherosclerosis society; and international association for the study of obesity. *Circulation* ,(16)120 , .1640-1645

Anstee, Q. M., Lawitz, E. J., Alkhouri, N., Wong, V. W., Romero-Gomez, M., Okanoue, T., . . . Younossi, Z. M. (2019). Noninvasive Tests Accurately Identify Advanced Fibrosis due to NASH: Baseline Data From the STELLAR Trials. *Hepatology, 70*(5), 1521-1530. doi:10.1002/hep.30842

Bardhan, I. R., Oh, J. H., Zheng, Z., & Kirksey, K. (2014). Predictive Analytics for Readmission of Patients with Congestive Heart Failure. *Information Systems Research, 26*(1), 19-39 .

Bedogni, G., Bellentani ,S., Miglioli, L., Masutti, F., Passalacqua, M., Castiglione, A., & Tiribelli, C. (2006). The Fatty Liver Index: a simple and accurate predictor of hepatic steatosis in the general population. *BMC Gastroenterol, 6*, 33. doi:1471-230X-6-33 [pii]

10.1186/1471-230X-6-33

Bedogni, G., Bellentani, S., Miglioli, L., Masutti, F., Passalacqua, M., Castiglione, A., & Tiribelli, C. (2006). The Fatty Liver Index: a simple and accurate predictor of hepatic steatosis in the general population. *BMC gastroenterology, 6* .33 ,(1)

Ben-Assuli, O., Jacobi, A., Goldman, O., Shenhar-Tsarfaty, S., Rogowski, O., Zeltser, D., Shapira, I., Berliner, S., & Zelber-Sagi, S. (2022). Stratifying Individuals into Non-Alcoholic Fatty Liver Disease Risk Levels using Time Series Machine Learning Models. *Journal of biomedical informatics*, 103986: 103981-103916. doi:https://doi.org/10.1016/j.jbi.2022.103986

Ben-Assuli, O., & Padman, R. (2020). Trajectories of Repeated Readmissions of Chronic Disease Patients: Risk Stratification, Profiling, and Prediction. *MIS quarterly, 44*(1), 201-226 .

Burckhardt, P., Nagin, D. S., & Padman, R. (2016). *Multi-Trajectory Models of Chronic Kidney Disease Progression.* Paper presented at the AMIA Annual Symposium Proceedings.

Cuthbertson, D. J., Weickert, M. O., Lythgoe, D., Sprung, V. S., Dobson, R., Shoajee-Moradie, F., Umpleby, M., Pfeiffer, A. F., Thomas, E. L., & Bell, J. D. (2014). External validation of the fatty liver index and lipid accumulation product indices, using 1H-magnetic resonance spectroscopy, to identify hepatic steatosis in healthy controls and obese, insulin-resistant individuals. *European Journal of Endocrinology, 171*(5), 561-569 .

DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, 837-845 .

Ekstedt, M., Hagstrom, H., Nasr, P., Fredrikson, M., Stal, P., Kechagias, S., & Hultcrantz, R. (2015). Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up. *Hepatology, 61*(5), 1547-1554. doi:10.1002/hep.27368

European Association for Study of, L., & Asociacion Latinoamericana para el Estudio del, H. (2015). EASL-ALEH Clinical Practice Guidelines: Non-invasive tests for evaluation of liver disease severity and prognosis. *J Hepatol, 63*(1), 237-264. doi:10.1016/j.jhep.2015.04.006

European Association for the Study of the, L., European Association for the Study of, D., & European Association for the Study of ,O. (2016). EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. *J Hepatol, 64*(6), 1388-1402. doi:10.1016/j.jhep.2015.11.004

Goldman, O., Ben-Assuli, O., Rogowski, O., Zeltser, D., Shapira, I., Berliner, S ,.Zelber-Sagi, S., & Shenhar-Tsarfaty, S. (2021). Non-alcoholic fatty liver and liver fibrosis predictive analytics: risk prediction and machine learning techniques for improved preventive medicine. *Journal of medical systems, 45*(2), 1-12 .

Gorunescu, F & ,.Belciug, S. (2014). Evolutionary strategy to develop learning-based decision systems. Application to breast cancer and liver fibrosis stadialization. *Journal of biomedical informatics, 49*, 112-118 .

Hagstrom, H., Nasr, P., Ekstedt, M., Hammar, U., Stal ,P., Hultcrantz, R., & Kechagias, S. (2017). Fibrosis stage but not NASH predicts mortality and time to development of severe liver disease in biopsy-proven NAFLD. *J Hepatol, 67*(6), 1265-1273. doi:10.1016/j.jhep.2017.07.027

Hillestad, R., Bigelow, J., Bower, A., Girosi, F., Meili, R., Scoville, R., & Taylor, R. (2005). Can electronic medical record systems transform health care? Potential health benefits, savings, and costs. *Health Affairs, 24*(5), 1103-1117 .

Kass, G. V. (1980). An exploratory technique for investigating large quantities of categorical data. *Journal of the Royal Statistical Society: Series C (Applied Statistics), 29*(2), 119-127 .

Kim, B. K., Park, J. Y., Ahn, S. H., Chon, C. Y., Kim, J. K., Paik, Y. H., Lee, K. S., Park, Y. N., & Han, K .H. (2010). Validation of FIB-4 and comparison with other simple noninvasive indices for predicting liver fibrosis and cirrhosis in hepatitis B virus-infected patients. *Liver International, 30*(4), 546-553 .

Koehler, E. M., Schouten, J. N., Hansen, B. E., Hofman, A., Stricker, B. H., & Janssen, H. L. (2013). External validation of the fatty liver index for identifying nonalcoholic fatty liver disease in a population-based study. *Clinical Gastroenterology and Hepatology, 11*(9), 1201-1204 .

Krumholz, H. M., Merrill, A. R., Schone, E. M., Schreiner, G. C., Chen, J., Bradley, E. H., Wang, Y., Wang, Y., Lin, Z., & Straube, B. M. (2009). Patterns of hospital performance in acute myocardial infarction and heart failure 30-day mortality and readmission. *Circulation* :*Cardiovascular Quality and Outcomes, 2*(5), 407-413 .

Lee, J.-H., Kim, D., Kim, H. J., Lee, C.-H., Yang, J. I., Kim, W., Kim, Y. J., Yoon, J.-H., Cho, S.-H., & Sung, M.-W. (2010). Hepatic steatosis index: A simple screening tool reflecting nonalcoholic fatty liver disease. *Digestive and Liver Disease, 7*(42), 503-508 .

Loomba, R., & Sanyal, A. J. (2013). The global NAFLD epidemic. *Nat Rev Gastroenterol Hepatol, 10*(11), 686-690. doi:10.1038/nrgastro.2013.171

Mitchell, S. E., Martin, J., Holmes, S., van Deusen Lukas, C., Cancino, R., Paasche-Orlow, M., Brach, C., & Jack, B. (2016). How Hospitals Reengineer Their Discharge Processes to Reduce Readmissions. *J Healthc Qual., 38*(2), 116-126 .

Modeler, I. S., & Guide, A. (2010). IBM Corporation. *NY, USA* .

Nagin ,D. S. (2005). *Group-based modeling of development*: Boston, MA: Harvard University Press.

Nagin, D. S., & Odgers, C. L. (2010). Group-based trajectory modeling in clinical research. *Annual Review of Clinical Psychology, 6*, 109-138 .

Netzer, O., Lattin, J .M., & Srinivasan, V. (2008). A hidden Markov model of customer relationship dynamics. *Marketing science, 27*(2), 185-204 .

Padman, R., Nagin, D., & Xie, Q. (2014). *Disease Progression and Risk Prediction for Chronic Kidney Disease: Analysis of Electronic Health Record Data using Group-Based Trajectory Models.* Paper presented at the Proceedings of Workshop on Information Systems & Technologies (WITS), Auckland, NZ.

Rabiner, L., & Juang, B. (1986). An introduction to hidden Markov models. *Ieee Assp Magazine* .4-16 ,(1)3 ,

Romero-Gomez, M., Zelber-Sagi, S., & Trenell, M. (2017). Treatment of NAFLD with diet, physical activity and exercise. *J Hepatol, 67*(4), 829-846. doi:10.1016/j.jhep.2017.05.016

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41-55 .

Ruhl, C., & Everhart, J. (2015). Fatty liver indices in the multiethnic United States National Health and Nutrition Examination Survey. *Alimentary pharmacology & therapeutics, 41*(1), 65-76 .

Semenov, A., Romov, P., Korolev, S., Yashkov, D., & Neklyudov, K. (2016). *Performance of machine learning algorithms in predicting game outcome from drafts in Dota 2.* Paper presented at the International Conference on Analysis of Images, Social Networks and Texts.

Shah, A. G., Lydecker, A., Murray, K., Tetri, B. N., Contos, M. J., Sanyal, A. J., & Network, N. C. R. (2009). Comparison of noninvasive markers of fibrosis in patients with nonalcoholic fatty liver disease. *Clinical Gastroenterology and Hepatology, 7*(10), 1104-1112 .

So-Armah, K. A., Lim, J. K., Re III, V. L., Tate, J. P., Chang, C.-C. H., Butt, A. A., Gibert, C. L., Rimland, D., Marconi, V. C., & Goetz, M. B. (2020). FIB-4 stage of liver fibrosis is associated with incident heart

failure with preserved, but not reduced, ejection fraction among people with and without HIV or hepatitis C. *Progress in cardiovascular diseases, 63*(2), 184-191 .

Sommerfeld, A. J., Althouse, A. D., Prince, J., & Hickey, G. W. (2016). Obstructive Sleep Apnea is Associated with Increased Readmissions in CHF Patients. *Journal of Cardiac Failure, 22*(8), S88. doi:10.1016/j.cardfail.2016.06.282

Stefan, N., Kantartzis, K., & Haring, H. U. (2008). Causes and metabolic consequences of Fatty liver. *Endocr Rev, 29*(7), 939-960. doi:er.2008-0009

Sterling, R. K., Lissen, E., Clumeck, N., Sola, R., Correa, M. C., Montaner, J., Sulkowski, M. S., Torriani, F. J., Dieterich, D. T., & Thomas, D. L. (2006). Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology, 43*(6), 1317-1325 .

Tapper, E. B., & Lok, A. S.-F. (2017). Use of liver imaging and biopsy in clinical practice. *New England Journal of Medicine, 377*(8), 756 .768-

Tsugawa, Y., Jena, A. B., Figueroa, J. F., Orav, E. J., Blumenthal, D. M., & Jha, A. K. (2017). Comparison of hospital mortality and readmission rates for Medicare patients treated by male vs female physicians. *JAMA internal medicine, 177*(2), 206-21 .3

van Walraven, C., Dhalla, I. A., Bell, C., Etchells, E., Stiell, I. G., Zarnke, K., Austin, P. C., & Forster, A. J. (2010). Derivation and validation of an index to predict early death or unplanned readmission after discharge from hospital to the community. *Canadian Medical Association Journal, 182*(6), 551-557 .

Wang, H., Robinson, R. D., Johnson, C., Zenarosa, N. R., Jayswal, R. D., Keithley, J., & Delaney, K. A. (2014). Using the LACE index to predict hospital readmissions in congestive heart failure patients. *BMC Cardiovascular Disorders, 14*(1), 97-104. doi:10.1186/1471-2261-14-97

Younossi, Z. M., Koenig, A. B., Abdelatif, D., Fazel, Y., Henry, L., & Wymer, M. (2016). Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. *Hepatology, 64*(1), 73-84. doi:10.1002/hep.28431

Zelber-Sagi, S., Lotan, R., Shlomai, A., Webb, M., Harrari, G., Buch, A., Nitzan Kaluski, D., Halpern, Z., & Oren, R. (2012). Predictors for incidence and remission of NAFLD in the general population during a seven-year prospective follow-up. *J Hepatol, 56*(5), 1145-1151. doi:10.1016/j.jhep.2011.12.011

Zelber-Sagi, S., Nitzan-Kaluski, D., Halpern, Z., & Oren, R. (2006). Prevalence of primary non-alcoholic fatty liver disease in a population-based study and its association with biochemical and anthropometric measures. *Liver Int, 26*(7), 856-863. doi:10.1111/j.1478-3231.2006.01311.x

Zelber-Sagi, S., Webb, M., Assy, N., Blendis, L., Yeshua, H., Leshno, M., Ratziu, V., Halpern, Z., Oren, R., & Santo, E. (2013). Comparison of fatty liver index with noninvasive methods for steatosis detection and quantification. *World journal of gastroenterology: WJG, 19*(1), 57 .

Zhang, Y., & Padman, R. (2015). Innovations in chronic care delivery using data-driven clinical pathways. *The American journal of managed care, 21*(12), e661-668 .

Zhuang, Y., Ding, H., Zhang, Y., Sun, H., Xu, C., & Wang, W. (2017). Two-dimensional shear-wave elastography performance in the noninvasive evaluation of liver fibrosis in patients with chronic hepatitis B: comparison with serum fibrosis indexes. *Radiology, 283*(3), 873-882 .