Nonalcoholic Fatty Liver Disease Incidence and Impact on Metabolic Burden and Death: a 20 Year-Community Study

Alina M. Allen, MD[1], Terry M. Therneau, PhD[2], Joseph J. Larson[2], Alexandra Coward[1], Virend K. Somers, MD, PhD[3], Patrick S. Kamath, MD[1]

[1]Department of Gastroenterology and Hepatology, Mayo Clinic, Rochester, Minnesota

[2]Department of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, Minnesota

[3]Department of Cardiovascular Diseases, Mayo Clinic, Rochester, Minnesota

Correspondence: Alina M. Allen, MD

Assistant Professor of Medicine

Department of Gastroenterology and Hepatology

Mayo Clinic

200 First Street SW, Rochester, MN

Email: allen.alina@mayo.edu

Phone: 507-284-3917

Fax: 507-284-0538

Abbreviations: BMI: body mass index; CI: confidence interval; CV: cardiovascular; DM: diabetes mellitus; HICDA: Hospital International Classification of Diseases Adapted codes; HTN: hypertension; ICD-9: International Classification of Diseases-9; MC: metabolic comorbidities; NHANES: National Health and Nutrition Examination Survey; NAFLD: nonalcoholic fatty liver disease; REP: Rochester Epidemiology Project; RR: relative risk; SD: standard deviation.

ABSTRACT

Recent population-based data on nonalcoholic fatty liver disease (NAFLD) epidemiology in general and incidence in particular, are lacking. We examined trends in NAFLD incidence in a US community, and the impact of NAFLD on incident metabolic comorbidities (MC), cardiovascular (CV) events and mortality. A community cohort of all adults diagnosed with NAFLD in Olmsted County, MN between 1997-2014 was constructed using the Rochester Epidemiology Project database. The yearly incidence rate was calculated. The impact of NAFLD on incident MC, CV events and mortality was studied using a multi-state model, with a 4:1 age and sex-matched general population as reference. We identified 3,869 NAFLD subjects (median age 53, 52% women) and 15,209 controls; median follow-up was 7 (1 to 20) years. NAFLD incidence increased 5-fold, from 62 to 329/100,000 person-years. The increase was highest (7-fold) in young adults, age 18-39 years. The 10-year mortality was higher in NAFLD subjects (10.2%) than controls (7.6%) (p<0.0001). NAFLD was an independent risk factor for incident MC and death. Mortality risk decreased as the number of incident MC increased: RR= 2.16 (95% CI 1.41-3.31), 1.99 (95% CI 1.48-2.66), 1.75 (95% CI 1.42-2.14) and 1.08 (95% CI 0.89-1.30) when 0, 1, 2, or 3 MC were present, respectively. The NAFLD impact on CV events was significant only in subjects without MC (RR=1.96, 95% CI=1.35-2.86). NAFLD reduced life expectancy by 4 years, with more time spent in high metabolic burden. *Conclusion:* The incidence of NAFLD diagnosis in the community has increased 5-fold, particularly in young adults. NAFLD is a consequence but also a precursor of MC. Incident MC attenuates the impact of NAFLD on death and annuls its impact on CV disease.

The unrelenting challenge of obesity has resulted in nonalcoholic fatty liver disease (NAFLD) becoming the most common cause of chronic liver disease in the industrialized countries. Based on extensive previous data, including a recent metaanalysis by Younossi et al.(1), the estimated prevalence of NAFLD in the United States is approximately 24%. The vast majority of large population-based studies assessing the burden of NAFLD are derived from the National Health and Nutrition Examination Survey (NHANES) III between 1988 and 1994, when obesity rates were estimated at 22.9%(2) of the population. However the prevalence of obesity has continued to increase steeply beyond this timeframe in all representative age and sex groups. Based on the most recent NHANES estimates from 2013-2014, 35% of men, 40% of women and 17% of children and adolescents are obese(3). In this context, the burden of obesity-related comorbidities, including NAFLD and metabolic syndrome, is expected to increase. However, robust epidemiologic data from recent years are lacking.

Previous studies of the natural history of NAFLD, particularly with regard to cardiovascular outcomes and mortality, have been limited by lack of large cohorts or sufficient longitudinal follow-up. Moreover, case definition based on abnormal aminotransferase levels or histological findings may introduce misclassification or spectrum bias, respectively, because the majority of NAFLD subjects have normal liver enzymes and liver biopsy is generally obtained only in subjects with more marked features of disease. Additionally, studies from certain restricted populations, such as those with diabetes mellitus or from referral centers, cannot be widely generalized. Thus, it remains unclear whether NAFLD is an independent risk factor for mortality beyond its association with obesity and metabolic syndrome.

This study aims to address these limitations by using the Rochester Epidemiology Project (REP) database which includes all the medical encounters and diagnoses of a large population-based cohort, with considerable longitudinal follow-up. The case ascertainment algorithm incorporates multiple diagnostic criteria used by medical providers, including blood tests, imaging, histology

4

and exclusion of alternative liver diseases, which ultimately led to a physician-derived diagnosis of NAFLD. Case identification was based on diagnosis codes that were consistent over time. We aimed to describe epidemiologic trends in incidence of NAFLD diagnosis, burden of metabolic comorbidities and cardiovascular events in Olmsted County in the last two decades. Additionally, we aimed to determine the role of NAFLD in incident metabolic comorbidities, such as diabetes mellitus (DM), hypertension (HTN) and hyperlipidemia, and its impact on mortality based on these incident comorbidities. These trends are critical to understanding the evolution of the burden of NAFLD in the community and the consequences for the medical and public health community, so as to allow inferences on the future health trajectory of the population.

METHODS

Data source

Population-based epidemiological research can be conducted in Olmsted County, Minnesota, because medical care is effectively self-contained within the community and there are only a few providers. As of 2010, the Olmsted County population included 144,428 people, of whom 85.7% were white and 48.9% were male. Currently, the medical providers include Mayo Clinic and its affiliated hospitals; Olmsted Medical Center and its affiliated hospital; and the Rochester Family Clinic, a private practitioner in the area. All medical information from each of these providers for each individual resident of Olmsted County is linked to a single identification number allowing linkage across health-care systems. The diagnoses recorded in billing data from each institution are indexed so that all individuals with any specific diagnosis can be identified across systems and the corresponding patient records can easily be retrieved from each site.

The system and infrastructure that collects, collates and indexes the data is the Rochester Epidemiology Project (REP)(4, 5), which has been funded by the National Institutes of Health since 1966. The data have been used in epidemiologic studies of a range of conditions including deep vein thrombosis, pulmonary embolism and epilepsy. Data from REP has also been used as a sampling frame for population-based studies, and as a source of common controls to allow examination of risk factors for various health conditions.

Patient selection and case definition

Using the REP, we constructed a community cohort of all adult subjects with NAFLD in Olmsted County, MN from 1997 to 2014. NAFLD cases were identified using Hospital International Classification of Diseases Adapted (HICDA) codes, a system developed at Mayo for research diagnosis coding and adapted by REP in 1976: HICDA 05710421 (fatty liver), 05710431

(nonalcoholic steatohepatitis - NASH). Additionally, the International Classification of Diseases (ICD)-9 codes ICD 9-CM 571.5 (cirrhosis of the liver without mention of alcohol), 571.8 (other chronic nonalcoholic liver disease), 571.9 (unspecified chronic liver disease without mention of alcohol) were used. From this cohort, we excluded subjects diagnosed with other liver diseases with specific diagnosis codes, such as viral hepatitis, alcoholic liver disease, alcohol use, cholestatic liver disease, etc (identified by the HICDA and ICD 9-CM codes listed in **Supplementary Table 1**). Subjects were ascertained as NAFLD cases if no alternative liver disease was identified prior to the index NAFLD diagnosis or during the follow-up. To avoid bias in long-term outcomes in this longitudinal study, subjects with short follow-up time of less than 1 year were excluded.

The remaining subjects were included in the final NAFLD cohort and followed longitudinally until death, last follow-up or the study end date (October 1st 2016). Individual chart review of 500 unique records (400 of the final NAFLD cohort and 100 of the excluded cases), including outpatient and hospital notes, images, laboratory studies, and histologic data, showed that the accuracy of case identification for NAFLD or other liver disease used in the exclusion criteria was 85%. Of 400 NAFLD cases identified by the code-based algorithm, 85% were true NAFLD by chart review, while 15% were alternative liver diseases. Of 100 excluded cases based on the code-based algorithm, 87% were true non-NAFLD liver disease and 13% were in fact NAFLD by chart review.

Cirrhosis was defined by ICD 9-CM 571.5 and FIB4 score >2.67 (6). High FIB4 value without a diagnostic code had low positive predictive value (data not shown). Comorbidities were defined based on combinations ICD 9-CM or HICDA codes (Supplementary Table 2), medications (Supplementary Table 3) and laboratory values, as follows: diabetes mellitus – diagnostic codes plus medications or laboratory values (fasting glucose ≥ 126 mg/dL or hemoglobin A1c ≥6.5%);

dyslipidemia – diagnostic codes plus medications or laboratory values (LDL cholesterol >100 mg/dL or triglycerides >150 mg/dL); hypertension – diagnostic codes plus medications.

Data Analysis

The crude yearly incidence rate of NAFLD in the community was calculated by the number of new NAFLD cases divided by the sum of "person-time" population at risk in Olmsted County, represented by the average adult population size multiplied by the duration of follow-up.

The impact of NAFLD on development of incident dysmetabolic comorbidities, such as DM, HTN or dyslipidemia, and death was examined using a 1:4 control population matched for age and sex as reference. These metabolic comorbidities were identified using HICDA and ICD 9-CM codes. Because NAFLD and metabolic syndrome have intertwined pathophysiology, the independent impact of NAFLD on death may vary with the number of dysmetabolic comorbidities developed as the subject progresses through the course of the disease. Since intermediate events change the natural history of the disease progression, the prognostic role of NAFLD may not be the same after development of DM, or after development of DM and HTN, respectively. Traditional regression analyses such as the Cox proportional hazards model, most often used to study the impact of NAFLD on incident metabolic syndrome or death, ignore these intermediate steps and yield a single "average" relative risk associated with NAFLD. Thus, to determine the impact of NAFLD on incident metabolic comorbidities and death based on each subject's current burden of the metabolic comorbidities, we used a multistate modeling approach(7, 8). In this 5 state model (**Supplementary Figure 1**), each NAFLD case and matched controls transition unidirectionally from a state of zero comorbidities (state 1) to a state of one (state 2), two (state 3) and three (state 4) comorbidities; transitions to a state of death (state 5) can occur from each of the other four states. Each subject starts in the state according to the combination of diagnoses identified in the REP database at the time of NAFLD diagnosis.

Controls enter at the same age as their matched case. For example, a NAFLD case with

coexistent hypertension at the time of NAFLD diagnosis starts in state 2, as does a control with

hypertension. These subjects transition to state 3 once they develop a new comorbidity (DM or

dyslipidemia). A window of minimum 7 days between new comorbidities or events was required

in order to minimize the risk of capturing simultaneous conditions as sequential while still

allowing robust longitudinal ascertainment of events. The model estimates the impact of

NAFLD on the rate of transition between states and the predicted mean time in each state, all

adjusted for age and sex.

Last, using the multi-state model we analyzed the impact of NAFLD on clinical cardiovascular

events (myocardial infarction, angina/ischemic heart disease, atrial fibrillation, cardiac arrest,

congestive heart failure and stroke). Statistical analyses were performed in SAS v9.4 (SAS

Institute; Cary, NC) and R statistical software, version 3.2.0 (R Foundation for Statistical

Computing, Vienna). The study was approved by the Institutional Review Boards of Mayo Clinic

and Olmsted Medical Center.

RESULTS

Identification of NAFLD cases

Between 1997 and 2014, a total of 5,427 subjects were identified using the inclusion codes (**Figure 1**). Of these, 1,558 were excluded based on identification of other diagnostic codes of alternative liver diseases (n= 1,222), follow-up time of less than 1 year (n= 327), lack of research authorization (n= 6) or inconclusive Olmsted County residency status at the time of diagnosis (n=3). Thus, the final NAFLD cohort comprised 3,869 subjects. Among these, 178 were diagnosed with cirrhosis during the study period.

Median age at diagnosis was 53 (IQR 42-63) years and 2,032 (52%) were women. Of these, 813 (21%) were diagnosed at age 18-39, 1,857 (48%) were diagnosed at age 40-59, and 1,199 (31%) were diagnosed at ≥60 years of age. The median follow-up was 7 (range 1 to 20) years.

Increasing NAFLD incidence and obesity rates in Olmsted County

The overall age and sex-adjusted incidence of NAFLD increased 5-fold, from 62/100,000 person-years in 1997 to 329/100,000 person-years in 2014 (**Figure 2A**). The magnitude and pattern of increase was similar for men (49/100,000 person years to 317/100,000 person years) and women (72/100,000 person years to 341/100,000 person years). **Figure 2B** illustrates the trends in NAFLD incidence stratified by age groups. Between 1997 and 2014 NAFLD incidence increased 7-fold among subjects 18-39 years old (from 20 to 140/100,000 person year), 6-fold among subjects 40-59 years old (from 70 to 407/100,000 person year) and 4-fold among subjects over 60 years old (from 125 to 515/100,000 person year).

To determine if the incidence trends were explained by true increases in incident cases versus increased provider awareness and imaging, we examined the temporal associations between rates of NAFLD and frequency of abdominal ultrasound testing in all Olmsted County residents during the study time frame. The age and sex-adjusted incidence of abdominal ultrasound

testing was 1,967/100,000 person years in 1997 and 1,959/100,000 person years in 2014. **Supplementary Figure 2** illustrates the increasing NAFLD incidence despite stable trends in abdominal ultrasound testing in Olmsted County residents.

To investigate if the increasing incidence of NAFLD corresponded to an increasing trend in obesity in the community, all height and weight values were extracted and the body mass index (BMI) was calculated in NAFLD subjects and a 4:1 age and sex-matched cohort (15,209 subjects) from the general population. Height and weight values were collected by medical personnel at the time of a medical visit. BMI data within 1 year from the time of NAFLD diagnosis or matching was available in 3,250 (84%) NAFLD subjects and 10,038 (66%) controls, respectively. The prevalence of obesity (BMI ≥ 30 kg/m^2) in the community increased from 28% in 1997 to 37% in 2014. The prevalence of obesity among NAFLD subjects remained constant, at 68% during the study period. **Supplementary Figure 3** shows that BMI in the general population has increased from 27.8 to 28.9 kg/m^2. Similarly, the BMI in NAFLD subjects at the time of diagnosis increased from 32.6 kg/m^{2}in 1997 to 34.2 kg/m^2 in 2008 when it stabilized.

Increasing burden of metabolic comorbidities present at the time of NAFLD diagnosis

The mean (± SD) age at the time of NAFLD diagnosis did not change significantly during the study period: 55 (± 16) years in 1997 versus 54 (±14) years in 2014. However, the proportion of NAFLD subjects with 2 or more dysmetabolic comorbidities present at the time of diagnosis increased over the study period from 21% to 53% (**Figure 3**). Thus the population-based trends in clinical- demographic characteristics at the time of NAFLD diagnosis over the last 2 decades are remarkable for higher BMI, more dysmetabolic burden, but similar mean age.

Impact of NAFLD on incident metabolic comorbidities and death

Baseline and incident metabolic comorbidities were compared between the 3,869 NAFLD subjects and a 4:1 age- and sex-matched cohort identified from the general population (15,209 subjects) (**Table 1**). Compared to controls, NAFLD cases had more dysmetabolic comorbid conditions at the time of diagnosis and after 10 years of follow-up. The number of subjects in each state at any given time, and the number of those who transition to higher states is shown in **Supplementary Table 4**. **Figure 4** shows the impact of NAFLD on incident metabolic comorbidities after NAFLD diagnosis and mortality in reference to age and sex-matched controls. In this multistate model, NAFLD subjects and their age- and sex- matched counterparts are followed as they transition unidirectionally towards death, directly or through intermediate states of progressive dysmetabolic burden. The dysmetabolic burden was quantified as one, two or three of the following comorbidities: diabetes, hypertension or dyslipidemia. Compared to controls, NAFLD subjects had a higher risk of developing incident dysmetabolic comorbidities. The impact of NAFLD is highest on the transition from zero to one metabolic comorbidity. For example, a subject with NAFLD without DM, HTN or dyslipidemia, was over twice (RR=2.62, 95% CI=2.31-2.96) more likely to develop one or more of these comorbidities than an age- and sex- matched control. A subject with NAFLD and either 1 of 3 comorbidities (DM, HTN or dyslipidemia), had a 1.67-fold higher risk (RR=1.67, 95% CI=1.51-1.85) of developing additional dysmetabolic comorbidities than an age- and sex- matched control with one comorbidity but without NAFLD. The impact of NAFLD was lower on the transition from a state of 2 to 3 metabolic comorbidities (RR=1.62, 95% CI=1.42-1.83). When stratified by cirrhosis status, the impact of NAFLD on incident metabolic comorbidities decreased with higher metabolic burden in a similar fashion: RR=2.60, 95% CI 2.30-2.95 to transition from 0 to 1+, RR= 1.66, 95% CI 1.49-1.84 to transition from 1 to 2+, and RR=1.53, 95% CI 1.34-1.74 to transition from 2 to 3 metabolic comorbidities.

The 10-year mortality was higher in NAFLD subjects (10.2%, n= 396) than controls (7.6%, n=1,155 (p<0.0001). We investigated the impact of NAFLD on mortality stratified by the presence and the severity of metabolic syndrome using the multi-state model described above. As opposed to traditional regression analysis, this model takes into consideration the presence and number of metabolic comorbidities not only at the time of NAFLD diagnosis, but also as they develop throughout the course of the disease. In this model, NAFLD subjects and matched controls transition towards death from any of the 4 dysmetabolic states: 0, 1, 2 or 3 metabolic comorbidities. As seen in **Figure 4**, NAFLD is an independent risk factor for death, but its impact decreases with the number of dysmetabolic conditions. A subject with NAFLD and no metabolic comorbidities has a 2-fold increased mortality risk compared to an age and sex-matched control with no metabolic comorbidities or NAFLD (RR=2.16, 95% CI=1.41-3.31). As expected, as the number of comorbidities increased, their respective attributable mortality risk diluted that of NAFLD alone: RR= 1.99 (95% CI 1.48-2.66), 1.75 (95% CI 1.42-2.14), when one or two comorbidities were present, respectively, and was insignificant in NAFLD subjects with three comorbidities: RR=1.08 (95% CI 0.89-1.30) .

When stratified by cirrhosis status to separate the effect of cirrhosis alone on mortality, the impact of NAFLD on death remained significant albeit lower, and decreased with the number of metabolic comorbidities in a similar fashion: RR=1.73 (95% CI 1.09-2.77), 1.56 (95% CI 1.12-2.16), 1.56 (95% CI 1.26-1.94), and 0.85 (95% CI 0.69-1.05) when 0, 1, 2 or 3 comorbidities were present, respectively.

Impact of NAFLD on time-in-state

Using the multistate model we determined the number of remaining years spent in each comorbid state (of 1,2 or 3 incident comorbidities), from an arbitrary age of 50 to death, in NAFLD subjects compared to age- and sex-matched controls from the general population.

Overall, the mean lifetime of non-cirrhotic NAFLD subjects is 4 years shorter than age- and sex-matched controls: female NAFLD vs female control: 32 vs 36 years; male NAFLD vs male control: 29 vs 33 years. The mean lifetime of NAFLD males is 3 years lower than NAFLD females. **Figure 5** illustrates that NAFLD subjects not only have shorter life expectancy, but also that most of their remaining time (75%) is spent in comorbid states (darker grey). This illustration is consistent with the model prediction that NAFLD is a risk factor for incident comorbidities and mortality.

Cardiovascular outcomes

At the index diagnosis, NAFLD subjects were more likely to have a history of cardiovascular (CV) events than the matched controls (28% vs 18%). At 10 years of follow-up 34% of NAFLD subjects and 22% of controls had CV events.

The impact of NAFLD on clinical CV events as subjects and controls develop incident metabolic comorbidities was estimated with a multi-state model, similar to that used for mortality assessment. The impact of NAFLD on CV events was highest in subjects without metabolic comorbidities, who had a 2-fold increased risk of CV events than an age and sex matched control with no metabolic comorbidities (RR=1.96, 95% CI=1.35-2.86, p<0.001). The clinical or statistical significance of NAFLD as a CV risk factor decreased or was no longer present as subjects and controls acquired metabolic conditions throughout their disease course and moved from higher states of metabolic burden: RR= 1.21 (95% CI 0.96-1.53, p=0.10), 1.24 (95% CI 1.05-1.47, p=0.01), and 1.02 (95% CI 0.86-1.21, p=0.76) when one, two or three comorbidities were present, respectively. When stratified by cirrhosis, the results were similar. **Table 2** summarizes the impact of NAFLD on death and clinical CV disease as estimated by the multi-state model.

DISCUSSION

Several important observations can be made in this large population study of 3,869 NAFLD subjects with 20 years of follow-up. First, the incidence of NAFLD diagnosis in the community has increased 5-fold among all age categories and at a similar rate in men and women, from 62/100,000 to 329/100,000 person years. The increase was disproportionally higher among young adults age 18-39 years, in whom the incidence of NAFLD increased 7-fold. Second, the increase in NAFLD diagnosis paralleled an increase in median BMI and obesity rates in the general population. Despite a shift towards a younger population, the clinical profile of the NAFLD patient has become more complex over the last 2 decades, due to the presence of more metabolic comorbidities at the time of NAFLD diagnosis. Last, NAFLD is a risk factor for incident dysmetabolic conditions and cardiovascular events. NAFLD is an independent risk factor for death, but impact decreases with the increase of incident metabolic comorbidities. The lifespan of NAFLD subjects is 4 years shorter than that of controls, and the majority of the remaining years are spent in states of metabolic comorbidities.

Whereas data on the estimated prevalence of NAFLD has been extensively published, there is a lack of data regarding the incidence of NAFLD in the community and the trends in incidence over time. The available data are mostly limited to Asian populations, where the incidence varied between 13.5% over 3-5 years in Hong Kong(9), 9.1% per year in southern China (10) and 19.0% over 7 years in Israel(11). Several criteria are required to accurately measure the incidence of a disease. These include a sufficient observation period of a well-defined population and the use of constant, reproducible measures of disease ascertainment. These criteria are seldom met, and therefore few studies are positioned to provide reliable incidence data. In this study, NAFLD was ascertained by disease coding, which implies that the subjects had imaging and/or laboratory evidence of disease during their medical evaluation. Thus, many Olmsted County residents who did not undergo the necessary testing remained undiagnosed,

leading to an underestimation of the true disease burden. Nonetheless, REP, a data-linkage resource that includes all medical records of the entire Olmsted County, offers a unique opportunity to gauge the trends in incident diagnosis of NAFLD in the population, which are otherwise difficult to estimate in large communities over a long period of time. It is possible that the improvement in ultrasound technology during the study period, with higher sensitivity to detect lower grades of hepatic steatosis, contributed to an overestimation of the magnitude of incidence rise, noted especially in the first half of the study period (1997-2007). Additionally, increased provider and radiologist awareness may lead to better screening and recognition of the disease. However the frequency of ultrasound use, as a surrogate of physician awareness, did not increase over the study period. Another possible cause for a spurious increase in incidence rates is a change in coding practices in the records-linkage system. However, we do not think that this was sizeable, because case ascertainment was based on an algorithm using the same HICDA and ICD-9 codes, which were introduced well in advance of the study start time.

In a previous Olmsted County analysis including 420 NAFLD subjects identified between the years 1980-2000, the adjusted incidence rate of NAFLD diagnosis was increasing, from an average of 4.2/100,000/year during 1980–1985 to 38.0/100,000/year during 1995–1999 (p <.001)(12). The current analysis shows that the increasing trend has continued over the following years, in parallel with increasing obesity rates in the community. In our large sample of 15,209 subjects from the community, the increasing trends in obesity rates from 28% to 37% are concordant with those reported by Centers of Disease Control based on National Health Examination Surveys(13),(14).

The steeper increase in NAFLD diagnosis in young adults compared to other age groups is alarming. NHANES analyses between 1988-1994 show that among young adolescents age 12 to 19 years, 10.5% were obese and 15.5% were overweight(15). The rates doubled through

2013-2014(3). Tracking adolescent obesity into adulthood during our study timeframe, it comes as no surprise that more young adults are diagnosed with NAFLD. Moreover, NAFLD is currently the most common liver abnormality in children age 2 to 19 years(16), with a prevalence as high as 38% in obese children. These data are concerning because longer duration of the disease increases the likelihood of progression to cirrhosis. The impact of NAFLD on liver-related mortality, morbidity, and healthcare burden could increase substantially in the upcoming years.

The clinical-demographic characteristics of the NAFLD patient have changed over the last two decades: despite a trend towards younger age at diagnosis, there is a steady increase in the prevalence and number of cardiovascular risk factors, such as diabetes mellitus, hypertension and hyperlipidemia. The proportion of patients with 2 or more metabolic comorbidities at the time of liver disease diagnosis has increased from 21% to 53%. Taken together, these findings suggest that the portrait of the liver transplant candidate is changing. As NASH is expected to become the leading cause of liver transplantation(17), the foreseeably growing proportion of transplant candidates with multiple comorbidities may raise many challenges for pretransplant selection and cardiovascular risk optimization in the perioperative period and long after liver transplantation.

Currently, there is insufficient evidence to change cardiovascular risk stratification in patients with NAFLD. The association between NAFLD and subclinical CV risk(18) and metabolic disease(19) has been demonstrated in several studies. However, whether NAFLD imparts an additive risk for hard clinical CV outcomes and mortality beyond established risk factors has been controversial (18, 20, 21, 22), partly because traditional regression analysis adjusts for significant covariates that are usually present at baseline. However, we have added important new information to the evidence that NAFLD is a complex and heterogeneous disease which can promote or accelerate the progression of CV risk factors(23-25). Thus, its impact on

mortality could vary throughout the natural course of the disease, a process which would not be accurately captured by a traditional regression model. Multistate models are particularly relevant for conditions and outcomes which have an event-related dependence, such as occurrence of an additional disease changing the risk of death associated with the disease of interest. In these progressive illness-death models (competing risks models are a sub-type of such multistate models), the transition intensities provide the hazards for movement from one state to another. Covariates may be incorporated in models through transition intensities to explain differences among individuals in the course of the illness.

The use of multistate models provides the opportunity to show that the impact of NAFLD on incident DM, HTN or dyslipidemia decreases as the individual progresses towards higher dysmetabolic burden. Similarly, the impact of NAFLD on death, measured by the transition from state 0 to death, state 1 to death, etc, decreases. In a higher dysmetabolic state, the influence of NAFLD on mortality is overtaken by the other metabolic conditions. NAFLD is a significant risk factor for clinic cardiovascular events only in the absence of metabolic syndrome. This modelling approach may bring out important biological insights which may be ignored when using an ordinary regression model, but may have important practical implications. Direct liver-related therapies may have a higher impact on hard outcomes such as mortality or cardiovascular disease if introduced before the onset of a high dysmetabolic burden. These preliminary observations represent important data for future prospective therapeutic trials. Meanwhile, multi-state modelling is a flexible tool for the study of NAFLD as an epidemiologic modifier, which offers additional information beyond that obtained by applying simpler survival models. Time-in-state, one of the additional benefits of using this modeling approach, shows that, even among the "healthier" subjects without metabolic comorbidities, NAFLD decreases the life expectancy by 4 years and most of the remaining years are spent in states of incident comorbidities. These outcomes have great repercussions for the healthcare burden of NAFLD,

as the management of incident DM, HTN and dyslipidemia involves outpatient visits and hospitalizations, which dominate most of the remaining lifetime of NAFLD subjects.

The strengths of this study consist of a large population based cohort with substantive longitudinal evaluation of outcomes up to 20 years, allowing sufficient time to detect epidemiologic trends and credible impact of NAFLD in a real-world setting. The disease ascertainment was based on system-linkage codes that resulted from a comprehensive medical evaluation of medical history, screening for excess alcohol consumption, laboratory, medications and imaging. We have not limited our inclusion criteria to biopsy-proven NAFLD to avoid bias towards a more aggressive disease spectrum. Similarly, we have not limited our cases to those identified based on abnormal liver enzymes because approximately 30% to 60% of patients with biopsy-confirmed nonalcoholic steatohepatitis have a normal ALT level (26, 27).

The use of clinical diagnosis codes for case ascertainment can limit disease specificity, especially in a diagnosis of exclusion such as NAFLD. In contrast to other ICD code-based studies of NAFLD, which used ICD codes as inclusion criteria only, we extended the case ascertainment algorithm by searching for an extensive list of diagnoses of alternative liver diseases, and excluding those subjects. Based on detailed chart review of clinical, laboratory and imaging parameters in a sizable random sample, this algorithm correctly identified true NAFLD cases with 85% accuracy. We believe that 15% is an acceptable proportion of misclassified cases given the limitations of individual chart review in large community based studies. Data on other cardiovascular risk factors such as smoking, genetic predisposition and family history was not available. Furthermore, the largely Caucasian Olmsted County population limits the generalizability to other races or ethnicities. In this study we did not aim to determine the cause of death in this NAFLD population, because cardiovascular-related deaths are already known to predominate, accounting for twice as many deaths as those that occur due to liver-related causes (12, 28). This work focused on the epidemiology of NAFLD and its

cardiometabolic impact; therefore liver-related outcomes were beyond the scope and constraints of this manuscript. These limitations notwithstanding, this study takes advantage of a unique population database to add important epidemiological data on NAFLD incidence and analyze its impact on outcomes using an innovative modeling technique.

The contemporary epidemiology of NAFLD is thus characterized by a substantial increase in incidence, with a shift in the burden of disease towards a younger population who presents with an increasingly large number of comorbid dysmetabolic conditions. These conditions will play a major role in the outcomes and health care utilization of the growing number of liver transplant candidates with complex cardiovascular barriers to life-saving liver transplantation. While the current understanding of these complex associations continues to evolve, general practitioners and medical subspecialists alike should be aware of the intricate relationships between NAFLD and incident metabolic diseases, as well as the dynamic impact of NAFLD on mortality which may have implications for the optimal timing for intervention. These findings represent a challenge to energize societal epidemiology to invest additional efforts towards awareness, early intervention, and optimal multidisciplinary management, which should ultimately translate into improved patient outcomes.

20

REFERENCES

1. Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M. Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. Hepatology 2016;64:73-84.

2. Flegal KM, Carroll MD, Ogden CL, Johnson CL. Prevalence and trends in obesity among US adults, 1999-2000. JAMA 2002;288:1723-1727.

3. Ogden CL, Carroll MD, Lawman HG, Fryar CD, Kruszon-Moran D, Kit BK, Flegal KM. Trends in Obesity Prevalence Among Children and Adolescents in the United States, 1988-1994 Through 2013-2014. JAMA 2016;315:2292-2299.

4. Rocca WA, Yawn BP, St Sauver JL, Grossardt BR, Melton LJ, 3rd. History of the Rochester Epidemiology Project: half a century of medical records linkage in a US population. Mayo Clinic proceedings 2012;87:1202-1213.

5. St Sauver JL, Grossardt BR, Yawn BP, Melton LJ, 3rd, Pankratz JJ, Brue SM, Rocca WA. Data resource profile: the Rochester Epidemiology Project (REP) medical records-linkage system. International journal of epidemiology 2012;41:1614-1624.

6. Shah AG1, Lydecker A, Murray K, Tetri BN, Contos MJ, Sanyal AJ. Comparison of noninvasive markers of fibrosis in patients with nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol. 2009 Oct;7(10):1104-12.

7. Andersen PK, Keiding N. Multi-state models for event history analysis. Statistical methods in medical research 2002;11:91-115.

8. Meira-Machado L, de Una-Alvarez J, Cadarso-Suarez C, Andersen PK. Multi-state models for the analysis of time-to-event data. Statistical methods in medical research 2009;18:195-222.

9.　　Wong VW, Wong GL, Yeung DK, Lau TK, Chan CK, Chim AM, Abrigo JM, et al. Incidence of non-alcoholic fatty liver disease in Hong Kong: a population study with paired proton-magnetic resonance spectroscopy. Journal of hepatology 2015;62:182-189.

10.　　Zhou YJ, Li YY, Nie YQ, Huang CM, Cao CY. Natural course of nonalcoholic fatty liver disease in southern China: a prospective cohort study. Journal of digestive diseases 2012;13:153-160.

11.　　Zelber-Sagi S, Salomone F, Yeshua H, Lotan R, Webb M, Halpern Z, Santo E, et al. Non-high-density lipoprotein cholesterol independently predicts new onset of non-alcoholic fatty liver disease. Liver international : official journal of the International Association for the Study of the Liver 2014;34:e128-135.

12.　　Adams LA, Lymp JF, St Sauver J, Sanderson SO, Lindor KD, Feldstein A, Angulo P. The natural history of nonalcoholic fatty liver disease: a population-based cohort study. Gastroenterology 2005;129:113-121.

13.　　Flegal KM, Carroll MD, Kuczmarski RJ, Johnson CL. Overweight and obesity in the United States: prevalence and trends, 1960-1994. International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity 1998;22:39-47.

14.　　Flegal KM, Kruszon-Moran D, Carroll MD, Fryar CD, Ogden CL. Trends in Obesity Among Adults in the United States, 2005 to 2014. JAMA 2016;315:2284-2291.

15.　　Ogden CL, Flegal KM, Carroll MD, Johnson CL. Prevalence and trends in overweight among US children and adolescents, 1999-2000. JAMA 2002;288:1728-1732.

16.　　Schwimmer JB, Deutsch R, Kahen T, Lavine JE, Stanley C, Behling C. Prevalence of fatty liver in children and adolescents. Pediatrics 2006;118:1388-1393.

17.　　Wong RJ, Aguilar M, Cheung R, Perumpail RB, Harrison SA, Younossi ZM, Ahmed A. Nonalcoholic steatohepatitis is the second leading etiology of liver disease among adults awaiting liver transplantation in the United States. Gastroenterology 2015;148:547-555.

18.   Mellinger JL, Pencina KM, Massaro JM, Hoffmann U, Seshadri S, Fox CS, O'Donnell CJ, et al. Hepatic steatosis and cardiovascular disease outcomes: An analysis of the Framingham Heart Study. Journal of hepatology 2015;63:470-476.

19.   Speliotes EK, Massaro JM, Hoffmann U, Vasan RS, Meigs JB, Sahani DV, Hirschhorn JN, et al. Fatty liver is associated with dyslipidemia and dysglycemia independent of visceral fat: the Framingham Heart Study. Hepatology 2010;51:1979-1987.

20.   Goessling W, Massaro JM, Vasan RS, D'Agostino RB, Sr., Ellison RC, Fox CS. Aminotransferase levels and 20-year risk of metabolic syndrome, diabetes, and cardiovascular disease. Gastroenterology 2008;135:1935-1944, 1944 e1931.

21.   Kim D, Kim WR, Kim HJ, Therneau TM. Association between noninvasive fibrosis markers and mortality among adults with nonalcoholic fatty liver disease in the United States. Hepatology 2013;57:1357-1365.

22.   Targher G, Byrne CD, Lonardo A, Zoppini G, Barbui C. Non-alcoholic fatty liver disease and risk of incident cardiovascular disease: A meta-analysis. Journal of hepatology 2016;65:589-600.

23.   Ma J, Hwang SJ, Pedley A, Massaro JM, Hoffmann U, Chung RT, Benjamin EJ, et al. Bi-directional analysis between fatty liver and cardiovascular disease risk factors. Journal of hepatology 2016.

24.   Fraser A, Harris R, Sattar N, Ebrahim S, Davey Smith G, Lawlor DA. Alanine aminotransferase, gamma-glutamyltransferase, and incident diabetes: the British Women's Heart and Health Study and meta-analysis. Diabetes care 2009;32:741-750.

25.   Sung KC, Wild SH, Byrne CD. Development of new fatty liver, or resolution of existing fatty liver, over five years of follow-up, and risk of incident hypertension. Journal of hepatology 2014;60:1040-1045.

26.   Mofrad P, Contos MJ, Haque M, Sargeant C, Fisher RA, Luketic VA, Sterling RK, et al.

Clinical and histologic spectrum of nonalcoholic fatty liver disease associated with normal ALT

values. Hepatology 2003;37:1286-1292.

27.   Amarapurkar DN, Patel ND. Clinical spectrum and natural history of non-alcoholic

steatohepatitis with normal alanine aminotransferase values. Tropical gastroenterology : official

journal of the Digestive Diseases Foundation 2004;25:130-134.

28.   Ekstedt M, Franzen LE, Mathiesen UL, Thorelius L, Holmqvist M, Bodemar G,

Kechagias S. Long-term follow-up of patients with NAFLD and elevated liver enzymes.

Hepatology 2006;44:865-873.

FIGURE LEGENDS

Figure 1. Identification of nonalcoholic fatty liver disease cases in Rochester Epidemiology

Project Database, using Hospital International Classification of Diseases Adapted (HICDA)

codes, a system developed at Mayo for research diagnosis coding, and International

Classification of Diseases (ICD)-9 codes.

Figure 2. A. The age and sex-adjusted incidence of NAFLD diagnosis in Olmsted County,

Minnesota increased 5-fold, from 62/100,000 person-years in 1997 to 329/100,000 person-

years in 2014. **B**. Incidence of NAFLD diagnosis by age groups. NAFLD incidence increased 7-

fold among subjects 18-39 years, 6-fold among subjects 40-59 years and 4-fold among subjects

over 60 years. Y-axis represents number of subjects per 100,000 person-years.

Figure 3. Proportion of NAFLD subjects with one, two or three metabolic comorbidities

(diabetes mellitus, hypertension or dyslipidemia) present at the time of diagnosis. The

proportion of subjects with multiple comorbidities at NAFLD diagnosis has increased since

1997.

Figure 4. The impact of NAFLD on incident metabolic comorbidities and mortality in reference

to age- and sex-matched controls. In this multistate model, NAFLD subjects and their age- and

sex- matched counterparts are followed longitudinally as they transition unidirectionally towards

death, directly or through intermediate states of progressive dysmetabolic burden. The relative

risk of NAFLD subjects (using controls as reference) of transitioning from one state to another is

illustrated in the arrows. Compared to controls, NAFLD subjects have a higher risk to develop

incident dysmetabolic comorbidities (top row, horizontal arrows). The impact of NAFLD on

mortality decreases as the number of dysmetabolic conditions increases (bottom row, vertical

arrows).

Figure 5. The time-in-state analysis illustrates the number of years spent in each of the four states of zero, one, two or three comorbidities (diabetes mellitus, hypertension and dyslipidemia) until death by NAFLD subjects and matched controls from the general population. The modelling was performed on NAFLD subjects age 50 and compared to their age- and sex-matched controls. The age of 50 was arbitrarily chosen as it is close to the median age in this cohort. Subjects and controls are followed longitudinally from age 50 to death, as they transition through intermediate states of comorbidities (progressively darker grey areas). The numbers in each box represent the number of years spent in each state. Compared to controls, NAFLD subjects develop higher dysmetabolic states sooner and spend proportionally more or their remaining time in states of metabolic comorbidities. The mean lifetime of NAFLD subjects is 4 years shorter than controls, in both males and females.

TABLES

Table 1. Demographic and clinical characteristics of NAFLD subjects compared to age- and

sex-matched controls from the general population.

Characteristics		NAFLD N= 3,869	Controls N= 15,209	P value
Age – median (IQR)		53 (42-63)	53 (43-64)	0.02
Female (%)		2,032 (52%)	7,973 (52%)	0.91
BMI – median (IQR) kg/m^2		33 (29-38)	28 (24-32)	<0.0001
Diabetes mellitus	at baseline	1,254 (32%)	1,614 (11%)	<0.0001
	after 10 years	1,532 (40%)	2,086 (14%)	<0.0001
Hypertension	at baseline	1,963 (51%)	4,164 (27%)	<0.0001
	after 10 years	2,334 (60%)	5,170 (34%)	<0.0001
Dyslipidemia	at baseline	2,914 (75%)	7,174 (47%)	<0.0001
	after 10 years	3,084 (80%)	7,979 (52%)	<0.0001
Cardiovascular disease	at baseline	1,102 (28%)	2,754 (18%)	<0.0001
	after 10 years	1,303 (34%)	3,304 (22%)	<0.0001

Table 2. Impact of NAFLD on mortality and clinical cardiovascular events based on baseline

and incident metabolic comorbidities in NAFLD subjects referenced to a 4:1 age-and sex-

matched general population. Bold fonts represent statistically significant results.

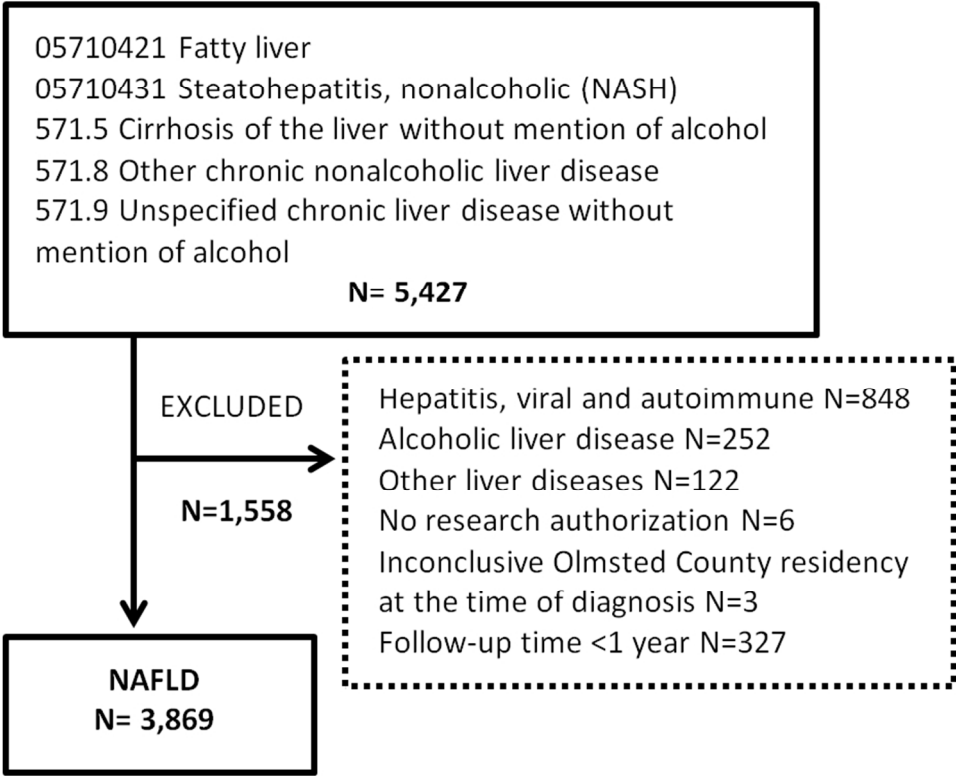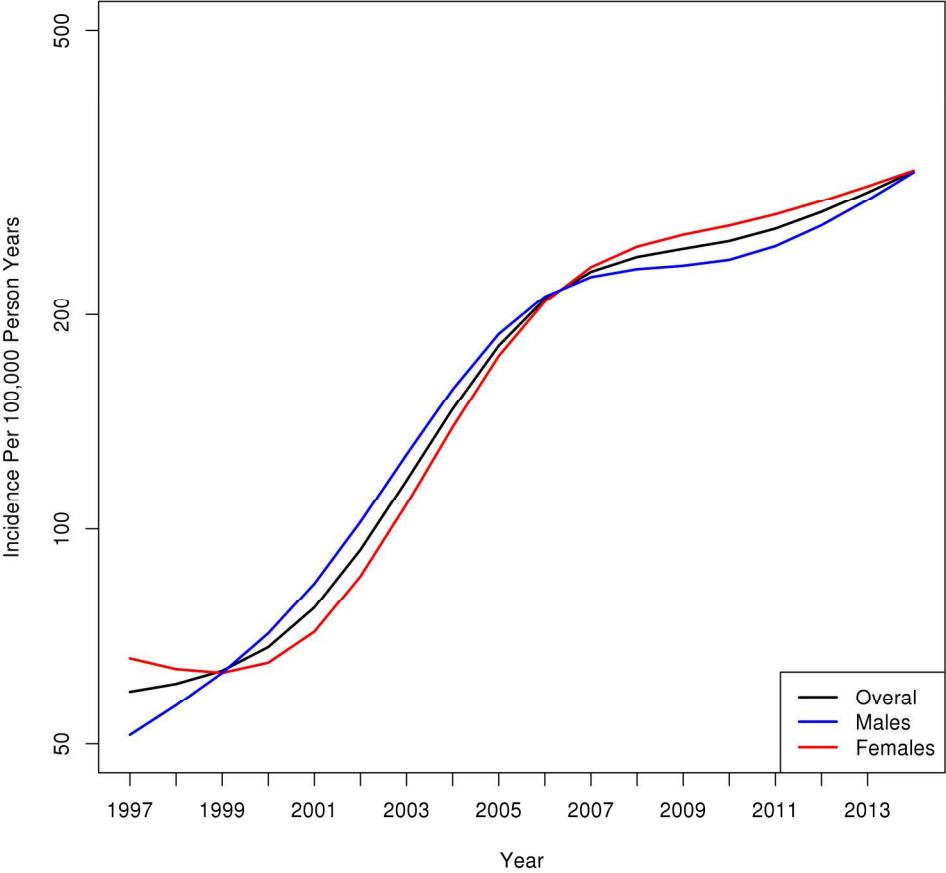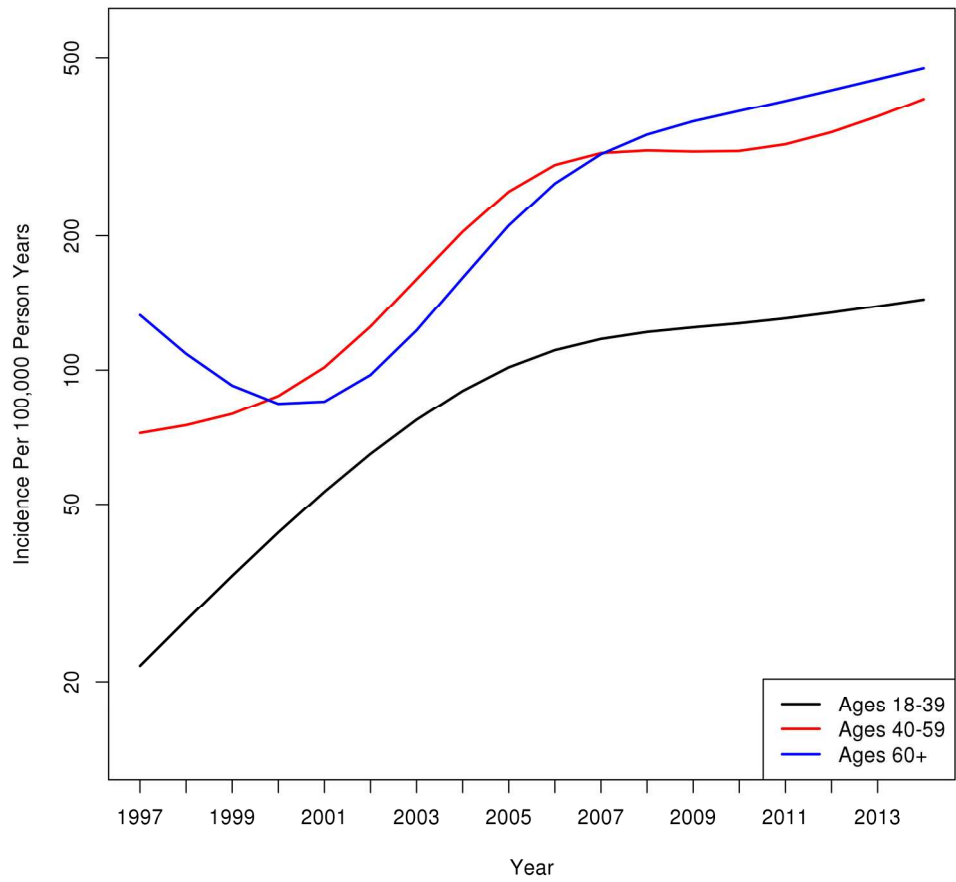|			No comorbidities	1 comorbidity	2 comorbidities	3 comorbidities
Relative risk	Mortality	Overall	**2.16**	**1.99**	**1.75**	1.08
		Stratified by cirrhosis	**1.73**	**1.56**	**1.56**	0.85
	Clinical CV events	Overall	**1.96**	1.21	**1.24**	1.02
		Stratified by cirrhosis	**2.01**	1.16	**1.23**	1.01

Figure 1

256x206mm (96 x 96 DPI)

Figure 2A
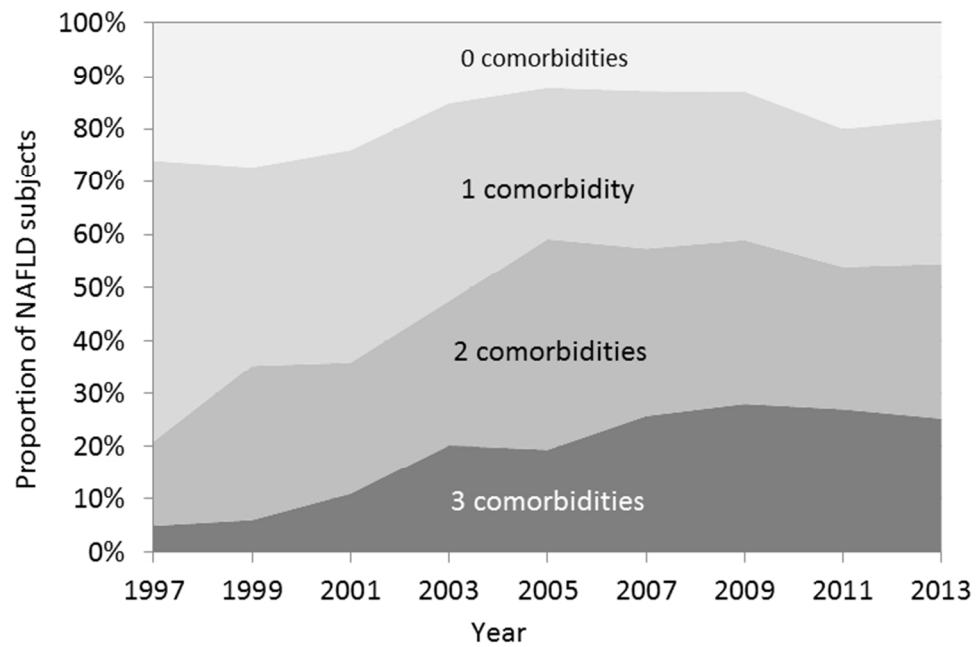
203x203mm (300 x 300 DPI)

Figure 2B
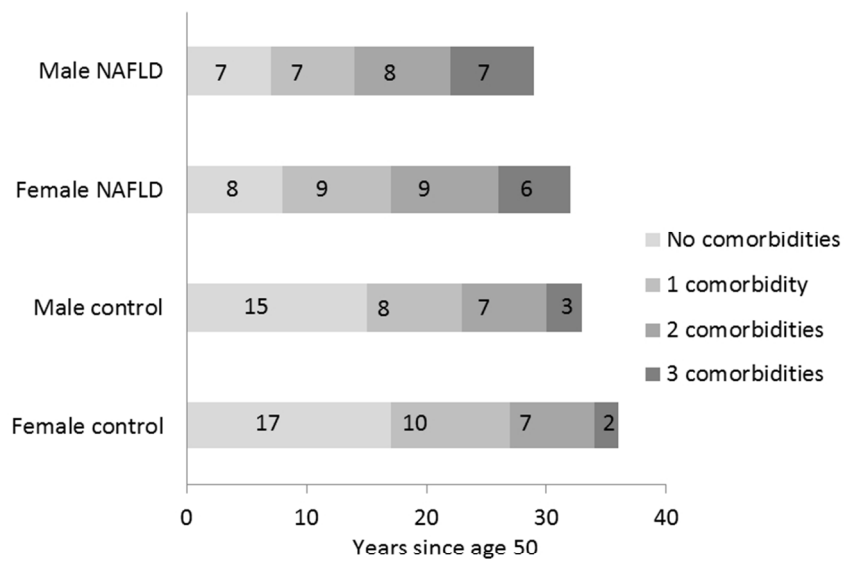
203x203mm (300 x 300 DPI)

Figure 3

210x143mm (96 x 96 DPI)

Figure 4

207x128mm (96 x 96 DPI)

Figure 5

254x190mm (96 x 96 DPI)