BMJ Open

Clinical outcome of nonalcoholic fatty liver disease: an eleven-year follow-up study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-054891
Article Type:	Original research
Date Submitted by the Author:	12-Oct-2021
Complete List of Authors:	Tang, Xiaoping ; Ningbo Zhenhai District Lianhua Hospital , Department of Prevention and Health Care Shi, Yanyan; Peking University Third Hospital, Research Center of Clinical Epidemiology Du, Juan ; Ningbo Zhenhai District Lianhua Hospital , Department of Internal Medicine Hu, Keming; Ningbo Zhenhai District Lianhua Hospital, Department of Prevention and Health Care Zhou, Tingting; Ningbo Zhenhai District Lianhua Hospital , Department of Internal Medicine Chen, Lan; Ningbo Zhenhai District Lianhua Hospital , Department of Radiology Zhang, Yanming; Ningbo Zhenhai District Lianhua Hospital , Center Laboratory Li, Fujun; Ningbo Zhenhai District Lianhua Hospital , Department of Radiology Zhang, Huier ; Ningbo Zhenhai District Lianhua Hospital , Center Laboratory Liebe, Roman ; Heinrich Heine University Düsseldorf, Clinic of Gastroenterology, Hepatology and Infectious Diseases,; Saarland University, Department of Medicine II, Saarland University Medical Center Meyer, Christoph ; Heidelberg University, Department of Medicine II, Medical Faculty Mannheim Dooley, Steven ; Heidelberg University, Department of Medicine II, Medical Faculty Mannheim Zhu, Zhongwei; Ningbo Zhenhai District Lianhua Hospital , Department of Surgery Weng, Hong-Lei ; Heidelberg University, Department of Medicine II, Medical Faculty Mannheim JIA, Jinzhu; Peking University, Department of Biostatistics, School of Public Health; Peking University, Center for Statistical Science Huang, Tong; Ningbo City First Hospital; Ningbo Zhenhai District Lianhua Hospital , Department of Prevention and Health Care
Keywords:	Epidemiology < INFECTIOUS DISEASES, Hypertension < CARDIOLOGY, DIABETES & ENDOCRINOLOGY, Health & safety < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Health informatics < BIOTECHNOLOGY & BIOINFORMATICS

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Clinical outcome of nonalcoholic fatty liver disease: an eleven-year follow-up study

Xiaoping Tang[1#], Yanyan Shi[2#], Juan Du[3#], Keming Hu[1], Tingting Zhou[3], Lan Chen[4], Yanming Zhang[5], Fujun Li[4], Huier Zhang[5], Roman Liebe[6,7], Christoph Meyer[8], Steven Dooley[8], Zhongwei Zhu[9], Hong-Lei Weng[8], Jinzhu Jia[10,11]*, Tong Huang[1,12]*

[1] Department of Prevention and Health Care, Ningbo Zhenhai Lianhua Hospital, Ningbo, China

[2] Research Center of Clinical Epidemiology, Peking University Third Hospital, Beijing, China

[3] Department of Internal Medicine, Ningbo Zhenhai Lianhua Hospital, Ningbo, China

[4] Department of Radiology, Ningbo Zhenhai Lianhua Hospital, Ningbo, China

[5] Center Laboratory, Ningbo Zhenhai Lianhua Hospital, Ningbo, China

[6] Clinic of Gastroenterology, Hepatology and Infectious Diseases, Heinrich Heine University, Düsseldorf, Germany;

[7] Department of Medicine II, Saarland University Medical Center, Saarland University, Homburg, Germany;

[8] Department of Medicine II, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany;

[9] Department of Surgery, Ningbo Zhenhai Lianhua Hospital, Ningbo, China

[10]Department of Biostatistics, School of Public Health, Peking University Health Science Center, Beijing, 100191, P.R. China

[11]Center for Statistical Science, Peking University, Beijing, 100871, P.R. China

[12]Center of health Management, Ningbo No.1 Hospital

# These authors contributed to the project equally.

* Co-senior authors

Corresponding authors:

Tong Huang, MD

Center of health Management, Ningbo No.1 Hospital

Address: No. 59, Liuting Street, Haishu District, Ningbo, Zhejiang Province, 315010, P.R. China

Email:nbbjzxht@163.com

Tel: +86-135868350952

Jinzhu Jia, PhD.
Department of Biostatistics, School of Public Health, Peking University Health Science Center
Address: No. 38, Xueyuan Road, Haidian District, Beijing, 100191, P.R. China
Email: jzjia@math.pku.edu.cn
Tel: +86-15801049187

Acknowledgements:

This project is supported by the National Key R&D Program of China No. 2017YFC0908103 (X.T., and J.D.). We are grateful to Drs. Changxi Chen, Jingyi Yuan and Yongjun Chen for support and discussion.

Author contributions:

Conception and design: Xiaoping Tang, Yanyan Shi, Juan Du, Hong-Lei Weng, Jinzhu Jia and Tong Huang

Ultrasonography: Lan Chen, Fujun Li

Blood assays: Yanming Zhang and Huier Zhang

Other examinations and data collection: Xiaoping Tang, Keming Hu, Tingting Zhou, Juan Du, Zhongwei Zhu and Tong Huang

Statistical analyses: Yanyan Shi and Jinzhu Jia

Drafting the article: Hong-Lei Weng

Reviewing and editing the article critically: Xiaoping Tang, Yanyan Shi, Christoph Meyer, Roman Liebe, Steven Dooley, Hong-Lei Weng, Jinzhu Jia and Tong Huang

Abstract

Objectives To clarify NAFLD prevalence, risk factors, and clinical outcome in China, a cohort of company employees was followed up for eleven years.

Design Retrospective cohort

Setting Between 2006-2016 in China

Participants 13032 company employees

Results Over eleven years, the prevalence of NAFLD increased from 17.2% to 32.4% (males 20.5% to 37% *versus* females 9.8% to 22.2%). Male peak prevalence was between 40 and 60 years of age, whereas highest prevalence in women was at an age of 60 years and older. Logistic and Cox regression revealed 16 risk factors. However, cause-effect analyses showed that only BMI, gender and triglycerides directly contributed to NAFLD development. Over an 11-year follow-up period, 12.6%, 37.7% and 14.2% of male NAFLD patients and 11.6%, 44.7% and 22.6% of female NAFLD patients developed diabetes, hypertension and hyperuricemia, respectively. Except one male patient who developed cirrhosis, no NAFLD patients progressed into severe liver disease.

Conclusion Diabetes, hypertension and hyperuricemia are the main clinical outcomes of NAFLD. Eleven years of NAFLD are not sufficient to cause severe liver disease. Age and obesity are direct risk factors for NAFLD. BMI, gender and triglycerides are three parameters directly reflecting the occurrence of NAFLD.

(Words: 194)

Strengths and limitations of this study

- The current study presented a dynamic NAFLD prevalence in an eastern Chinese community.
- By analyzing cause-effect link, only three parameters BMI, gender and triglycerides were confirmed to directly reflect the occurrence of NAFLD.
- In contrast to the current dogma, severe liver diseases are not the clinical outcomes of NAFLD within eleven years.
- Metabolic syndrome such as diabetes, hypertension and hyperuricemia are the main consequences of NAFLD.
- The current study is a single-center observation. Therefore, a multiple-center study is required to confirm the conclusions.

Introduction

Non-alcoholic fatty liver disease (NAFLD) is one of the most common causes of chronic liver disease globally [1]. The global prevalence of NAFLD is currently around 25% [2,3]. NAFLD is predicted to become the most frequent indication for liver transplantation by 2030 in Western countries [4]. An analysis based on 18 million patients in four European cohorts showed that NAFLD and non-alcoholic steatohepatitis (NASH) increase the risk of end-stage liver diseases, e.g., cirrhosis and hepatocellular carcinoma (HCC) [5]. Of note, NAFLD is not only a disease restricted to the liver, but also affects extra-hepatic organs. NAFLD is tightly associated with the occurrence of type 2 diabetes mellitus (T2DM), cardiovascular (CVD) and cardiac diseases, and chronic kidney disease (CKD) [4].

In China, the incidence of NAFLD has been increasing over the last two decades. A recent meta-analysis, based on 392 studies between 2008 and 2018, showed the national incidence of NAFLD in China to be at 29.2% [6]. In Shanghai, the adult incidence of NAFLD has increased from 14.04% in 1995 to 43.65% in 2015 [2]. Being a vast country, Chinese living in different areas vary widely in lifestyle and economic status. Thus, the epidemiology, natural history and clinical outcomes of NAFLD in different areas of the country are worth further investigation.

It is well accepted that viral hepatitis is a major reason for progressing chronic liver diseases, e.g., fibrosis, cirrhosis and ultimately, hepatocellular carcinoma (HCC). With regard to NAFLD, incidence and severity of associated chronic liver disease outcomes has not been monitored in large Chinese cohorts yet – especially over a long-time span. The current study therefore describes the prevalence of NAFLD in a large Eastern Chinese community over eleven years (2006 - 2016). We focused on three aspects: (1) annual prevalence of NAFLD, (2) risk factors of NAFLD, and (3) intra- and extra-hepatic clinical outcomes of NAFLD. Noteworthy, NAFLD prevalence is higher in males and peaking between 40 – 60 years, whereas in females, NAFLD if most frequently observed at an age above 60 years. Diabetes, hypertension and hyperuricemia are the main clinical outcomes of NAFLD. Intriguingly, only 1 out of 696 NAFLD patients developed liver cirrhosis within 11-year follow-up and none progress into liver cancer.

Methods

Patient and public involvement

No patient involved in this study

Design and participants

In this retrospective study we analyzed the "annual health examination database" of the Zhenhai Lianhua Hospital from 2006 to 2016. This hospital is affiliated to Sinopec Zhenhai Refining & Chemical Company. Supported by the company, all employees were offered the opportunity to go to this hospital for an annual health examination. During 11 years, a total 13,032 employees received health examinations. From 2006 to 2016, 11689, 11706, 11584, 9521, 9592, 9725, 9710, 9869, 9718, 9702 and 9706 persons received health examinations (**Figure 1**). To describe the longitudinal NAFLD occurrence in this cohort, we excluded subjects with the following conditions: (1) viral hepatitis B and C infection, which were identified by blood virus measurements (HBV-DNA and HCV-RNA), and (2) alcoholic liver disease, which was defined as previously described [7,8]. NAFLD was defined as the presence of hepatic steatosis, determined by ultrasonography.

The study protocol was approved by the Ethics Committee of Zhenhai Lianhua Hospital ([2016]001). Informed consent was obtained from all subjects.

Measures

Supplementary Table 1 shows all parameters measured in the annual health examinations. Ultrasonography was performed by the same three experienced doctors (L.C., F. L., and J.Y.) with an Ultrasongraph B, GE, Voluson 730 pro. Blood biochemistry and HBV serum levels were measured by an Olympus AU640 autoanalyzer (Olympus, Kobe, Japan) and an ImmunoAssay Analyzer VitrosECI (JOHNSON-JOHNSON, USA), respectively. All methods were carried out in accordance with relevant guidelines and regulations.

Statistical analysis

For population characteristics, variables were described as means and standard deviation (SD) or proportions as appropriate. Student's t-test or nonparametric test was used to analyze differences between two groups as mentioned. Chi-square test was used to verify the differences of nominative variables between two groups. Multivariate analysis of risk factors for NAFLD was performed using logistic regression analysis. Combined receiver operating characteristic curve (ROC) and area under curve (AUC) analyses were used to evaluate the diagnostic performance of biomarkers based on the logistic regression model. Multivariate Cox regression model was performed to calculate hazard ratios of variables to identify independent prognostic variables. First order Markov models were used to analyze the cause-effect link between

NAFLD and risk factors. L1 penalized logistic regression was applied to select predictive predictors. R package "glmnet" contains functions to select predictors using L1 penalized logistic regression. Statistical analyses were performed using SPSS 22.0 and R version 3.5.3. *P*-values that were less than 0.05 were considered statistically significant. Figures were generated by R package such as 'forestplot', 'ROCR', 'bnlearn', or 'survival'.

Results

Prevalence of NAFLD from 2006 to 2016

We retrospectively analyzed 9786, 9852, 9827, 8026, 8225, 8309, 8311, 8552, 8442, 8463 and 8436 persons who received health examinations from 2006 to 2016. **Supplementary Table 2** shows the eleven-year annual NAFLD prevalence in this population. In 2006, NAFLD was diagnosed in 17.2% of persons, and gradually increased over the examination period to 19% (2007), 22% (2008), 22.4% (2009), 22.7% (2010), 23.4% (2011), 24.5% (2012), 25.4% (2013), 27.9% (2014), 30.8% (2015), and 32.4% (2016), respectively (**Supplementary Table 2**). Both males and females demonstrated continuously increasing NAFLD prevalence (**Supplementary Table 2**). Compared to females, male Chinese demonstrated significantly higher NAFLD prevalence, e.g., in 2006, the prevalence of NAFLD in males and females was 20.5% and 9.8%, respectively. Eleven years later, the prevalence had increased to 37% in males and 22.2% in females (**Supplementary Table 2**). Noteworthy, the prevalence of NAFLD in male and female was correlating with age. The peak prevalence of NAFLD in men emerged in those aged between 40 and 60 years. In 2006, the prevalence of NAFLD in men aged between 40 - 50 and 50 - 60 years was 24.2% and 24.6%, respectively. In 2016, prevalence reached 42.8% and 46.6% (**Supplementary Table 2**) for men. Distinct from males, the peak NAFLD prevalence in females emerged at an age above 60 years. In 2006, the prevalence of NAFLD in women older than 60 and 70 years was 25.6% and 22.9%, respectively. Eleven years later, these values had increased to 53.4% and 30.9% (**Supplementary Table 2**).

Among the observed population, 5606 persons received annual health examinations for 11 years, and thus prevalence of NAFLD was analyzed in these individuals. As shown in **Table 1**, the prevalence of NAFLD increased from 17% in 2006 to 35.2% in 2016. The highest prevalence rates for NAFLD in 3795 men emerged in those older than 40 years. In 2006, the NAFLD prevalence in males aged between 40 - 50, 50 – 60, and 60 – 70 years was 24.2%, 25.4% and 19.8%, respectively. In 2016, these values reached 43.3%, 46.8% and 42.3% (**Table**

1). Different from males, the peak NAFLD prevalence in 1811 females emerged at an age of more than 60 years. In 2006, the prevalence of NAFLD in women older than 60 or 70 years was 22.1% and 28%, respectively. Eleven years later, these values had increased to 31.8% and 28.6% (**Table 1**).

BMI and NAFLD incidence

Given the tight link between obesity and NAFLD, we paid special attention to the population with high BMI. We focused on the 5606 persons with complete follow-up and analyzed the prevalence of NAFLD in those with BMI >25. In total, out of the 5606 persons, 2445 presented with a BMI of >25. The prevalence of NAFLD in this overweighted population was far higher than in the general population. In 2006, 45.2% of individuals (n=1104; male *vs.* female: 47.3% *vs.* 37.1%) with BMI >25 were suffering from NAFLD (**Supplementary Table 3**). In 2016, values reached 67.1% (n=1414; male *vs.* female: 69% *vs.* 59.2%, **Supplementary Table 3**). Impressively, the NAFLD prevalence in both genders was very high at any age, even in those below the age of 30 years. In 2006, among 213 overweighted men, younger than 30 years, 52.6% were also diagnosed for NAFLD (**Supplementary Table 3**). This number increased to 63% in 2016 (**Supplementary Table 3**). In 15 (in 2006) and 16 (in 2016) overweighted women aged less than 30 years, the NAFLD prevalence was 20% and 43.8%, respectively (**Supplementary Table 3**). In those older than 40 years, NAFLD prevalence increased from 36.6 – 45.4% in 2006 to 53 – 65.6% (**Supplementary Table 3**).

Risk factors relevant to NAFLD occurrence

Next, we analyzed risk factors relevant to NAFLD occurrence. Logistic regression analysis was performed on 26 parameters, including gender, age, BMI, albumin, white globulin ratio (WGR), white blood cell count (WBC), low density lipoprotein (LDL), triglycerides (TG), high density lipoprotein (HDL), glutamyl transpeptidase (GLT), alanine transaminase (ALT), aspartate transaminase (AST), creatinine (Cr), alkaline phosphatase (ALP), blood urea nitrogen (Bun), uric acid (UA), blood glucose (Glu), systolic blood pressure (SBp), diastolic blood pressure (DBp), Blood sedimentation (ESR), hemoglobin (HGB), platelets (PLT), Apolipoprotein A1 (ApoA1), Apolipoprotein B2 (ApoB), total bilirubin (TB), total protein (TP). We performed variable selection by penalized Logistic regression using R package glmnet. Cross validation selected 16 variables as potential predictors. These were BMI, albumin, WBC, TG, HDL, GLT, ALT, Cr, UA, Glu, SBp, DBp, ESR, HGB, PLT and ApoB (**Supplementary Table 4**). The corresponding forest-plot is shown in **Figure 2A**. Among these variables, ApoB and BMI

displayed the most robust positive correlation with NAFLD occurrence, while HDL had a strong negative correlation with NAFLD incidence (**Supplementary Table 4**). The AUC of these variables for NAFLD is 0.88 (see ROC curve in **Figure 2B**). We further performed a time dependent Cox regression to calculate the hazard ratios of these parameters for NAFLD occurrence. Cox regression confirmed that the 16 parameters were significantly relevant to NAFLD incidence (**Supplementary Table 5** and **Figure 2C**). Furthermore, ApoB and HDL were the most robust positive and negative risk factors for NAFLD (**Figure 2C**).

Cause-effect link between risk factors and NAFLD occurrence

Although the aforementioned parameters were regarded as "risk factors" according to statistical models, it did not necessarily mean that all of them contributed to NAFLD occurrence. Based on 11 years of longitudinal data, it was possible to construct a dynamic Bayesian network to identify the risk factors most relevant to NAFLD occurrence. As shown in **Figure 3**, these parameters constituted a complicated, but clear intercross paradigm. Only three parameters, BMI, gender and TG directly pointed to NAFLD. In addition, ApoB impacted the incidence of NAFLD through contributing to TG. Furthermore, LDL can indirectly contribute to NAFLD through influencing ApoB. Very impressively, the dynamic Bayesian network pointed out that NAFLD directly leads to alterations of seven parameters: ALT, DBp, SBp, TP, albumin, HGB and UA. Intriguingly, age, GLT, AST, Cr, and BUN did not interact with any other parameter in our model, indicating that these factors by incidence correlate, but not any causal interaction is existing.

Outcome of NAFLD

Subsequently, we examined clinical outcomes of NAFLD over the eleven years. **Table 2** summarizes the incidence of intra- and extrahepatic diseases of 696 NAFLD and 222 NASH patients during the follow-up period. Among the total NAFLD and NASH population, only 1 male NAFLD patient developed liver cirrhosis within the eleven years. However, this time span witnessed significantly increased extrahepatic diseases, including diabetes, hypertension and hyperuricemia. In the NAFLD population, there were 64 (12.6%) men and 22 (11.6%) women, 191 (37.7%) men and 85 (44.7%) women, 72 (14.2%) men and 43 (22.6%) women who developed into type II diabetes, hypertension and hyperuricemia, respectively (**Table 2**).

Given that 2006 is the starting point of data collection, patients diagnosed for NAFLD in this year did not mean that their disease had just manifested, but rather patients could have suffered for more years. To clarify the exact clinical outcomes of NAFLD over one decade, we focused

on the following two cohorts of individuals with annual health examinations for 11 years: (1) patients who were diagnosed as non-NAFLD in 2006, but were NAFLD in 2007 (new NAFLD cohort); and (2) who were non-NAFLD in both 2006 and 2007 (non-NAFLD cohort). As shown in **Table 2**, 185 new NAFLD cases (138 men and 47 women) and 4547 non-NAFLD (2786 men and 1761 women) persons were found in 2007. Between 2007 and 2016, neither NAFLD nor non-NAFLD individuals developed liver cirrhosis or cancer. However, the one-decade follow-up shows different prevalence of diabetes, hypertension and hyperuricemia. In NAFLD patients, there were 14 (10.1%) men and 5 (10.6%) women, 47 (34.1%) men and 21 (44.7%) women, 34 (24.6%) men and 8 (17%) women who developed type II diabetes, hypertension and hyperuricemia, respectively (**Table 2**). In non-NAFLD individuals, 157 (5.6%) men and 54 (3.4%) women, 259 (9.3%) men and 324 (18.4%), 284 (10.2%) men and 84 women (4.8%) developed into type II diabetes, hypertension and hyperuricemia, respectively (**Table 2**). For all three diseases, statistically significant differences were determined between the two cohorts of population (all P <0.05, **Table 2**). These results suggest that diabetes, hypertension and hyperuricemia are the main clinical outcomes of NAFLD.

Discussion

This eleven-year follow-up retrospective study reports the following: (1) NAFLD prevalence has substantially increased in the examined Eastern Chinese population. (2) The prevalence of NAFLD differs by gender and age. Middle-aged men and elderly women are the two populations at highest risk for NAFLD. (3) Gender, BMI and triglycerides are the parameters directly associated with NAFLD occurrence. Regardless of gender and age, persons with high BMI (≥25) have a high risk for NAFLD development. (4) NAFLD directly leads to alterations of seven clinical parameters: ALT, DBp, SBp, TP, albumin, HGB and UA. (5) Within 11 years, a significant part of the NAFLD population develops three clinically relevant diseases: type 2 diabetes mellitus, hypertension and hyperuricemia. (6) Within 11 years, NAFLD does not cause severe liver disease, such as cirrhosis or HCC, in patients.

The most impressive observation of the current study is that among 918 diseased persons (696 NAFLD and 222 NASH), no patient progressed towards HCC and only 1 male NAFLD patient developed liver cirrhosis within the 11 years. Furthermore, among 185 new NAFLD cases diagnosed in 2007, none developed liver cirrhosis or liver cancer. Liver cirrhosis and HCC are commonly regarded as the most severe and costly clinical outcomes of NAFLD [9]. In USA and Europe, it is estimated that 10-15% of NAFLD patients develop advanced fibrosis [10]. In China,

a study with biopsy-proven NAFLD revealed 1.97%-2.97% cirrhosis prevalence [11]. In addition, NAFLD is regarded as the third-most common cause of cancer-related death worldwide [12]. In a study based on 4949 US patients with HCC, 701 patients had NAFLD [13]. It was estimated that the cumulative incidence of HCC among patients with NAFLD and cirrhosis ranges from 2.4% to 12.8% over a median follow-up period of 3.2 to 7.2 years [14] (Global Health Observatory) data. Mortality and global health estimates were obtained from: http://www.who.int/gho/mortality_burden_disease/en/. Last accessed on 1/7/2020.). Given that the above conclusions were based on cross-sectional investigations and statistical models, it has been unknown to date over which period a NAFLD patient develops liver cirrhosis or HCC (personal risk assessment). Our 11-year follow-up provides therefore a valuable and comprehensive dataset. In this study, most patients were diagnosed with NAFLD when they received a routine health examination. Before the examination, these people did not have any symptoms or signs of NAFLD. Therefore, they belong to NAFLD patients at a very early stage (although for 2006, the duration of pre-existing NAFLD cannot be determined). Except for a single person, no serious liver problems were observed within this time period. These data suggest that for the vast majority of patients with early stage NAFLD, 11 years are not sufficient to develop liver cirrhosis or cancer. Nasr et al followed up 129 NAFLD patients with varying fibrosis stages on two occasions (mean time 13.7 and 9.3 years). Liver biopsy analyses showed that 9.3% of patients developed end-stage liver disease and 34% advanced fibrosis [15]. The NAFLD patients observed by Nasr et al. actually belonged to the NASH category, because they suffered from fibrosis and elevated ALT and/or AST levels. As a study based on healthy examination, liver biopsy is impossible for such a study. Very likely, the current cohort included a portion of NASH patients. They also did not show significant progression towards cirrhosis or HCC was monitored.

In contrast to hepatic complications, patients with NAFLD showed a significant risk for the development of extrahepatic diseases, including diabetes, hypertension and hyperuricemia. In 696 NAFLD patients, eleven years witnessed the development of 86 cases (12.4%) type II diabetes, 276 (40%) cases of hypertension and 115 patients with (16.5%) hyperuricemia, respectively. Interestingly, in 222 NASH patients, the prevalence of these three diseases was 12 (5.4%), 46 (20.7%), and 33 (14.9%) only. In general, men had a higher probability to develop these diseases than women. These results are consistent with previous reports from USA and Europe [16-18].

To date, there are a large number of studies investigating risk factors for NAFLD [19]. These studies tried to identify single, or multiple combined biomarkers to predict NAFLD occurrence. Given that most studies were based on cross-sectional designs, or with only short follow-up periods, it is difficult to clarify the causality between the proposed predictors and NAFLD morbidity. Our eleven-year dataset provides a chance to shed led on this issue. Here, the dynamic causal relationships between variables, including risk parameters and clinical outcomes, were identified by a first order Markov model, which was displayed by a dynamic Bayes network. The dynamic Bayes model discriminates causal relationship through time sequence. When a variable change is closely related to a previous variance alteration, a causal relationship between the two variables is assumed. Based on Logistic and Cox regression and dynamic Bayesian network analyses, we confirmed three direct risk factors for NAFLD occurrence: gender, BMI and TG. These findings are supported by the following data: (1) Men have higher NAFLD prevalence than women in this population (37% *versus* 22.2% in 2016); (2) In overwgithed people with a BMI >25, NAFLD prevalence reached 69% in males and 59.2% in females. Given that triglycerides are a major energy source, but are leading to obesity, it is not surprising that this parameter directly reflects the risk for NAFLD development. These findings provide robust evidence supporting the use of BMI to monitor or predict NAFLD.

Conclusion

This 11-year follow-up study documents the rapid increase in NAFLD prevalence in an Eastern Chinese population. In contrast to previous reports, our observation does not observe that one decade of NAFLD is sufficient to lead to severe hepatic clinical outcomes. It is worthy to note that our population was biased towards physically fit and active people in full employment, while previous studies were often based on hospital populations, who suffered from negative selection bias and thus came up with higher estimates. In addition, given there are differences in NAFLD profiles between Eastern and Western populations, it would be interesting to know the natural development of NAFLD in a Western population. A key point for clarifying the true history of NAFLD is to follow a population starting from the early phases of the disease. Consistent with previous studies, NAFLD does indeed induce multiple extra-hepatic diseases relevant to the metabolic syndrome. In the future, follow-up of the current cohort for another one and two decades will provide further valuable data to clarify the extended natural history of NAFLD. Last but not least, a large portion of the men and women in this study were educated above the average and have a position in the company that gave them the availability of better

food choices as well as regular sport. On the other hand, the relatively low sensitivity of ultrasound for the detection of liver fat might underestimate the incidence of NAFLD in this cohort.

References

[1] Younossi ZM, Tampi R, Priyadarshini M, et al. Burden of Illness and Economic Model for Patients With Nonalcoholic Steatohepatitis in the United States. Hepatology, 2019, 69: 564-572.

[2] Younossi ZM, Loomba R, Anstee QM, et al. Diagnostic modalities for nonalcoholic fatty liver disease, nonalcoholic steatohepatitis, and associated fibrosis. Hepatology, 2018, 68: 349-360.

[3] Fan JG, Kim SU, Wong VW. New trends on obesity and NAFLD in Asia. J Hepatol, 2017, 67: 862-873.

[4] Byrne CD, Targher G. NAFLD: a multisystem disease. J Hepatol, 2015, 62: S47-64.

[5] Alexander M, Loomis AK, van der Lei J, et al. Risks and clinical predictors of cirrhosis and hepatocellular carcinoma diagnoses in adults with diagnosed NAFLD: real-world study of 18 million patients in four European cohorts. BMC Med, 2019, 17: 95.

[6] Zhou F, Zhou J, Wang W, et al. Unexpected Rapid Increase in the Burden of NAFLD in China From 2008 to 2018: A Systematic Review and Meta-Analysis. Hepatology, 2019, 70: 1119-1133.

[7] EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. J Hepatol, 2016, 64: 1388-1402.

[8] Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of nonalcoholic fatty liver disease: Practice guidance from the American Association for the Study of Liver Diseases. Hepatology, 2018, 67: 328-357.

[9] Younossi Z, Tacke F, Arrese M, et al. Global Perspectives on Nonalcoholic Fatty Liver Disease and Nonalcoholic Steatohepatitis. Hepatology, 2019, 69: 2672-2682.

[10] Younossi ZM, Koenig AB, Abdelatif D, et al. Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. Hepatology, 2016, 64: 73-84.

[11] Shen F, Zheng RD, Shi JP, et al. Impact of skin capsular distance on the performance of controlled attenuation parameter in patients with chronic liver disease. Liver Int, 2015, 35: 2392-2400.

[12] Ascha MS, Hanouneh IA, Lopez R, et al. The incidence and risk factors of hepatocellular carcinoma in patients with nonalcoholic steatohepatitis. Hepatology, 2010, 51: 1972-1978.

[13] Younossi ZM, Otgonsuren M, Henry L, et al. Association of nonalcoholic fatty liver disease (NAFLD) with hepatocellular carcinoma (HCC) in the United States from 2004 to 2009. Hepatology, 2015, 62: 1723-1730.

[14] White DL, Kanwal F, El-Serag HB. Association between nonalcoholic fatty liver disease and risk for hepatocellular cancer, based on systematic review. Clin Gastroenterol Hepatol, 2012, 10: 1342-1359.e1342.

[15] Nasr P, Ignatova S, Kechagias S, et al. Natural history of nonalcoholic fatty liver disease: A prospective follow-up study with serial biopsies. Hepatol Commun, 2018, 2: 199-210.

[16] Estes C, Razavi H, Loomba R, et al. Modeling the epidemic of nonalcoholic fatty liver disease demonstrates an exponential increase in burden of disease. Hepatology, 2018, 67: 123-133.

[17] Williamson RM, Price JF, Glancy S, et al. Prevalence of and risk factors for hepatic steatosis and nonalcoholic Fatty liver disease in people with type 2 diabetes: the Edinburgh Type 2 Diabetes Study. Diabetes Care, 2011, 34: 1139-1144.

[18] Targher G, Bertolini L, Padovani R, et al. Prevalence of nonalcoholic fatty liver disease and its association with cardiovascular disease among type 2 diabetic patients. Diabetes Care, 2007, 30: 1212-1218.

[19] Miyake T, Kumagi T, Furukawa S, et al. Non-alcoholic fatty liver disease: factors associated with its presence and onset. J Gastroenterol Hepatol, 2013, 28 Suppl 4: 71-78.

Figure legends

Figure 1. Flow chart depicting the enrollment of a population with non-alcoholic fatty liver disease (NAFLD) for follow-up in Ningbo Zhenhai Lianhua Hospital, China.

Figure 2. Penalized logistic regression and Cox regression analysis weres performed for risk factors and hazard ratios of non-alcoholic fatty liver disease (NAFLD). The following parameters were available from 5606 participants: Gender, age, BMI, albumin, white globulin ratio (WGR), white blood cell (WBC), low density lipoprotein (LDL), triglycerides (TG), high density lipoprotein (HDL), glutamyl transpeptidase (GLT), alanine transaminase (ALT), aspartate transaminase (AST), creatinine (Cr), alkaline phosphatase (ALP), blood urea nitrogen (Bun), Uric Acid (UA), blood glucose (Glu), systolic blood pressure (SBp), diastolic blood pressure (DBp), blood sedimentation (ESR), hemoglobin (HGB), platelets (PLT), Apolipoprotein A1 (ApoA1), Apolipoprotein B2 (ApoB), total bilirubin (TB), total protein (TP). Cross validation selected 16 variables to be potential predictors. The corresponding forest-plot is shown in (**A**). The AUC of these above 16 variables for NAFLD is 0.88 **(B)**. Cox regression confirmed that the 16 variables were relevant for NAFLD incidence, including BMI, albumin, WBC, TG, HDL, GLT, ALT, Cr, UA, Glu, SBp, DBp, ESR, HGB, PLT and ApoB. The corresponding forest-plot is shown (**C**).

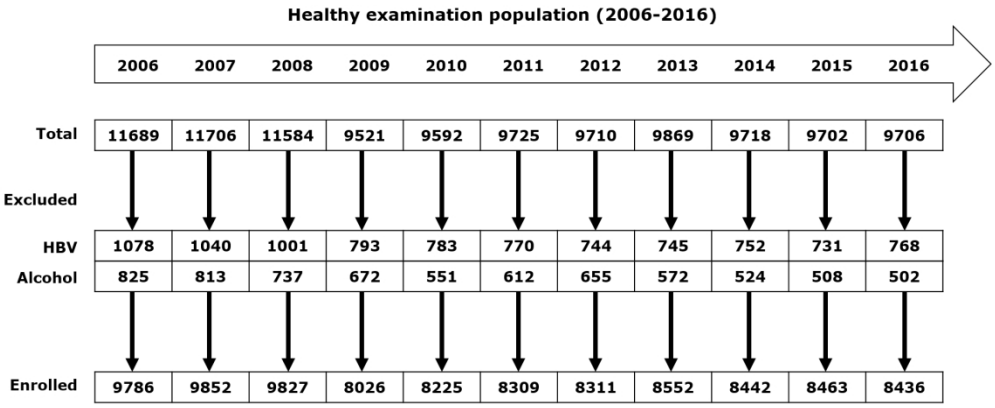
Figure 3. Dynamic Bayesian network analyses were performed to show the cause-effect link between non-alcoholic fatty liver disease (NAFLD) and its potential risk factors. Three variables, BMI, gender and trigylcerides directly pointed to NAFLD. ApoB impacted on the incidence of NAFLD through TG abundance. LDL indirectly contributed to NAFLD through ApoB. NAFLD directly led to alterations of seven clinical parameters: ALT, DBp, SBp, TP, albumin, HGB and UA.

Table 1. Prevalence of NAFLD in 5606 persons with 11-year follow-up (2006-2016)

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Total (n)	5606	5606	5606	5606	5606	5606	5606	5606	5606	5606	5606
NAFLD (n)	951	1068	1247	1322	1378	1466	1524	1574	1700	1883	1976
(%)	17	19.1	22.2	23.6	24.6	26.2	27.2	28.1	30.3	33.6	35.2
Male (n)	3795	3795	3795	3795	3795	3795	3795	3795	3795	3795	3795
NAFLD (n)	776	871	1014	1075	1117	1206	1244	1283	1387	1517	1569
(%)	20.4	23	26.7	28.3	29.4	31.8	32.8	33.8	36.5	40	41.3
≤ 30ys (n)	668	466	366	297	246	204	142	67	14	4	0
NAFLD (n)	97	71	57	46	35	35	30	17	3	1	0
(%)	14.5	15.2	15.6	15.5	14.2	17.2	21.1	25.4	21.4	25	0
>30, ≤ 40ys (n)	1212	1315	1342	1258	1180	1091	1034	981	896	778	668
NAFLD (n)	231	286	349	349	338	330	330	331	331	301	261
(%)	19.1	21.7	26	27.7	28.6	30.2	31.9	33.7	36.9	38.7	39.1
>40, ≤ 50ys (n)	873	863	842	928	988	1022	1097	1128	1170	1201	1212
NAFLD (n)	211	215	239	278	314	353	388	411	454	524	525
(%)	24.2	24.9	28.4	30	31.8	34.5	35.4	36.4	38.8	43.6	43.3
>50, ≤ 60ys (n)	562	623	696	741	776	826	796	798	820	851	873
NAFLD (n)	143	184	231	262	276	314	292	286	320	363	409
(%)	25.4	29.5	33.2	35.4	35.6	38	36.7	35.8	39	42.7	46.8
>60, ≤ 70ys (n)	368	382	373	348	346	356	398	465	496	521	562
NAFLD (n)	73	85	97	95	96	108	125	156	178	210	238
(%)	19.8	22.3	26	27.3	27.7	30.3	31.4	33.5	35.9	40.3	42.3
>70ys	112	146	176	223	259	296	328	356	399	440	480
NAFLD (n)	21	30	41	45	58	66	79	82	101	118	136
(%)	18.8	20.5	23.3	20.2	22.4	22.3	24.1	23	25.3	26.8	28.3
Female (n)	1811	1811	1811	1811	1811	1811	1811	1811	1811	1811	1811
NAFLD (n)	175	197	233	247	261	260	280	291	313	366	407
(%)	9,7	10,9	12,9	13,6	14,4	14,4	15,5	16,1	17,3	20,2	22,5
≤ 30ys (n)	85	44	22	13	9	9	5	2	0	0	0
NAFLD (n)	3	1	1	1	0	0	0	0	0	0	0
(%)	3.5	2.3	4.5	7.7	0	0	0	0	0	0	0
>30, ≤ 40ys (n)	582	578	541	480	403	334	277	213	173	124	85
NAFLD (n)	17	21	25	22	24	18	18	13	15	14	12
(%)	2.9	3.6	4.6	4.6	6	5.4	6.5	6.1	8.7	11.3	14.1
>40, ≤ 50ys (n)	485	469	482	496	541	580	612	638	617	608	582
NAFLD (n)	31	31	35	41	47	54	61	66	65	81	86
(%)	6.4	6.6	7.3	8.3	8.7	9.3	10	10.3	10.5	13.3	14.8
>50, ≤ 60ys (n)	365	400	422	450	461	469	456	452	467	476	485
NAFLD (n)	56	67	85	86	90	88	88	83	88	98	109
(%)	15.3	16.8	20.1	19.1	19.5	18.8	19.3	18.4	18.8	20.6	22.5
>60, ≤ 70ys (n)	244	260	262	267	266	266	280	301	314	337	365
NAFLD (n)	54	62	61	66	69	66	70	81	82	104	116
(%)	22.1	23.8	23.3	24.7	25.9	24.8	25	26.9	26.1	30.9	31.8
>70ys (n)	50	60	82	105	131	153	181	205	240	266	294
NAFLD (n)	14	15	26	31	31	34	43	48	63	69	84
(%)	28	25	31.7	29.5	23.7	22.2	23.8	23.4	26.3	25.9	28.6

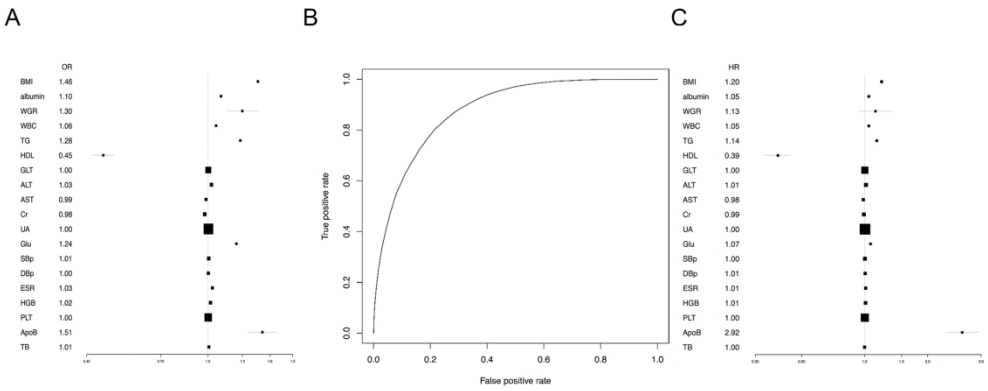
Table 2. Clinical outcome of NAFLD patients

2006 - 2016								
		Cirrhosis	HCC	Diabetes n (%)		Hypertension n (%)		Hyperuricemia n (%)	
Male (n=506)		0	0	64 (12.6)		191 (37.7)		72 (14.2)	
Female (n=190)		0	0	22 (11.6)		85 (44.7)		43 (22.6)	
2007 – 2016 (outcome of new NFLAD)									
		Cirrhosis	HCC	Diabetes		Hypertension		Hyperuricemia	
				n (%)	*P*-value	n (%)	*P*-value	n (%)	*P*-value
Male	NAFLD n=138	0	0	14 (10.1)	0.028	47(34.1)	< 0.001	34 (24.6)	< 0.001
	Non-NAFLD n=2786	0	0	157 (5.6)		259 (9.3)		284 (10.2)	
Female	NAFLD n=47	0	0	5 (10.6)	0.014	21 (44.7)	< 0.001	8 (17)	< 0.001
	Non-NAFLD n=1761	0	0	54 (3.1)		324 (18.4)		84 (4.8)	

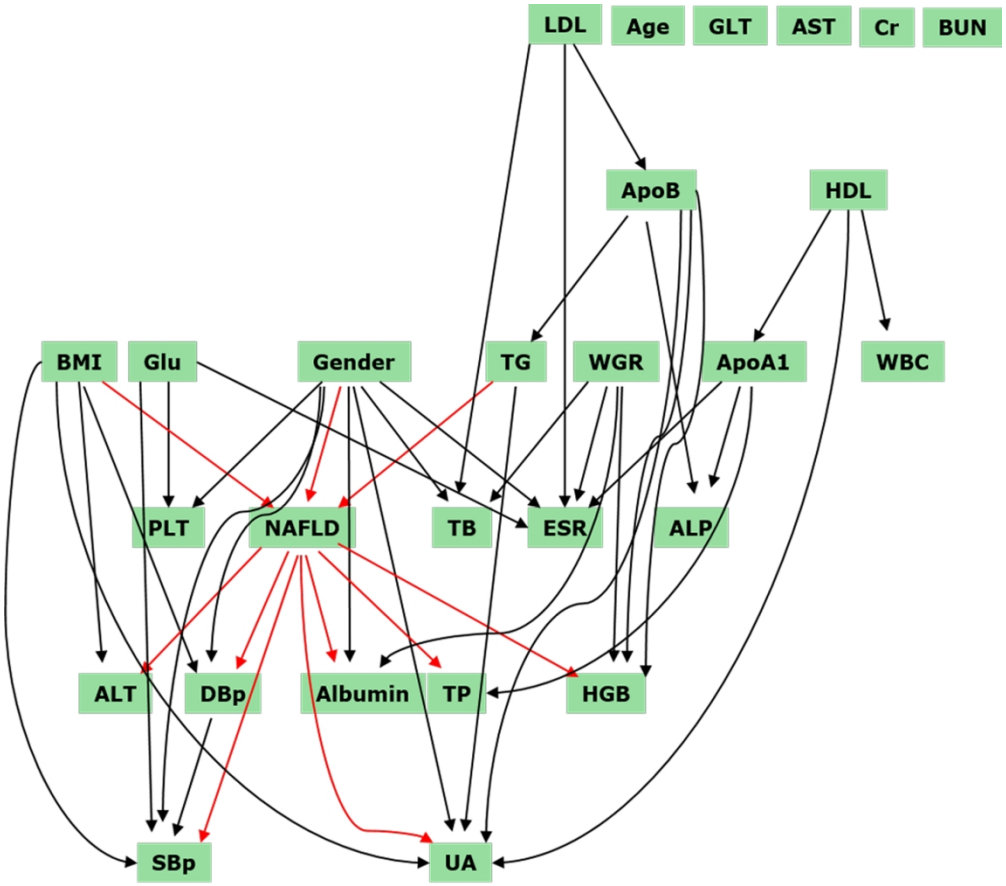


Flow chart depicting the enrollment of a population with non-alcoholic fatty liver disease (NAFLD) for follow-up in Ningbo Zhenhai Lianhua Hospital, China.

335x136mm (150 x 150 DPI)



120x59mm (600 x 600 DPI)



Dynamic Bayesian network analyses were performed to show the cause-effect link between non-alcoholic fatty liver disease (NAFLD) and its potential risk factors. Three variables, BMI, gender and trigylcerides directly pointed to NAFLD. ApoB impacted on the incidence of NAFLD through TG abundance. LDL indirectly contributed to NAFLD through ApoB. NAFLD directly led to alterations of seven clinical parameters: ALT, DBp, SBp, TP, albumin, HGB and UA.

522x461mm (96 x 96 DPI)

Nonalcoholic fatty liver disease increased risk of metabolic diseases, but not severe liver disease: an eleven-year follow-up study

Xiaoping Tang, Yanyan Shi, Juan Du, Keming Hu, Tingting Zhou, Lan Chen, Yanming Zhang, Fujun Li, Huier Zhang, Roman Liebe, Christoph Meyer, Steven Dooley, Zhongwei Zhu, Hong-Lei Weng, Jinzhu Jia, Tong Huang

Supplementary Table 1. Parameters measured in the annual health examinations.

Age
Albumin
AKP (alkaline phosphatase)
ALT (alanine transaminase)
ApoA1 (Apolipoprotein A1)
ApoB (Apolipoprotein B2)
AST (aspartate transaminase)
BMI (Body mass index)
BLRV (whole blood low shear reduced viscosity)
BLRI (relative index of whole blood low shear)
BHRV (whole blood high shear reduced viscosity)
BHRI (relative index of whole blood high shear)
BVV200 (Whole blood viscosity value)
BUN (blood urea nitrogen)
BUS (ultrasound prompt)
CRP (high sensitive C-reactive protein)
Cr (creatinine)
CA (carotid atherosclerosis)
DBIL (Direct bilirubin)
DBp (diastolic blood pressure)
DM (type II diabetes)
ESR (Blood sedimentation)
ESRKV (Blood sedimentation equation K value)
Gender
GLT (glutamyl transpeptidase)
Glucose
HBP (Hypertension)
HBX (red blood cell deformation index TK)
HCT (Hematocrit)
HCY (Homocysteine)
HDL (high density lipoprotein C)
Height
HGB (hemoglobin)
LDL (low density lipoprotein C)
LVH (left ventricular hypertrophy)
MPV (mean platelet volume)
NAFLD (non-alcoholic fatty liver disease)
PhyExa (physical examination results)
PV (plasma viscosity)
PDW (Platelet distribution width)
PLT (platelet)
PCT (prothrombin consumption time)
RBC (red blood cell count)
SBp (systolic blood pressure)
TB (Total Bilirubin)
TC (Total cholesterol)
TG (Triglyceride)

TP (total protein)
UA (uric acid)
Waist
Weight
WGR (white globulin ratio)
WBC (white blood cell count)

Supplementary Table 2. Prevalence of NAFLD in an eastern Chinese population (2006-2016)

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Total (n)	9786	9852	9827	8026	8225	8309	8311	8552	8442	8463	8436
NAFLD (n)	1687	1875	2161	1797	1871	1943	2033	2169	2356	2610	2734
(%)	17.2	19.0	22.0	22.4	22.7	23.4	24.5	25.4	27.9	30.8	32.4
Male (n)	6834	6872	6857	5468	5640	5692	5680	5891	5820	5812	5816
NAFLD (n)	1399	1555	1790	1419	1488	1562	1616	1734	1896	2081	2152
(%)	20.5	22.6	26.1	26.0	26.4	27.4	28.5	29.4	32.6	35.8	37.0
≤ 30ys (n)	1196	938	729	826	980	1092	1068	1163	1013	964	927
NAFLD (n)	155	112	100	78	83	94	108	157	171	208	200
(%)	13.0	11.9	13.7	9.4	8.5	8.6	10.1	13.5	16.9	21.6	21.6
>30, ≤ 40ys (n)	2144	2292	2353	1504	1395	1248	1150	1096	1028	970	953
NAFLD (n)	419	502	590	412	386	375	358	367	372	357	344
(%)	19.5	21.9	25.1	27.4	27.7	30.0	31.1	33.5	36.2	36.8	36.1
>40, ≤ 50ys (n)	1480	1465	1450	1107	1183	1211	1279	1299	1339	1360	1347
NAFLD (n)	358	376	429	334	380	418	456	478	517	595	577
(%)	24.2	25.7	29.6	30.2	32.1	34.5	35.7	36.8	38.6	43.8	42.8
>50, ≤ 60ys (n)	1084	1180	1314	1033	1055	1068	1008	979	978	984	978
NAFLD (n)	267	341	415	352	364	394	361	350	385	416	456
(%)	24.6	28.9	31.6	34.1	34.5	36.9	35.8	35.8	39.4	42.3	46.6
>60, ≤ 70ys (n)	635	655	622	553	547	574	632	774	834	879	937
NAFLD (n)	137	147	165	151	159	172	202	243	292	331	382
(%)	21.6	22.4	26.5	27.3	29.1	30.0	32.0	31.4	35.0	37.7	40.8
>70ys	295	342	389	445	480	499	543	580	628	655	674
NAFLD (n)	63	77	91	92	116	109	131	139	159	174	193
(%)	21.4	22.5	23.4	20.7	24.2	21.8	24.1	24.0	25.3	26.6	28.6
Female (n)	2952	2980	2970	2558	2585	2617	2631	2661	2622	2651	2620
NAFLD (n)	288	320	371	378	383	381	417	435	460	529	582
(%)	9.8	10.7	12.5	14.8	14.8	14.6	15.8	16.3	17.5	20.0	22.2
≤ 30ys (n)	209	167	147	143	213	251	248	270	239	244	214
NAFLD (n)	5	4	4	1	0	1	2	8	9	14	13
(%)	2.4	2.4	2.7	0.7	0.0	0.4	0.8	3.0	3.8	5.7	6.1
>30, ≤ 40ys (n)	934	924	869	596	496	415	348	276	232	198	177
NAFLD (n)	33	29	36	28	28	24	23	17	19	20	20
(%)	3.5	3.1	4.1	4.7	5.6	5.8	6.6	6.2	8.2	10.1	11.3
>40, ≤ 50ys (n)	801	785	776	627	666	701	733	754	707	704	658
NAFLD (n)	54	60	63	62	69	76	85	88	84	98	99
(%)	6.7	7.6	8.1	9.9	10.4	10.8	11.6	11.7	11.9	13.9	15.0
>50, ≤ 60ys (n)	536	604	651	643	647	659	641	638	665	680	689
NAFLD (n)	78	100	133	131	132	128	132	132	131	157	171
(%)	14.6	16.6	20.4	20.4	20.4	19.4	20.6	20.7	19.7	23.1	24.8
>60, ≤ 70ys (n)	367	368	367	358	347	347	380	417	438	461	494
NAFLD (n)	94	96	92	99	95	92	106	117	123	140	159
(%)	25.6	26.1	25.1	27.7	27.4	26.5	27.9	28.1	28.1	30.4	32.2
>70ys	105	132	160	191	216	244	281	306	341	364	388
NAFLD (n)	24	31	43	57	59	60	69	73	94	100	120
(%)	22.9	23.5	26.9	29.8	27.3	24.6	24.6	23.9	27.6	27.5	30.9

Supplementary Table 3. Prevalence of NAFLD in obese persons (BMI > 25) (2006-2016)

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Total (n)	2445	2674	2749	2119	2314	2271	2218	2593	2159	2121	2106
NAFLD											
(n)	1104	1256	1397	1103	1231	1218	1239	1387	1309	1384	1414
(%)	45.2	47	50.8	52.1	53.2	53.6	55.9	53.5	60.6	65.3	67.1
Male (n)	1927	2163	2232	1663	1822	1813	1744	2086	1728	1716	1699
NAFLD											
(n)	912	1053	1180	885	1002	1001	1010	1153	1086	1147	1173
(%)	47.3	48.7	52.9	53.2	55	55.2	57.9	55.3	62.8	66.8	69
≤ 30ys (n)	213	202	174	156	190	227	220	293	224	212	200
NAFLD											
(n)	112	86	67	53	64	70	79	115	119	132	126
(%)	52.6	42.6	38.5	34	33.7	30.8	35.9	39.2	53.1	62.3	63
>30, ≤											
40ys (n)	549	640	689	427	428	375	346	407	310	302	291
NAFLD											
(n)	269	325	388	259	250	229	226	247	229	211	206
(%)	49	50.8	56.3	60.7	58.4	61.1	65.3	60.7	73.9	69.9	70.8
>40, ≤											
50ys (n)	433	456	468	353	414	423	431	517	443	437	425
NAFLD											
(n)	210	225	263	185	248	264	276	321	293	323	310
(%)	48.5	49.3	56.2	52.4	59.9	62.4	64	62.1	66.1	73.9	72.9
>50, ≤											
60ys (n)	351	447	491	364	390	392	353	381	312	309	303
NAFLD											
(n)	173	240	280	222	236	247	209	226	207	217	232
(%)	49.3	53.7	57	61	60.5	63	59.2	59.3	66.3	70.2	76.6
>60, ≤											
70ys (n)	265	274	249	213	213	209	222	291	263	262	281
NAFLD											
(n)	98	111	113	102	113	113	133	155	150	159	189
(%)	37	40.5	45.4	47.9	53.1	54.1	59.9	53.3	57	60.7	67.3
>70ys	116	144	161	150	187	187	172	197	176	194	199
NAFLD											
(n)	50	66	69	64	91	78	87	89	88	105	110
(%)	43.1	45.8	42.9	42.7	48.7	41.7	50.6	45.2	50	54.1	55.3
Female (n)	518	511	517	456	492	458	474	507	431	405	407
NAFLD											
(n)	192	203	217	218	229	217	229	234	223	237	241
(%)	37.1	39.7	42	47.8	46.5	47.4	48.3	46.2	51.7	58.5	59.2
≤ 30ys (n)	15	5	11	9	6	12	13	15	13	14	16
NAFLD											
(n)	3	1	2	0	0	1	2	5	5	5	7
(%)	20	20	18.2	0	0	8.3	15.4	33.3	38.5	35.7	43.8
>30, ≤											
40ys (n)	69	78	75	46	50	36	36	31	21	20	18
NAFLD											
(n)	20	18	22	18	22	14	17	9	10	11	11
(%)	29	23.1	29.3	39.1	44	38.9	47.2	29	47.6	55	61.1
>40, ≤											
50ys (n)	101	107	109	83	90	88	94	115	89	76	66
NAFLD											
(n)	37	41	41	34	36	39	44	48	45	46	35
(%)	36.6	38.3	37.6	41	40	44.3	46.8	41.7	50.6	60.5	53
>50, ≤											
60ys (n)	141	146	146	151	155	137	121	125	101	92	93
NAFLD											
(n)	49	56	63	72	72	66	56	61	53	62	61
(%)	34.8	38.4	43.2	47.7	46.5	48.2	46.3	48.8	52.5	67.4	65.6
>60, ≤											
70ys (n)	141	127	115	103	115	105	120	127	109	105	105
NAFLD											
(n)	64	66	58	59	65	59	68	67	57	63	66
(%)	45.4	52	50.4	57.3	56.5	56.2	56.7	52.8	52.3	60	62.9
>70ys (n)	51	48	61	64	76	80	90	94	98	98	109
NAFLD											
(n)	19	21	31	35	34	38	42	44	53	50	61
(%)	37.3	43.8	50.8	54.7	44.7	47.5	46.7	46.8	54.1	51	56

Supplementary Table 4. Logistic regression for risk factors of NAFLD

	Estimate	Std. Error	z value	OR	2.5%CI	97.5%CI	Pr(>|z|)
(Intercept)	-21.63	0.3151	-68.631	4.05E-10	2.18E-10	7.49E-10	< 0.001***
BMI	0.3791	0.005224	72.563	1.460966	1.446132	1.476054	< 0.001***
albumin	0.0994	0.005635	17.64	1.104513	1.092391	1.116791	< 0.001***
WGR	0.2666	0.06118	4.357	1.305464	1.157849	1.471676	< 0.001***
WBC	0.06149	0.008274	7.432	1.063423	1.046307	1.080801	< 0.001***
TG	0.2447	0.01213	20.17	1.277201	1.247335	1.308081	< 0.001***
HDL	-0.7862	0.04247	-18.511	0.45558	0.419127	0.495053	< 0.001***
GLT	0.002947	0.000358	8.242	1.002952	1.002255	1.003661	< 0.001***
ALT	0.02717	0.001056	25.738	1.027538	1.025425	1.029676	< 0.001***
AST	-0.0142	0.001834	-7.745	0.985896	0.982336	0.989418	< 0.001***
Cr	-0.02375	0.001025	-23.177	0.976528	0.974562	0.978485	< 0.001***
ALP	0.000909	0.000547	1.661	1.000909	0.999835	1.001982	0.09664
UA	0.00433	0.000176	24.542	1.004339	1.003992	1.004687	< 0.001***
Glu	0.2159	0.01197	18.044	1.241037	1.212327	1.270567	< 0.001***
SBp	0.006338	0.001014	6.252	1.006358	1.00436	1.008359	< 0.001***
DBp	0.003561	0.001546	2.303	1.003567	1.000532	1.006614	0.021265*
ESR	0.03401	0.001808	18.808	1.034593	1.030931	1.038265	< 0.001***
HGB	0.01966	0.00121	16.252	1.019853	1.017441	1.022277	< 0.001***
PLT	0.002588	0.000249	10.401	1.002592	1.002102	1.003081	< 0.001***
ApoB	0.4143	0.05413	7.654	1.513337	1.361002	1.6827	< 0.001***
TB	0.00778	0.002057	3.782	1.00781	1.003739	1.011865	< 0.001***

Supplementary Table 5. Cox regression for risk factors of NAFLD

	se(coef)	z value	coef	HR	lower .95	upper .95	Pr(>|z|)
BMI	0.005327	35.102	0.186996	1.205622	1.1930993	1.2182761	< 0.001***
Albumin	0.009655	4.656	0.04495	1.045976	1.0263691	1.0659575	< 0.001***
WGR	0.096408	1.326	0.12788	1.136417	0.9407519	1.3727778	0.184691
WBC	0.010893	4.087	0.044521	1.045527	1.0234412	1.0680884	< 0.001***
TG	0.011283	11.755	0.132633	1.141831	1.1168563	1.1673638	< 0.001***
HDL	0.071682	-13.214	-0.94719	0.387829	0.3369955	0.4463302	< 0.001***
GLT	0.000494	2.601	0.001285	1.001285	1.0003166	1.0022553	0.009302**
ALT	0.001478	9.309	0.013757	1.013853	1.0109201	1.0167934	< 0.001***
AST	0.003074	-4.965	-0.01526	0.984853	0.978937	0.990805	< 0.001***
Cr	0.001677	-5.609	-0.00941	0.990637	0.9873861	0.9938985	< 0.001***
ALP	0.000871	1.451	0.001263	1.001264	0.9995565	1.0029745	0.146888
UA	0.000277	11.14	0.003085	1.003089	1.0025451	1.0036338	< 0.001***
Glu	0.016397	3.771	0.061837	1.063788	1.0301446	1.0985309	0.000162***
SBp	0.001643	1.485	0.00244	1.002443	0.9992205	1.0056749	0.137485
DBp	0.002517	2.383	0.005996	1.006014	1.0010643	1.0109889	0.01719*
ESR	0.003054	3.429	0.010473	1.010528	1.0044965	1.0165956	0.000606***
HGB	0.002056	4.507	0.009269	1.009312	1.0052523	1.0133885	< 0.001***
PLT	0.000389	3.318	0.001291	1.001292	1.0005286	1.0020554	0.000905***
ApoB	0.092419	11.569	1.069192	2.913026	2.4303962	3.4914963	< 0.001***
TB	0.003316	-0.216	-0.00072	0.999283	0.9928097	1.0057994	0.828863

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	
Objectives	3	State specific objectives, including any prespecified hypotheses	
Methods			
Study design	4	Present key elements of study design early in the paper	
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	
Bias	9	Describe any efforts to address potential sources of bias	
Study size	10	Explain how the study size was arrived at	
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	
		(*b*) Describe any methods used to examine subgroups and interactions	
		(*c*) Explain how missing data were addressed	
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	
		(*e*) Describe any sensitivity analyses	

Continued on next page

Results			
Participants	13 *	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	
		(b) Give reasons for non-participation at each stage	
		(c) Consider use of a flow diagram	
Descriptive data	14 *	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	
		(b) Indicate number of participants with missing data for each variable of interest	
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	
Outcome data	15 *	*Cohort study*—Report numbers of outcome events or summary measures over time	
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	
		Cross-sectional study—Report numbers of outcome events or summary measures	
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	
		(*b*) Report category boundaries when continuous variables were categorized	
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	
Discussion			
Key results	18	Summarise key results with reference to study objectives	
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	
Generalisability	21	Discuss the generalisability (external validity) of the study results	
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Clinical outcome of nonalcoholic fatty liver disease: an eleven-year follow-up study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-054891.R1
Article Type:	Original research
Date Submitted by the Author:	17-Mar-2022
Complete List of Authors:	Tang, Xiaoping ; Ningbo Zhenhai District Lianhua Hospital , Department of Prevention and Health Care Shi, Yanyan; Peking University Third Hospital, Research Center of Clinical Epidemiology Du, Juan ; Ningbo Zhenhai District Lianhua Hospital , Department of Internal Medicine Hu, Keming; Ningbo Zhenhai District Lianhua Hospital, Department of Prevention and Health Care Zhou, Tingting; Ningbo Zhenhai District Lianhua Hospital , Department of Internal Medicine Chen, Lan; Ningbo Zhenhai District Lianhua Hospital , Department of Radiology Zhang, Yanming; Ningbo Zhenhai District Lianhua Hospital , Center Laboratory Li, Fujun; Ningbo Zhenhai District Lianhua Hospital , Department of Radiology Zhang, Huier ; Ningbo Zhenhai District Lianhua Hospital , Center Laboratory Liebe, Roman ; Heinrich Heine University Düsseldorf, Clinic of Gastroenterology, Hepatology and Infectious Diseases,; Saarland University, Department of Medicine II, Saarland University Medical Center Meyer, Christoph ; Heidelberg University, Department of Medicine II, Medical Faculty Mannheim Dooley, Steven ; Heidelberg University, Department of Medicine II, Medical Faculty Mannheim Zhu, Zhongwei; Ningbo Zhenhai District Lianhua Hospital , Department of Surgery Weng, Hong-Lei ; Heidelberg University, Department of Medicine II, Medical Faculty Mannheim JIA, Jinzhu; Peking University, Department of Biostatistics, School of Public Health; Peking University, Center for Statistical Science Huang, Tong; Ningbo City First Hospital; Ningbo Zhenhai District Lianhua Hospital , Department of Prevention and Health Care
Primary Subject Heading:	Infectious diseases
Secondary Subject Heading:	Infectious diseases
Keywords:	Epidemiology < INFECTIOUS DISEASES, Hypertension < CARDIOLOGY,

	DIABETES & ENDOCRINOLOGY, Health & safety < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Health informatics < BIOTECHNOLOGY & BIOINFORMATICS

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Clinical outcome of nonalcoholic fatty liver disease: an eleven-year follow-up study

Xiaoping Tang[1#], Yanyan Shi[2#], Juan Du[3#], Keming Hu[1], Tingting Zhou[3], Lan Chen[4], Yanming Zhang[5], Fujun Li[4], Huier Zhang[5], Roman Liebe[6,7], Christoph Meyer[8], Steven Dooley[8], Zhongwei Zhu[9], Hong-Lei Weng[8], Jinzhu Jia[10,11]*, Tong Huang[1,12]*

[1] Department of Prevention and Health Care, Ningbo Zhenhai Lianhua Hospital, Ningbo, China

[2] Research Center of Clinical Epidemiology, Peking University Third Hospital, Beijing, China

[3] Department of Internal Medicine, Ningbo Zhenhai Lianhua Hospital, Ningbo, China

[4] Department of Radiology, Ningbo Zhenhai Lianhua Hospital, Ningbo, China

[5] Center Laboratory, Ningbo Zhenhai Lianhua Hospital, Ningbo, China

[6] Clinic of Gastroenterology, Hepatology and Infectious Diseases, Heinrich Heine University, Düsseldorf, Germany;

[7] Department of Medicine II, Saarland University Medical Center, Saarland University, Homburg, Germany;

[8] Department of Medicine II, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany;

[9] Department of Surgery, Ningbo Zhenhai Lianhua Hospital, Ningbo, China

[10]Department of Biostatistics, School of Public Health, Peking University Health Science Center, Beijing, 100191, P.R. China

[11]Center for Statistical Science, Peking University, Beijing, 100871, P.R. China

[12]Center of health Management, Ningbo No.1 Hospital

# These authors contributed to the project equally.

* Co-senior authors

Corresponding authors:

Tong Huang, MD

Center of health Management, Ningbo No.1 Hospital

Address: No. 59, Liuting Street, Haishu District, Ningbo, Zhejiang Province, 315010, P.R. China

Email: nbbjzxht@163.com

Tel: +86-135868350952

Jinzhu Jia, PhD.
Department of Biostatistics, School of Public Health, Peking University Health Science Center
Address: No. 38, Xueyuan Road, Haidian District, Beijing, 100191, P.R. China
Email: jzjia@math.pku.edu.cn
Tel: +86-15801049187

Acknowledgements:

This project is supported by the National Key R&D Program of China No. 2017YFC0908103 (X.T., and J.D.). We are grateful to Drs. Changxi Chen, Jingyi Yuan and Yongjun Chen for support and discussion.

Abstract

Objectives To clarify NAFLD prevalence, risk factors, and clinical outcome in an exemplary Chinese population, a cohort of company employees was followed up for eleven years.

Design Retrospective cohort

Setting Between 2006-2016 in China

Participants 13032 company employees

Results Over eleven years, the prevalence of NAFLD increased from 17.2% to 32.4% (males 20.5% to 37% *versus* females 9.8% to 22.2%). Male peak prevalence was between 40 and 60 years of age, whereas highest prevalence in women was at an age of 60 years and older. Logistic and Cox regression revealed 16 risk factors, including BMI, albumin, WBC, TG, HDL, GLT, ALT, Cr, UA, Glu, SBp, DBp, ESR, HGB, PLT and ApoB. The AUC of these variables for NAFLD is 0.88. However, cause-effect analyses showed that only BMI, gender and triglycerides directly contributed to NAFLD development. Over an 11-year follow-up period, 12.6%, 37.7% and 14.2% of male NAFLD patients and 11.6%, 44.7% and 22.6% of female NAFLD patients developed diabetes, hypertension and hyperuricemia, respectively. Except one male patient who developed cirrhosis, no NAFLD patients progressed into severe liver disease.

Conclusion Diabetes, hypertension and hyperuricemia are the main clinical outcomes of NAFLD. Eleven years of NAFLD are not sufficient to cause severe liver disease. Age and obesity are direct risk factors for NAFLD. BMI, gender and triglycerides are three parameters directly reflecting the occurrence of NAFLD.

(Words: 221)

Strengths and limitations of this study

- This study dynamically follows up NAFLD prevalence in an eastern Chinese community for eleven years.
- The study adopted First order Markov models to evaluate the cause-effect link between NAFLD and risk factors.
- The relatively low sensitivity of ultrasound for the detection of liver fat might underestimate the incidence of NAFLD in this cohort.
- Given that the current study is a single-center observation, multiple-center studies are required to confirm the conclusions in the future.

- The study population is a highly select, relatively homogenous group of well-educated professionals in privileged social positions and permanent employment. Thus, the conclusions might not be transferable to the general Chinese population.

Introduction

Non-alcoholic fatty liver disease (NAFLD) is one of the most common causes of chronic liver disease globally [1]. The global prevalence of NAFLD is currently around 25% [2,3]. NAFLD is predicted to become the most frequent indication for liver transplantation by 2030 in Western countries [4]. An analysis based on 18 million patients in four European cohorts showed that NAFLD and non-alcoholic steatohepatitis (NASH) increase the risk of end-stage liver diseases, e.g., cirrhosis and hepatocellular carcinoma (HCC) [5]. Of note, NAFLD is not only a disease restricted to the liver, but also affects extra-hepatic organs. NAFLD is tightly associated with the occurrence of type 2 diabetes mellitus (T2DM), cardiovascular (CVD) and cardiac diseases, and chronic kidney disease (CKD) [4].

In China, the incidence of NAFLD has been increasing over the last two decades. A recent meta-analysis, based on 392 studies between 2008 and 2018, showed the national incidence of NAFLD in China to be at 29.2% [6]. In Shanghai, the adult incidence of NAFLD has increased from 14.04% in 1995 to 43.65% in 2015 [2]. Being a vast country, Chinese living in different areas vary widely in lifestyle and economic status. Thus, the epidemiology, natural history and clinical outcomes of NAFLD in different areas of the country are worth further investigation.

It is well accepted that viral hepatitis is a major reason for progressive chronic liver diseases, e.g., fibrosis, cirrhosis and ultimately, hepatocellular carcinoma (HCC). With regard to NAFLD, incidence and severity of associated chronic liver disease outcomes has not been monitored in large Chinese cohorts yet – especially over an extended time span. The current study therefore describes the prevalence of NAFLD in a large Eastern Chinese community over eleven years (2006 - 2016). We focused on three aspects: (1) annual prevalence of NAFLD, (2) risk factors of NAFLD, and (3) intra- and extra-hepatic clinical outcomes of NAFLD.

Methods

Patient and public involvement

No patients were involved in this study

Design and participants

In this retrospective study we analyzed the "annual health examination database" of the Zhenhai Lianhua Hospital from 2006 to 2016. This hospital is affiliated to Sinopec Zhenhai Refining & Chemical Company. Supported by the company, all employees were offered the opportunity to

go to this hospital for an annual health examination. During 11 years, a total 13,032 employees received health examinations. From 2006 to 2016, 11689, 11706, 11584, 9521, 9592, 9725, 9710, 9869, 9718, 9702 and 9706 persons received health examinations (**Figure 1**). To describe the longitudinal NAFLD occurrence in this cohort, we excluded subjects with the following conditions: (1) viral hepatitis B and C infection, which were identified by blood virus measurements (HBV-DNA and HCV-RNA), and (2) alcoholic liver disease, which was defined as previously described [7,8]. NAFLD was defined as the presence of hepatic steatosis, determined by ultrasonography.

The study protocol was approved by the Ethics Committee of Zhenhai Lianhua Hospital ([2016]001). Informed consent was obtained from all subjects.

Measures

Supplementary Table 1 shows all parameters measured in the annual health examinations. Ultrasonography was performed by the same three experienced doctors (L.C., F. L., and J.Y.) with an Ultrasongraph B, GE, Voluson 730 pro. Blood biochemistry and HBV serum levels were measured by an Olympus AU640 autoanalyzer (Olympus, Kobe, Japan) and an ImmunoAssay Analyzer VitrosECI (JOHNSON-JOHNSON, USA), respectively. All methods were carried out in accordance with relevant guidelines and regulations.

Statistical analysis

For population characteristics, variables were described as means and standard deviation (SD) or proportions as appropriate. Student's t-test or nonparametric test was used to analyze differences between two groups as mentioned. Chi-square test was used to verify the differences of nominative variables between two groups. Multivariate analysis of risk factors for NAFLD was performed using logistic regression analysis. Combined receiver operating characteristic curve (ROC) and area under curve (AUC) analyses were used to evaluate the diagnostic performance of biomarkers based on the logistic regression model. Multivariate Cox regression model was performed to calculate hazard ratios of variables to identify independent prognostic variables. First order Markov models were used to analyze the cause-effect link between NAFLD and risk factors. L1 penalized logistic regression was applied to select predictive predictors. R package "glmnet" contains functions to select predictors using L1 penalized logistic regression. Statistical analyses were performed using SPSS 22.0 and R version 3.5.3. *P*-values that were less than 0.05 were considered statistically significant. Figures were generated by R package such as 'forestplot', 'ROCR', 'bnlearn', or 'survival'.

Results

Prevalence of NAFLD from 2006 to 2016

We retrospectively analyzed 9786, 9852, 9827, 8026, 8225, 8309, 8311, 8552, 8442, 8463 and 8436 persons who received health examinations from 2006 to 2016. **Supplementary Table 2** shows the eleven-year annual NAFLD prevalence in this population. In 2006, NAFLD was diagnosed in 17.2% of persons, and gradually increased over the examination period to 19% (2007), 22% (2008), 22.4% (2009), 22.7% (2010), 23.4% (2011), 24.5% (2012), 25.4% (2013), 27.9% (2014), 30.8% (2015), and 32.4% (2016), respectively (**Supplementary Table 2**). Both males and females demonstrated continuously increasing NAFLD prevalence (**Supplementary Table 2**). Compared to females, male Chinese demonstrated significantly higher NAFLD prevalence, e.g., in 2006, the prevalence of NAFLD in males and females was 20.5% and 9.8%, respectively. Eleven years later, the prevalence had increased to 37% in males and 22.2% in females (**Supplementary Table 2**). Noteworthy, the prevalence of NAFLD in male and female was correlating with age. The peak prevalence of NAFLD in men emerged in those aged between 40 and 60 years. In 2006, the prevalence of NAFLD in men aged between 40 - 50 and 50 - 60 years was 24.2% and 24.6%, respectively. In 2016, prevalence reached 42.8% and 46.6% (**Supplementary Table 2**) for men. Distinct from males, the peak NAFLD prevalence in females emerged at an age above 60 years. In 2006, the prevalence of NAFLD in women older than 60 and 70 years was 25.6% and 22.9%, respectively. Eleven years later, these values had increased to 53.4% and 30.9% (**Supplementary Table 2**).

Among the observed population, 5606 persons received annual health examinations for 11 years, and thus prevalence of NAFLD was analyzed in these individuals. As shown in **Table 1**, the prevalence of NAFLD increased from 17% in 2006 to 35.2% in 2016. The highest prevalence rates for NAFLD in 3795 men emerged in those older than 40 years. In 2006, the NAFLD prevalence in males aged between 40 - 50, 50 – 60, and 60 – 70 years was 24.2%, 25.4% and 19.8%, respectively. In 2016, these values reached 43.3%, 46.8% and 42.3% (**Table 1**). Different from males, the peak NAFLD prevalence in 1811 females emerged at an age of more than 60 years. In 2006, the prevalence of NAFLD in women older than 60 or 70 years was 22.1% and 28%, respectively. Eleven years later, these values had increased to 31.8% and 28.6% (**Table 1**).

BMI and NAFLD incidence

Given the tight link between obesity and NAFLD, we paid special attention to the population with high BMI. We focused on the 5606 persons with complete follow-up and analyzed the prevalence of NAFLD in those with BMI >25. In total, out of the 5606 persons, 2445 presented with a BMI of >25. The prevalence of NAFLD in this overweight subpopulation was far higher than in the general population. In 2006, 45.2% of individuals (n=1104; male *vs.* female: 47.3% *vs.* 37.1%) with BMI >25 were suffering from NAFLD (**Supplementary Table 3**). In 2016, values reached 67.1% (n=1414; male *vs.* female: 69% *vs.* 59.2%, **Supplementary Table 3**). Impressively, the NAFLD prevalence in both genders was very high at any age, even in those below the age of 30 years. In 2006, among 213 overweight men, younger than 30 years, 52.6% were also diagnosed for NAFLD (**Supplementary Table 3**). This number increased to 63% in 2016 (**Supplementary Table 3**). In 2006, there were 15 overweight women aged less than 30 years. Among them, 3 presented as NAFLD (20%). In 2016, 7 out of 16 overweight women aged less than 30 years were identified. The NALFD prevalence had increased to 43.8% (**Supplementary Table 3**). In those older than 40 years, NAFLD prevalence increased from 36.6 – 45.4% in 2006 to 53 – 65.6% (**Supplementary Table 3**).

Risk factors relevant to NAFLD occurrence

Next, we analyzed risk factors relevant to NAFLD occurrence. Logistic regression analysis was performed on 26 parameters, including gender, age, BMI, albumin, white globulin ratio (WGR), white blood cell count (WBC), low density lipoprotein (LDL), triglycerides (TG), high density lipoprotein (HDL), glutamyl transpeptidase (GLT), alanine transaminase (ALT), aspartate transaminase (AST), creatinine (Cr), alkaline phosphatase (ALP), blood urea nitrogen (Bun), uric acid (UA), blood glucose (Glu), systolic blood pressure (SBp), diastolic blood pressure (DBp), Blood sedimentation (ESR), hemoglobin (HGB), platelets (PLT), Apolipoprotein A1 (ApoA1), Apolipoprotein B2 (ApoB), total bilirubin (TB), total protein (TP). We performed variable selection by penalized Logistic regression using R package glmnet. Cross validation selected 16 variables as potential predictors. These were BMI, albumin, WBC, TG, HDL, GLT, ALT, Cr, UA, Glu, SBp, DBp, ESR, HGB, PLT and ApoB (**Supplementary Table 4**). The corresponding forest-plot is shown in **Figure 2A**. Among these variables, ApoB and BMI displayed the most robust positive correlation with NAFLD occurrence, while HDL had a strong negative correlation with NAFLD incidence (**Supplementary Table 4**). The AUC of these variables for NAFLD is 0.88 (see ROC curve in **Figure 2B**). We further performed a time dependent Cox regression to calculate the hazard ratios of these parameters for NAFLD

occurrence. Cox regression confirmed that the 16 parameters were significantly relevant to NAFLD incidence (**Supplementary Table 5** and **Figure 2C**). Furthermore, ApoB and HDL were the most robust positive and negative risk factors for NAFLD (**Figure 2C**).

Cause-effect link between risk factors and NAFLD occurrence

Although the aforementioned parameters were regarded as "risk factors" according to statistical models, it did not necessarily mean that all of them contributed to NAFLD occurrence. Based on 11 years of longitudinal data, it was possible to construct a dynamic Bayesian network to identify the risk factors most relevant to NAFLD occurrence. As shown in **Figure 3**, these parameters constituted a complicated, but clear intercross paradigm. Only three parameters, BMI, gender and TG directly pointed to NAFLD. In addition, ApoB impacted the incidence of NAFLD through contributing to TG. Furthermore, LDL can indirectly contribute to NAFLD through influencing ApoB. Very impressively, the dynamic Bayesian network pointed out that NAFLD directly leads to alterations of seven parameters: ALT, DBp, SBp, TP, albumin, HGB and UA. Intriguingly, age, GLT, AST, Cr, and BUN did not interact with any other parameter in our model, indicating that these factors by incidence correlate, but not any causal interaction is existing.

Outcome of NAFLD

Subsequently, we examined clinical outcomes of NAFLD over the eleven years. **Table 2** summarizes the incidence of intra- and extrahepatic diseases of 696 NAFLD and 222 NASH patients during the follow-up period. Among the total NAFLD and NASH population, only 1 male NAFLD patient developed liver cirrhosis within the eleven years. However, this time span witnessed significantly increased extrahepatic diseases, including diabetes, hypertension and hyperuricemia. In the NAFLD population, there were 64 (12.6%) men and 22 (11.6%) women, 191 (37.7%) men and 85 (44.7%) women, 72 (14.2%) men and 43 (22.6%) women who developed into type II diabetes, hypertension and hyperuricemia, respectively (**Table 2**).

Given that 2006 is the starting point of data collection, patients diagnosed for NAFLD in this year had by no means just manifested their disease, but rather patients had possibly developed NAFLD several years prior to inclusion. To clarify the exact clinical outcomes of NAFLD over one decade, we focused on the following two cohorts of individuals with annual health examinations for 11 years: (1) patients who were diagnosed as non-NAFLD in 2006, but were NAFLD in 2007 (new NAFLD cohort); and (2) who were non-NAFLD in both 2006 and 2007 (non-NAFLD cohort). As shown in **Table 2**, 185 new NAFLD cases (138 men and 47 women)

and 4547 non-NAFLD (2786 men and 1761 women) persons were found in 2007. Between 2007 and 2016, neither NAFLD nor non-NAFLD individuals developed liver cirrhosis or cancer. However, the one-decade follow-up reveals different prevalences of diabetes, hypertension and hyperuricemia: In NAFLD patients, there were 14 (10.1%) men and 5 (10.6%) women, 47 (34.1%) men and 21 (44.7%) women, 34 (24.6%) men and 8 (17%) women who developed type II diabetes, hypertension and hyperuricemia, respectively (**Table 2**). In non-NAFLD individuals, 157 (5.6%) men and 54 (3.4%) women, 259 (9.3%) men and 324 (18.4%), 284 (10.2%) men and 84 women (4.8%) developed type II diabetes, hypertension and hyperuricemia, respectively (**Table 2**). For all three diseases, statistically significant differences were determined between the two cohorts of population (all P <0.05, **Table 2**). These results suggest that diabetes, hypertension and hyperuricemia are the main clinical outcomes of NAFLD.

Discussion

This eleven-year follow-up retrospective study reports the following: (1) NAFLD prevalence has substantially increased in the examined Eastern Chinese population. (2) The prevalence of NAFLD differs by gender and age. Middle-aged men and elderly women are the two populations at highest risk for NAFLD. (3) Gender, BMI and triglycerides are the parameters directly associated with NAFLD occurrence. Regardless of gender and age, persons with high BMI (≥25) have a high risk for NAFLD development. (4) NAFLD directly leads to alterations of seven clinical parameters: ALT, DBp, SBp, TP, albumin, HGB and UA. (5) Within 11 years, a significant part of the NAFLD population develops three clinically relevant diseases: type 2 diabetes mellitus, hypertension and hyperuricemia. (6) Within 11 years, NAFLD does not cause severe liver disease, such as cirrhosis or HCC, in patients.

The most impressive observation of the current study is that among 918 diseased persons (696 NAFLD and 222 NASH), no patient progressed towards HCC and only 1 male NAFLD patient developed liver cirrhosis within the 11 years. Furthermore, among 185 new NAFLD cases diagnosed in 2007, none developed liver cirrhosis or liver cancer. Liver cirrhosis and HCC are commonly regarded as the most severe and costly clinical outcomes of NAFLD [9]. In USA and Europe, it is estimated that 10-15% of NAFLD patients develop advanced fibrosis [10]. In China, a study with biopsy-proven NAFLD revealed 1.97%-2.97% cirrhosis prevalence [11]. In addition, NAFLD is regarded as the third-most common cause of cancer-related death worldwide [12]. In a study based on 4949 US patients with HCC, 701 patients had NAFLD [13].

It was estimated that the cumulative incidence of HCC among patients with NAFLD and cirrhosis ranges from 2.4% to 12.8% over a median follow-up period of 3.2 to 7.2 years [14] (Global Health Observatory) data. Mortality and global health estimates were obtained from: http://www.who.int/gho/mortality_burden_disease/en/. Last accessed on 1/7/2020.). Given that the above conclusions were based on cross-sectional investigations and statistical models, it has been unknown to date over which period a NAFLD patient develops liver cirrhosis or HCC (personal risk assessment). Our 11-year follow-up provides therefore a valuable and comprehensive dataset. In this study, most patients were diagnosed with NAFLD when they received a routine health examination. Before the examination, these people did not have any symptoms or signs of NAFLD. Therefore, they belong to NAFLD patients at a very early stage (although for 2006, the duration of pre-existing NAFLD cannot be determined). Except for a single person, no serious liver problems were observed within this time period. These data suggest that for the vast majority of patients with early stage NAFLD, 11 years are not sufficient to develop liver cirrhosis or cancer. Nasr et al followed up 129 NAFLD patients with varying fibrosis stages on two occasions (mean time 13.7 and 9.3 years). Liver biopsy analyses showed that 9.3% of patients developed end-stage liver disease and 34% advanced fibrosis [15]. The NAFLD patients observed by Nasr et al. actually belonged to the NASH category, because they suffered from fibrosis and elevated ALT and/or AST levels. As our study was based on examinations of healthy, liver biopsy is not justifiable. Very likely, the current cohort included a portion of NASH patients. They also did not show significant progression towards cirrhosis or HCC was monitored.

In contrast to hepatic complications, patients with NAFLD showed a significant risk for the development of extrahepatic diseases, including diabetes, hypertension and hyperuricemia. In 696 NAFLD patients, eleven years witnessed the development of 86 cases (12.4%) type II diabetes, 276 (40%) cases of hypertension and 115 patients with (16.5%) hyperuricemia, respectively. Interestingly, in 222 NASH patients, the prevalence of these three diseases was 12 (5.4%), 46 (20.7%), and 33 (14.9%) only. In general, men had a higher probability to develop these diseases than women. These results are consistent with previous reports from USA and Europe [16-18]. Whether NAFLD is associated with the risk of severe heart or brain diseases such as acute myocardial infarction (AMI) and stroke is worth further investigation. A recent matched cohort study analyzed databases from four European countries, which included 17.7 million patients with NAFLD or NASH [19]. These patients had a mean follow-up of 2.1 to 5.5 years. The study showed that the diagnosis of NAFLD appears not to be associated with AMI

or stroke risk after adjustment for established cardiovascular risk factors. Nevertheless, the authors mentioned that cardiovascular risk assessment in adults with a diagnosis of NAFLD is important [19]. Follow-up for 5 years might be not sufficient to reach a conclusion for this issue.

An important issue is the cause-and-effect relationship between NAFLD and its clinical outcomes such as diabetes, hypertension and hyperuricemia. A dynamic Bayesian network in the current study provides direct evidence on this issue: NAFLD directly results in alterations of several parameters, including DBp, SBp and UA, suggesting that NAFLD directly contributes to the occurrence of hypertension and hyperuricemia. The underlying mechanisms require further investigation.

The current dynamic Bayesian network analysis does not confirm a direct cause-and-effect relationship between NAFLD and type 2 diabetes mellitus. There are plenty of studies showing the close relationship between type 2 diabetes and NALFD [20]. Pathophysiologically, insulin resistance is a key event in both NAFLD and diabetes progression [20]. However, genome-wide association studies have not yet identified the exact impact of insulin resistance on the variants associated with NAFLD severity [20,21]. Clarification of the cause-and-effect relationship between NAFLD and diabetes requires further long-term follow-up studies.

To date, there are a large number of studies investigating risk factors for NAFLD [22]. These studies tried to identify single, or multiple combined biomarkers to predict NAFLD occurrence. Given that most studies were based on cross-sectional designs, or with only short follow-up periods, it is difficult to clarify the causality between the proposed predictors and NAFLD morbidity. Our eleven-year dataset provides a chance to shed led on this issue. Here, the dynamic causal relationships between variables, including risk parameters and clinical outcomes, were identified by a first order Markov model, which was displayed by a dynamic Bayes network. The dynamic Bayes model discriminates causal relationship through time sequence. When a variable change is closely related to a previous variance alteration, a causal relationship between the two variables is assumed. Based on Logistic and Cox regression and dynamic Bayesian network analyses, we confirmed three direct risk factors for NAFLD occurrence: gender, BMI and TG. These findings are supported by the following data: (1) Men have higher NAFLD prevalence than women in this population (37% *versus* 22.2% in 2016); (2) In overweight people with a BMI >25, NAFLD prevalence reached 69% in males and 59.2% in females. Given that triglycerides are a major energy source, but are leading to obesity, it is

not surprising that this parameter directly reflects the risk for NAFLD development. These findings provide robust evidence supporting the use of BMI to monitor or predict NAFLD.

Conclusion

This 11-year follow-up study documents the rapid increase in NAFLD prevalence in an Eastern Chinese population. In contrast to previous reports, our observation does not observe that one decade of NAFLD is sufficient to lead to severe hepatic clinical outcomes. It is worthy to note that our population were biased because they are on the well-off, well-educated side of the Chinese people, while previous studies were often based on hospital populations, who suffered from negative selection bias and thus came up with higher estimates. In addition, given there are differences in NAFLD profiles between Eastern and Western populations, it would be interesting to know the natural development of NAFLD in a Western population. A key point for clarifying the true history of NAFLD is to follow a population starting from the early phases of the disease. Consistent with previous studies, NAFLD is tightly associated with multiple extra-hepatic diseases relevant to the metabolic syndrome. In the future, follow-up of the current cohort for another one and two decades will provide further valuable data to clarify the extended natural history of NAFLD. Last but not least, a large portion of the men and women in this study were educated above the average and have a position in the company that gave them the availability of better food choices as well as regular sport. On the other hand, the relatively low sensitivity of ultrasound for the detection of liver fat might underestimate the incidence of NAFLD in this cohort.

Contributorship:

Conception and design: Xiaoping Tang, Yanyan Shi, Juan Du, Hong-Lei Weng, Jinzhu Jia and Tong Huang

Ultrasonography: Lan Chen, Fujun Li

Blood assays: Yanming Zhang and Huier Zhang

Other examinations and data collection: Xiaoping Tang, Keming Hu, Tingting Zhou, Juan Du, Zhongwei Zhu and Tong Huang

Statistical analyses: Yanyan Shi and Jinzhu Jia

Drafting the article: Hong-Lei Weng

Reviewing and editing the article critically: Xiaoping Tang, Yanyan Shi, Christoph Meyer, Roman Liebe, Steven Dooley, Hong-Lei Weng, Jinzhu Jia and Tong Huang

Funding Statement

The project is supported by National Key R&D Program of China (2017YFC0908103).

Competing of Interests

The authors declare that there is no conflict of interest.

Ethics approval

The study protocol was approved by the Ethics Committee of Zhenhai Lianhua Hospital No. 2016(001). Informed consent was obtained from all subjects.

Data sharing

All data relevant to the study are included in the article or uploaded as supplementary information.

References

[1] Younossi ZM, Tampi R, Priyadarshini M, et al. Burden of Illness and Economic Model for Patients With Nonalcoholic Steatohepatitis in the United States. Hepatology, 2019, 69: 564-572.

[2] Younossi ZM, Loomba R, Anstee QM, et al. Diagnostic modalities for nonalcoholic fatty liver disease, nonalcoholic steatohepatitis, and associated fibrosis. Hepatology, 2018, 68: 349-360.

[3] Fan JG, Kim SU, Wong VW. New trends on obesity and NAFLD in Asia. J Hepatol, 2017, 67: 862-873.

[4] Byrne CD, Targher G. NAFLD: a multisystem disease. J Hepatol, 2015, 62: S47-64.

[5] Alexander M, Loomis AK, van der Lei J, et al. Risks and clinical predictors of cirrhosis and hepatocellular carcinoma diagnoses in adults with diagnosed NAFLD: real-world study of 18 million patients in four European cohorts. BMC Med, 2019, 17: 95.

[6] Zhou F, Zhou J, Wang W, et al. Unexpected Rapid Increase in the Burden of NAFLD in China From 2008 to 2018: A Systematic Review and Meta-Analysis. Hepatology, 2019, 70: 1119-1133.

[7] EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. J Hepatol, 2016, 64: 1388-1402.

[8] Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of nonalcoholic fatty liver disease: Practice guidance from the American Association for the Study of Liver Diseases. Hepatology, 2018, 67: 328-357.

[9] Younossi Z, Tacke F, Arrese M, et al. Global Perspectives on Nonalcoholic Fatty Liver Disease and Nonalcoholic Steatohepatitis. Hepatology, 2019, 69: 2672-2682.

[10] Younossi ZM, Koenig AB, Abdelatif D, et al. Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. Hepatology, 2016, 64: 73-84.
[11] Shen F, Zheng RD, Shi JP, et al. Impact of skin capsular distance on the performance of controlled attenuation parameter in patients with chronic liver disease. Liver Int, 2015, 35: 2392-2400.
[12] Ascha MS, Hanouneh IA, Lopez R, et al. The incidence and risk factors of hepatocellular carcinoma in patients with nonalcoholic steatohepatitis. Hepatology, 2010, 51: 1972-1978.
[13] Younossi ZM, Otgonsuren M, Henry L, et al. Association of nonalcoholic fatty liver disease (NAFLD) with hepatocellular carcinoma (HCC) in the United States from 2004 to 2009. Hepatology, 2015, 62: 1723-1730.
[14] White DL, Kanwal F, El-Serag HB. Association between nonalcoholic fatty liver disease and risk for hepatocellular cancer, based on systematic review. Clin Gastroenterol Hepatol, 2012, 10: 1342-1359.e1342.
[15] Nasr P, Ignatova S, Kechagias S, et al. Natural history of nonalcoholic fatty liver disease: A prospective follow-up study with serial biopsies. Hepatol Commun, 2018, 2: 199-210.
[16] Estes C, Razavi H, Loomba R, et al. Modeling the epidemic of nonalcoholic fatty liver disease demonstrates an exponential increase in burden of disease. Hepatology, 2018, 67: 123-133.
[17] Williamson RM, Price JF, Glancy S, et al. Prevalence of and risk factors for hepatic steatosis and nonalcoholic Fatty liver disease in people with type 2 diabetes: the Edinburgh Type 2 Diabetes Study. Diabetes Care, 2011, 34: 1139-1144.
[18] Targher G, Bertolini L, Padovani R, et al. Prevalence of nonalcoholic fatty liver disease and its association with cardiovascular disease among type 2 diabetic patients. Diabetes Care, 2007, 30: 1212-1218.
[19] Alexander M, Loomis AK, van der Lei J, et al. Non-alcoholic fatty liver disease and risk of incident acute myocardial infarction and stroke: findings from matched cohort study of 18 million European adults. Bmj, 2019, 367: l5367.
[20] Tilg H, Moschen AR, Roden M. NAFLD and diabetes mellitus. Nat Rev Gastroenterol Hepatol, 2017, 14: 32-42.
[21] Zhou Y, Llauradó G, Orešič M, et al. Circulating triacylglycerol signatures and insulin sensitivity in NAFLD associated with the E167K variant in TM6SF2. J Hepatol, 2015, 62: 657-663.
[22] Miyake T, Kumagi T, Furukawa S, et al. Non-alcoholic fatty liver disease: factors associated with its presence and onset. J Gastroenterol Hepatol, 2013, 28 Suppl 4: 71-78.

Figure legends

Figure 1. Flow chart depicting the enrollment of a population with non-alcoholic fatty liver disease (NAFLD) for follow-up in Ningbo Zhenhai Lianhua Hospital, China.

Figure 2. Penalized logistic regression and Cox regression analysis were performed for risk factors and hazard ratios of non-alcoholic fatty liver disease (NAFLD). The following parameters were available from 5606 participants: Gender, age, BMI, albumin, white globulin

ratio (WGR), white blood cell (WBC), low density lipoprotein (LDL), triglycerides (TG), high density lipoprotein (HDL), glutamyl transpeptidase (GLT), alanine transaminase (ALT), aspartate transaminase (AST), creatinine (Cr), alkaline phosphatase (ALP), blood urea nitrogen (Bun), Uric Acid (UA), blood glucose (Glu), systolic blood pressure (SBp), diastolic blood pressure (DBp), blood sedimentation (ESR), hemoglobin (HGB), platelets (PLT), Apolipoprotein A1 (ApoA1), Apolipoprotein B2 (ApoB), total bilirubin (TB), total protein (TP). Cross validation selected 16 variables to be potential predictors. The corresponding forest-plot is shown in (**A**). The AUC of these above 16 variables for NAFLD is 0.88 **(B)**. Cox regression confirmed that the 16 variables were relevant for NAFLD incidence, including BMI, albumin, WBC, TG, HDL, GLT, ALT, Cr, UA, Glu, SBp, DBp, ESR, HGB, PLT and ApoB. The corresponding forest-plot is shown (**C**).

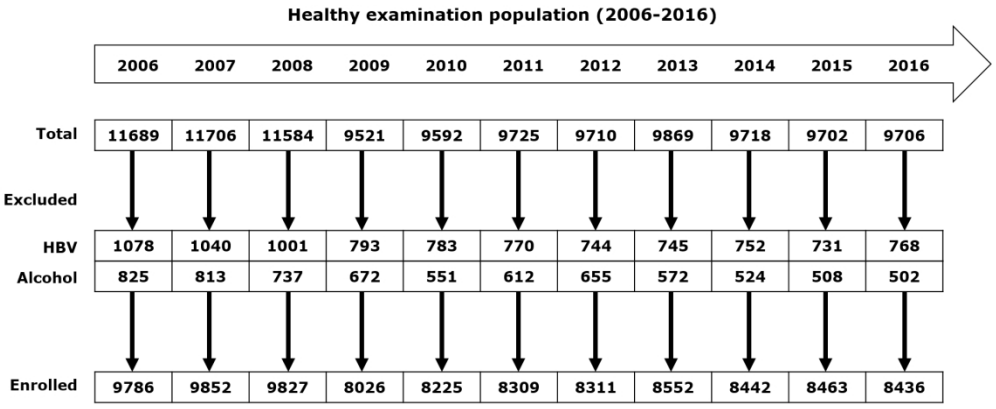
Figure 3. Dynamic Bayesian network analyses were performed to show the cause-effect link between non-alcoholic fatty liver disease (NAFLD) and its potential risk factors. Three variables, BMI, gender and trigylcerides directly pointed to NAFLD. ApoB impacted on the incidence of NAFLD through TG abundance. LDL indirectly contributed to NAFLD through ApoB. NAFLD directly led to alterations of seven clinical parameters: ALT, DBp, SBp, TP, albumin, HGB and UA.

Table 1. Prevalence of NAFLD in 5606 persons with 11-year follow-up (2006-2016)

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Total (n)	5606	5606	5606	5606	5606	5606	5606	5606	5606	5606	5606
NAFLD (n)	951	1068	1247	1322	1378	1466	1524	1574	1700	1883	1976
(%)	17	19.1	22.2	23.6	24.6	26.2	27.2	28.1	30.3	33.6	35.2
Male (n)	3795	3795	3795	3795	3795	3795	3795	3795	3795	3795	3795
NAFLD (n)	776	871	1014	1075	1117	1206	1244	1283	1387	1517	1569
(%)	20.4	23	26.7	28.3	29.4	31.8	32.8	33.8	36.5	40	41.3
≤ 30ys (n)	668	466	366	297	246	204	142	67	14	4	0
NAFLD (n)	97	71	57	46	35	35	30	17	3	1	0
(%)	14.5	15.2	15.6	15.5	14.2	17.2	21.1	25.4	21.4	25	0
>30, ≤ 40ys (n)	1212	1315	1342	1258	1180	1091	1034	981	896	778	668
NAFLD (n)	231	286	349	349	338	330	330	331	331	301	261
(%)	19.1	21.7	26	27.7	28.6	30.2	31.9	33.7	36.9	38.7	39.1
>40, ≤ 50ys (n)	873	863	842	928	988	1022	1097	1128	1170	1201	1212
NAFLD (n)	211	215	239	278	314	353	388	411	454	524	525
(%)	24.2	24.9	28.4	30	31.8	34.5	35.4	36.4	38.8	43.6	43.3
>50, ≤ 60ys (n)	562	623	696	741	776	826	796	798	820	851	873
NAFLD (n)	143	184	231	262	276	314	292	286	320	363	409
(%)	25.4	29.5	33.2	35.4	35.6	38	36.7	35.8	39	42.7	46.8
>60, ≤ 70ys (n)	368	382	373	348	346	356	398	465	496	521	562
NAFLD (n)	73	85	97	95	96	108	125	156	178	210	238
(%)	19.8	22.3	26	27.3	27.7	30.3	31.4	33.5	35.9	40.3	42.3
>70ys	112	146	176	223	259	296	328	356	399	440	480
NAFLD (n)	21	30	41	45	58	66	79	82	101	118	136
(%)	18.8	20.5	23.3	20.2	22.4	22.3	24.1	23	25.3	26.8	28.3
Female (n)	1811	1811	1811	1811	1811	1811	1811	1811	1811	1811	1811
NAFLD (n)	175	197	233	247	261	260	280	291	313	366	407
(%)	9,7	10,9	12,9	13,6	14,4	14,4	15,5	16,1	17,3	20,2	22,5
≤ 30ys (n)	85	44	22	13	9	9	5	2	0	0	0
NAFLD (n)	3	1	1	1	0	0	0	0	0	0	0
(%)	3.5	2.3	4.5	7.7	0	0	0	0	0	0	0
>30, ≤ 40ys (n)	582	578	541	480	403	334	277	213	173	124	85
NAFLD (n)	17	21	25	22	24	18	18	13	15	14	12
(%)	2.9	3.6	4.6	4.6	6	5.4	6.5	6.1	8.7	11.3	14.1
>40, ≤ 50ys (n)	485	469	482	496	541	580	612	638	617	608	582
NAFLD (n)	31	31	35	41	47	54	61	66	65	81	86
(%)	6.4	6.6	7.3	8.3	8.7	9.3	10	10.3	10.5	13.3	14.8
>50, ≤ 60ys (n)	365	400	422	450	461	469	456	452	467	476	485
NAFLD (n)	56	67	85	86	90	88	88	83	88	98	109
(%)	15.3	16.8	20.1	19.1	19.5	18.8	19.3	18.4	18.8	20.6	22.5
>60, ≤ 70ys (n)	244	260	262	267	266	266	280	301	314	337	365
NAFLD (n)	54	62	61	66	69	66	70	81	82	104	116
(%)	22.1	23.8	23.3	24.7	25.9	24.8	25	26.9	26.1	30.9	31.8
>70ys (n)	50	60	82	105	131	153	181	205	240	266	294
NAFLD (n)	14	15	26	31	31	34	43	48	63	69	84
(%)	28	25	31.7	29.5	23.7	22.2	23.8	23.4	26.3	25.9	28.6

Table 2. Clinical outcome of NAFLD patients

2006 - 2016									
		Cirrhosis	HCC	Diabetes n (%)		Hypertension n (%)		Hyperuricemia n (%)	
Male (n=506)		0	0	64 (12.6)		191 (37.7)		72 (14.2)	
Female (n=190)		0	0	22 (11.6)		85 (44.7)		43 (22.6)	
2007 – 2016 (outcome of new NFLAD)									
		Cirrhosis	HCC	Diabetes		Hypertension		Hyperuricemia	
				n (%)	*P*-value	n (%)	*P*-value	n (%)	*P*-value
Male	NAFLD n=138	0	0	14 (10.1)	0.028	47(34.1)	< 0.001	34 (24.6)	< 0.001
	Non-NAFLD n=2786	0	0	157 (5.6)		259 (9.3)		284 (10.2)	
Female	NAFLD n=47	0	0	5 (10.6)	0.014	21 (44.7)	< 0.001	8 (17)	< 0.001
	Non-NAFLD n=1761	0	0	54 (3.1)		324 (18.4)		84 (4.8)	



335x136mm (150 x 150 DPI)

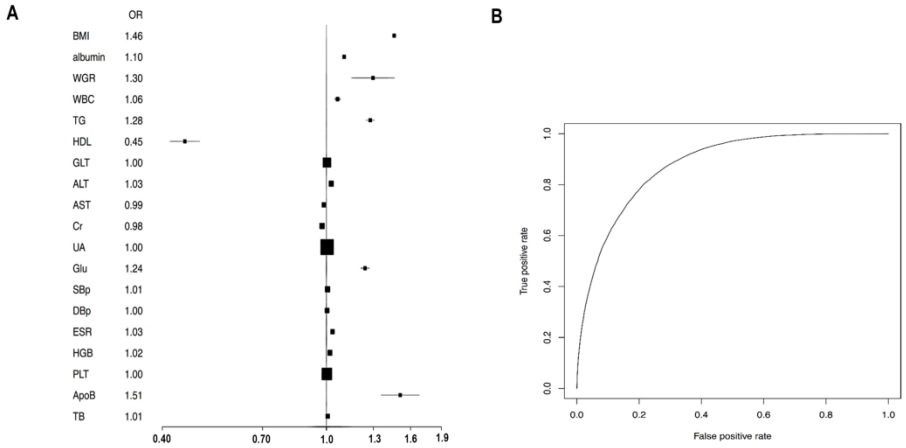
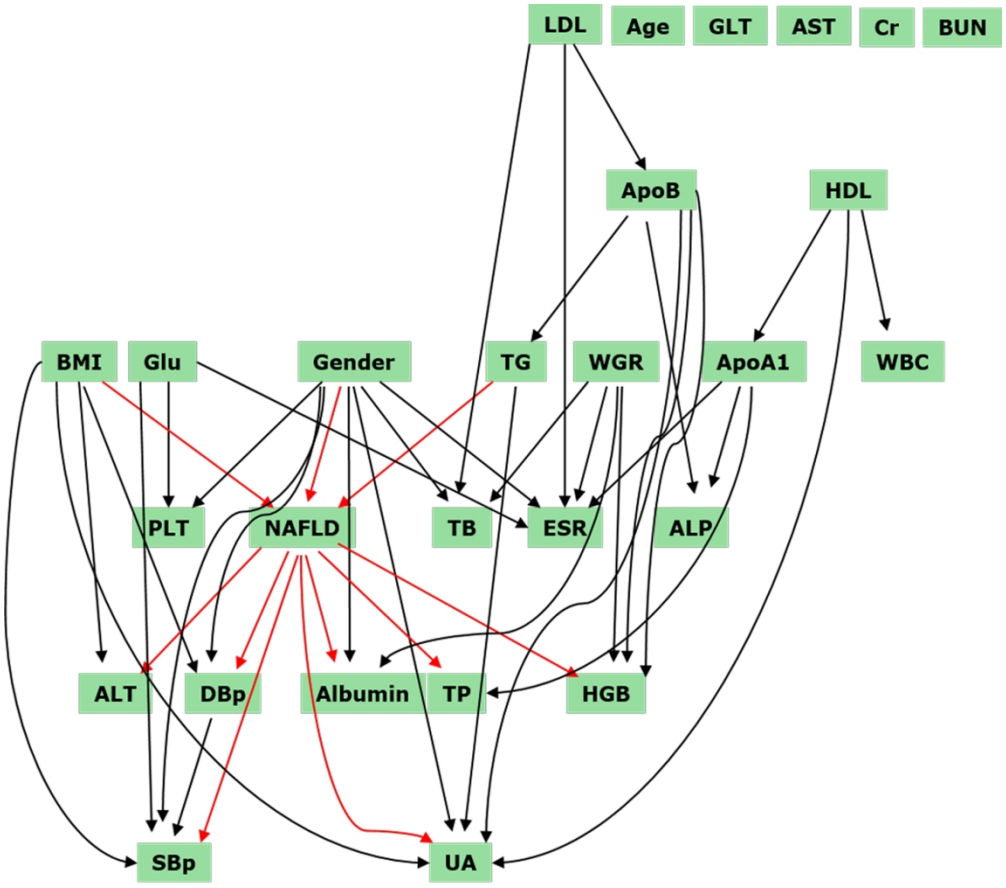


Figure 2

122x63mm (600 x 600 DPI)



522x461mm (96 x 96 DPI)

Nonalcoholic fatty liver disease increased risk of metabolic diseases, but not severe liver disease: an eleven-year follow-up study

Xiaoping Tang, Yanyan Shi, Juan Du, Keming Hu, Tingting Zhou, Lan Chen, Yanming Zhang, Fujun Li, Huier Zhang, Roman Liebe, Christoph Meyer, Steven Dooley, Zhongwei Zhu, Hong-Lei Weng, Jinzhu Jia, Tong Huang

Supplementary Table 1. Parameters measured in the annual health examinations.

Age
Albumin
AKP (alkaline phosphatase)
ALT (alanine transaminase)
ApoA1 (Apolipoprotein A1)
ApoB (Apolipoprotein B2)
AST (aspartate transaminase)
BMI (Body mass index)
BLRV (whole blood low shear reduced viscosity)
BLRI (relative index of whole blood low shear)
BHRV (whole blood high shear reduced viscosity)
BHRI (relative index of whole blood high shear)
BVV200 (Whole blood viscosity value)
BUN (blood urea nitrogen)
BUS (ultrasound prompt)
CRP (high sensitive C-reactive protein)
Cr (creatinine)
CA (carotid atherosclerosis)
DBIL (Direct bilirubin)
DBp (diastolic blood pressure)
DM (type II diabetes)
ESR (Blood sedimentation)
ESRKV (Blood sedimentation equation K value)
Gender
GLT (glutamyl transpeptidase)
Glucose
HBP (Hypertension)
HBX (red blood cell deformation index TK)
HCT (Hematocrit)
HCY (Homocysteine)
HDL (high density lipoprotein C)
Height
HGB (hemoglobin)
LDL (low density lipoprotein C)
LVH (left ventricular hypertrophy)
MPV (mean platelet volume)
NAFLD (non-alcoholic fatty liver disease)
PhyExa (physical examination results)
PV (plasma viscosity)
PDW (Platelet distribution width)
PLT (platelet)
PCT (prothrombin consumption time)
RBC (red blood cell count)
SBp (systolic blood pressure)
TB (Total Bilirubin)
TC (Total cholesterol)
TG (Triglyceride)

TP (total protein)
UA (uric acid)
Waist
Weight
WGR (white globulin ratio)
WBC (white blood cell count)

Supplementary Table 2. Prevalence of NAFLD in an eastern Chinese population (2006-2016)

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Total (n)	9786	9852	9827	8026	8225	8309	8311	8552	8442	8463	8436
NAFLD (n)	1687	1875	2161	1797	1871	1943	2033	2169	2356	2610	2734
(%)	17.2	19.0	22.0	22.4	22.7	23.4	24.5	25.4	27.9	30.8	32.4
Male (n)	6834	6872	6857	5468	5640	5692	5680	5891	5820	5812	5816
NAFLD (n)	1399	1555	1790	1419	1488	1562	1616	1734	1896	2081	2152
(%)	20.5	22.6	26.1	26.0	26.4	27.4	28.5	29.4	32.6	35.8	37.0
≤ 30ys (n)	1196	938	729	826	980	1092	1068	1163	1013	964	927
NAFLD (n)	155	112	100	78	83	94	108	157	171	208	200
(%)	13.0	11.9	13.7	9.4	8.5	8.6	10.1	13.5	16.9	21.6	21.6
>30, ≤ 40ys (n)	2144	2292	2353	1504	1395	1248	1150	1096	1028	970	953
NAFLD (n)	419	502	590	412	386	375	358	367	372	357	344
(%)	19.5	21.9	25.1	27.4	27.7	30.0	31.1	33.5	36.2	36.8	36.1
>40, ≤ 50ys (n)	1480	1465	1450	1107	1183	1211	1279	1299	1339	1360	1347
NAFLD (n)	358	376	429	334	380	418	456	478	517	595	577
(%)	24.2	25.7	29.6	30.2	32.1	34.5	35.7	36.8	38.6	43.8	42.8
>50, ≤ 60ys (n)	1084	1180	1314	1033	1055	1068	1008	979	978	984	978
NAFLD (n)	267	341	415	352	364	394	361	350	385	416	456
(%)	24.6	28.9	31.6	34.1	34.5	36.9	35.8	35.8	39.4	42.3	46.6
>60, ≤ 70ys (n)	635	655	622	553	547	574	632	774	834	879	937
NAFLD (n)	137	147	165	151	159	172	202	243	292	331	382
(%)	21.6	22.4	26.5	27.3	29.1	30.0	32.0	31.4	35.0	37.7	40.8
>70ys	295	342	389	445	480	499	543	580	628	655	674
NAFLD (n)	63	77	91	92	116	109	131	139	159	174	193
(%)	21.4	22.5	23.4	20.7	24.2	21.8	24.1	24.0	25.3	26.6	28.6
Female (n)	2952	2980	2970	2558	2585	2617	2631	2661	2622	2651	2620
NAFLD (n)	288	320	371	378	383	381	417	435	460	529	582
(%)	9.8	10.7	12.5	14.8	14.8	14.6	15.8	16.3	17.5	20.0	22.2
≤ 30ys (n)	209	167	147	143	213	251	248	270	239	244	214
NAFLD (n)	5	4	4	1	0	1	2	8	9	14	13
(%)	2.4	2.4	2.7	0.7	0.0	0.4	0.8	3.0	3.8	5.7	6.1
>30, ≤ 40ys (n)	934	924	869	596	496	415	348	276	232	198	177
NAFLD (n)	33	29	36	28	28	24	23	17	19	20	20
(%)	3.5	3.1	4.1	4.7	5.6	5.8	6.6	6.2	8.2	10.1	11.3
>40, ≤ 50ys (n)	801	785	776	627	666	701	733	754	707	704	658
NAFLD (n)	54	60	63	62	69	76	85	88	84	98	99
(%)	6.7	7.6	8.1	9.9	10.4	10.8	11.6	11.7	11.9	13.9	15.0
>50, ≤ 60ys (n)	536	604	651	643	647	659	641	638	665	680	689
NAFLD (n)	78	100	133	131	132	128	132	132	131	157	171
(%)	14.6	16.6	20.4	20.4	20.4	19.4	20.6	20.7	19.7	23.1	24.8
>60, ≤ 70ys (n)	367	368	367	358	347	347	380	417	438	461	494
NAFLD (n)	94	96	92	99	95	92	106	117	123	140	159
(%)	25.6	26.1	25.1	27.7	27.4	26.5	27.9	28.1	28.1	30.4	32.2
>70ys	105	132	160	191	216	244	281	306	341	364	388
NAFLD (n)	24	31	43	57	59	60	69	73	94	100	120
(%)	22.9	23.5	26.9	29.8	27.3	24.6	24.6	23.9	27.6	27.5	30.9

Supplementary Table 3. Prevalence of NAFLD in obese persons (BMI > 25) (2006-2016)

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Total (n)	2445	2674	2749	2119	2314	2271	2218	2593	2159	2121	2106
NAFLD											
(n)	1104	1256	1397	1103	1231	1218	1239	1387	1309	1384	1414
(%)	45.2	47	50.8	52.1	53.2	53.6	55.9	53.5	60.6	65.3	67.1
Male (n)	1927	2163	2232	1663	1822	1813	1744	2086	1728	1716	1699
NAFLD											
(n)	912	1053	1180	885	1002	1001	1010	1153	1086	1147	1173
(%)	47.3	48.7	52.9	53.2	55	55.2	57.9	55.3	62.8	66.8	69
≤ 30ys (n)	213	202	174	156	190	227	220	293	224	212	200
NAFLD											
(n)	112	86	67	53	64	70	79	115	119	132	126
(%)	52.6	42.6	38.5	34	33.7	30.8	35.9	39.2	53.1	62.3	63
>30, ≤											
40ys (n)	549	640	689	427	428	375	346	407	310	302	291
NAFLD											
(n)	269	325	388	259	250	229	226	247	229	211	206
(%)	49	50.8	56.3	60.7	58.4	61.1	65.3	60.7	73.9	69.9	70.8
>40, ≤											
50ys (n)	433	456	468	353	414	423	431	517	443	437	425
NAFLD											
(n)	210	225	263	185	248	264	276	321	293	323	310
(%)	48.5	49.3	56.2	52.4	59.9	62.4	64	62.1	66.1	73.9	72.9
>50, ≤											
60ys (n)	351	447	491	364	390	392	353	381	312	309	303
NAFLD											
(n)	173	240	280	222	236	247	209	226	207	217	232
(%)	49.3	53.7	57	61	60.5	63	59.2	59.3	66.3	70.2	76.6
>60, ≤											
70ys (n)	265	274	249	213	213	209	222	291	263	262	281
NAFLD											
(n)	98	111	113	102	113	113	133	155	150	159	189
(%)	37	40.5	45.4	47.9	53.1	54.1	59.9	53.3	57	60.7	67.3
>70ys	116	144	161	150	187	187	172	197	176	194	199
NAFLD											
(n)	50	66	69	64	91	78	87	89	88	105	110
(%)	43.1	45.8	42.9	42.7	48.7	41.7	50.6	45.2	50	54.1	55.3
Female (n)	518	511	517	456	492	458	474	507	431	405	407
NAFLD											
(n)	192	203	217	218	229	217	229	234	223	237	241
(%)	37.1	39.7	42	47.8	46.5	47.4	48.3	46.2	51.7	58.5	59.2
≤ 30ys (n)	15	5	11	9	6	12	13	15	13	14	16
NAFLD											
(n)	3	1	2	0	0	1	2	5	5	5	7
(%)	20	20	18.2	0	0	8.3	15.4	33.3	38.5	35.7	43.8
>30, ≤											
40ys (n)	69	78	75	46	50	36	36	31	21	20	18
NAFLD											
(n)	20	18	22	18	22	14	17	9	10	11	11
(%)	29	23.1	29.3	39.1	44	38.9	47.2	29	47.6	55	61.1
>40, ≤											
50ys (n)	101	107	109	83	90	88	94	115	89	76	66
NAFLD											
(n)	37	41	41	34	36	39	44	48	45	46	35
(%)	36.6	38.3	37.6	41	40	44.3	46.8	41.7	50.6	60.5	53
>50, ≤											
60ys (n)	141	146	146	151	155	137	121	125	101	92	93
NAFLD											
(n)	49	56	63	72	72	66	56	61	53	62	61
(%)	34.8	38.4	43.2	47.7	46.5	48.2	46.3	48.8	52.5	67.4	65.6
>60, ≤											
70ys (n)	141	127	115	103	115	105	120	127	109	105	105
NAFLD											
(n)	64	66	58	59	65	59	68	67	57	63	66
(%)	45.4	52	50.4	57.3	56.5	56.2	56.7	52.8	52.3	60	62.9
>70ys (n)	51	48	61	64	76	80	90	94	98	98	109
NAFLD											
(n)	19	21	31	35	34	38	42	44	53	50	61
(%)	37.3	43.8	50.8	54.7	44.7	47.5	46.7	46.8	54.1	51	56

Supplementary Table 4. Logistic regression for risk factors of NAFLD

	Estimate	Std. Error	z value	OR	2.5%CI	97.5%CI	Pr(>|z|)
(Intercept)	-21.63	0.3151	-68.631	4.05E-10	2.18E-10	7.49E-10	< 0.001***
BMI	0.3791	0.005224	72.563	1.460966	1.446132	1.476054	< 0.001***
albumin	0.0994	0.005635	17.64	1.104513	1.092391	1.116791	< 0.001***
WGR	0.2666	0.06118	4.357	1.305464	1.157849	1.471676	< 0.001***
WBC	0.06149	0.008274	7.432	1.063423	1.046307	1.080801	< 0.001***
TG	0.2447	0.01213	20.17	1.277201	1.247335	1.308081	< 0.001***
HDL	-0.7862	0.04247	-18.511	0.45558	0.419127	0.495053	< 0.001***
GLT	0.002947	0.000358	8.242	1.002952	1.002255	1.003661	< 0.001***
ALT	0.02717	0.001056	25.738	1.027538	1.025425	1.029676	< 0.001***
AST	-0.0142	0.001834	-7.745	0.985896	0.982336	0.989418	< 0.001***
Cr	-0.02375	0.001025	-23.177	0.976528	0.974562	0.978485	< 0.001***
ALP	0.000909	0.000547	1.661	1.000909	0.999835	1.001982	0.09664
UA	0.00433	0.000176	24.542	1.004339	1.003992	1.004687	< 0.001***
Glu	0.2159	0.01197	18.044	1.241037	1.212327	1.270567	< 0.001***
SBp	0.006338	0.001014	6.252	1.006358	1.00436	1.008359	< 0.001***
DBp	0.003561	0.001546	2.303	1.003567	1.000532	1.006614	0.021265*
ESR	0.03401	0.001808	18.808	1.034593	1.030931	1.038265	< 0.001***
HGB	0.01966	0.00121	16.252	1.019853	1.017441	1.022277	< 0.001***
PLT	0.002588	0.000249	10.401	1.002592	1.002102	1.003081	< 0.001***
ApoB	0.4143	0.05413	7.654	1.513337	1.361002	1.6827	< 0.001***
TB	0.00778	0.002057	3.782	1.00781	1.003739	1.011865	< 0.001***

Supplementary Table 5. Cox regression for risk factors of NAFLD

	se(coef)	z value	coef	HR	lower .95	upper .95	Pr(>|z|)
BMI	0.005327	35.102	0.186996	1.205622	1.1930993	1.2182761	< 0.001***
Albumin	0.009655	4.656	0.04495	1.045976	1.0263691	1.0659575	< 0.001***
WGR	0.096408	1.326	0.12788	1.136417	0.9407519	1.3727778	0.184691
WBC	0.010893	4.087	0.044521	1.045527	1.0234412	1.0680884	< 0.001***
TG	0.011283	11.755	0.132633	1.141831	1.1168563	1.1673638	< 0.001***
HDL	0.071682	-13.214	-0.94719	0.387829	0.3369955	0.4463302	< 0.001***
GLT	0.000494	2.601	0.001285	1.001285	1.0003166	1.0022553	0.009302**
ALT	0.001478	9.309	0.013757	1.013853	1.0109201	1.0167934	< 0.001***
AST	0.003074	-4.965	-0.01526	0.984853	0.978937	0.990805	< 0.001***
Cr	0.001677	-5.609	-0.00941	0.990637	0.9873861	0.9938985	< 0.001***
ALP	0.000871	1.451	0.001263	1.001264	0.9995565	1.0029745	0.146888
UA	0.000277	11.14	0.003085	1.003089	1.0025451	1.0036338	< 0.001***
Glu	0.016397	3.771	0.061837	1.063788	1.0301446	1.0985309	0.000162***
SBp	0.001643	1.485	0.00244	1.002443	0.9992205	1.0056749	0.137485
DBp	0.002517	2.383	0.005996	1.006014	1.0010643	1.0109889	0.01719*
ESR	0.003054	3.429	0.010473	1.010528	1.0044965	1.0165956	0.000606***
HGB	0.002056	4.507	0.009269	1.009312	1.0052523	1.0133885	< 0.001***
PLT	0.000389	3.318	0.001291	1.001292	1.0005286	1.0020554	0.000905***
ApoB	0.092419	11.569	1.069192	2.913026	2.4303962	3.4914963	< 0.001***
TB	0.003316	-0.216	-0.00072	0.999283	0.9928097	1.0057994	0.828863

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	
Objectives	3	State specific objectives, including any prespecified hypotheses	
Methods			
Study design	4	Present key elements of study design early in the paper	
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	
Bias	9	Describe any efforts to address potential sources of bias	
Study size	10	Explain how the study size was arrived at	
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	
		(*b*) Describe any methods used to examine subgroups and interactions	
		(*c*) Explain how missing data were addressed	
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	
		(*e*) Describe any sensitivity analyses	

Continued on next page

Results			
Participants	13 *	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	
		(b) Give reasons for non-participation at each stage	
		(c) Consider use of a flow diagram	
Descriptive data	14 *	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	
		(b) Indicate number of participants with missing data for each variable of interest	
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	
Outcome data	15 *	*Cohort study*—Report numbers of outcome events or summary measures over time	
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	
		Cross-sectional study—Report numbers of outcome events or summary measures	
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	
		(*b*) Report category boundaries when continuous variables were categorized	
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	
Discussion			
Key results	18	Summarise key results with reference to study objectives	
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	
Generalisability	21	Discuss the generalisability (external validity) of the study results	
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Clinical outcome of nonalcoholic fatty liver disease: an eleven-year follow-up study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-054891.R2
Article Type:	Original research
Date Submitted by the Author:	07-May-2022
Complete List of Authors:	Tang, Xiaoping ; Ningbo Zhenhai District Lianhua Hospital , Department of Prevention and Health Care Shi, Yanyan; Peking University Third Hospital, Research Center of Clinical Epidemiology Du, Juan ; Ningbo Zhenhai District Lianhua Hospital , Department of Internal Medicine Hu, Keming; Ningbo Zhenhai District Lianhua Hospital, Department of Prevention and Health Care Zhou, Tingting; Ningbo Zhenhai District Lianhua Hospital , Department of Internal Medicine Chen, Lan; Ningbo Zhenhai District Lianhua Hospital , Department of Radiology Zhang, Yanming; Ningbo Zhenhai District Lianhua Hospital , Center Laboratory Li, Fujun; Ningbo Zhenhai District Lianhua Hospital , Department of Radiology Zhang, Huier ; Ningbo Zhenhai District Lianhua Hospital , Center Laboratory Liebe, Roman ; Heinrich Heine University Düsseldorf, Clinic of Gastroenterology, Hepatology and Infectious Diseases,; Saarland University, Department of Medicine II, Saarland University Medical Center Meyer, Christoph ; Heidelberg University, Department of Medicine II, Medical Faculty Mannheim Dooley, Steven ; Heidelberg University, Department of Medicine II, Medical Faculty Mannheim Zhu, Zhongwei; Ningbo Zhenhai District Lianhua Hospital , Department of Surgery Weng, Hong-Lei ; Heidelberg University, Department of Medicine II, Medical Faculty Mannheim JIA, Jinzhu; Peking University, Department of Biostatistics, School of Public Health; Peking University, Center for Statistical Science Huang, Tong; Ningbo City First Hospital; Ningbo Zhenhai District Lianhua Hospital , Department of Prevention and Health Care
Primary Subject Heading:	Gastroenterology and hepatology
Secondary Subject Heading:	Cardiovascular medicine, Diabetes and endocrinology
Keywords:	Epidemiology < INFECTIOUS DISEASES, Hypertension < CARDIOLOGY,

	DIABETES & ENDOCRINOLOGY, Health & safety < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Health informatics < BIOTECHNOLOGY & BIOINFORMATICS

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Clinical outcome of nonalcoholic fatty liver disease: an eleven-year follow-up study

Xiaoping Tang[1#], Yanyan Shi[2#], Juan Du[3#], Keming Hu[1], Tingting Zhou[3], Lan Chen[4], Yanming Zhang[5], Fujun Li[4], Huier Zhang[5], Roman Liebe[6,7], Christoph Meyer[8], Steven Dooley[8], Zhongwei Zhu[9], Hong-Lei Weng[8], Jinzhu Jia[10,11]*, Tong Huang[1,12]*

[1] Department of Prevention and Health Care, Ningbo Zhenhai Lianhua Hospital, Ningbo, China

[2] Research Center of Clinical Epidemiology, Peking University Third Hospital, Beijing, China

[3] Department of Internal Medicine, Ningbo Zhenhai Lianhua Hospital, Ningbo, China

[4] Department of Radiology, Ningbo Zhenhai Lianhua Hospital, Ningbo, China

[5] Center Laboratory, Ningbo Zhenhai Lianhua Hospital, Ningbo, China

[6] Clinic of Gastroenterology, Hepatology and Infectious Diseases, Heinrich Heine University, Düsseldorf, Germany;

[7] Department of Medicine II, Saarland University Medical Center, Saarland University, Homburg, Germany;

[8] Department of Medicine II, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany;

[9] Department of Surgery, Ningbo Zhenhai Lianhua Hospital, Ningbo, China

[10]Department of Biostatistics, School of Public Health, Peking University Health Science Center, Beijing, 100191, P.R. China

[11]Center for Statistical Science, Peking University, Beijing, 100871, P.R. China

[12]Center of health Management, Ningbo No.1 Hospital

# These authors contributed to the project equally.

* Co-senior authors

Corresponding authors:

Tong Huang, MD

Center of health Management, Ningbo No.1 Hospital

Address: No. 59, Liuting Street, Haishu District, Ningbo, Zhejiang Province, 315010, P.R. China

Email: nbbjzxht@163.com

Tel: +86-13586835095

Jinzhu Jia, PhD.
Department of Biostatistics, School of Public Health, Peking University Health Science Center
Address: No. 38, Xueyuan Road, Haidian District, Beijing, 100191, P.R. China
Email: jzjia@math.pku.edu.cn
Tel: +86-15801049187

Competing interests:

Authors do not have conflict of interest.

Funding statement:

This project is supported by the National Key R&D Program of China No. 2017YFC0908103 (X.T., and J.D.).

Data availability statement

Data are available upon reasonable request.

Acknowledgements:

We are grateful to Drs. Changxi Chen, Jingyi Yuan and Yongjun Chen for support and discussion.

Author contributions:

Conception and design: Xiaoping Tang, Yanyan Shi, Juan Du, Hong-Lei Weng, Jinzhu Jia and Tong Huang

Ultrasonography: Lan Chen, Fujun Li

Blood assays: Yanming Zhang and Huier Zhang

Other examinations and data collection: Xiaoping Tang, Keming Hu, Tingting Zhou, Juan Du, Zhongwei Zhu and Tong Huang

Statistical analyses: Yanyan Shi and Jinzhu Jia

Drafting the article: Hong-Lei Weng

Reviewing and editing the article critically: Xiaoping Tang, Yanyan Shi, Christoph Meyer, Roman Liebe, Steven Dooley, Hong-Lei Weng, Jinzhu Jia and Tong Huang

Abstract

Objectives To clarify NAFLD prevalence, risk factors, and clinical outcome in an exemplary Chinese population, a cohort of company employees was followed up for eleven years.

Design Retrospective cohort study

Setting Between 2006-2016 in China

Participants 13032 company employees

Results Over eleven years, the prevalence of NAFLD increased from 17.2% to 32.4% (males 20.5% to 37% *versus* females 9.8% to 22.2%). Male peak prevalence was between 40 and 60 years of age, whereas highest prevalence in women was at an age of 60 years and older. Logistic and Cox regression revealed 16 risk factors, including body mass index, albumin, white blood cell, triglycerides, high density lipoprotein, glutamyl transpeptidase, alanine transaminase, creatinine, urea acid, glucose, systolic blood pressure, diastolic blood pressure, Blood sedimentation, hemoglobin, platelet, and Apolipoprotein B2 ($P< 0.05$ for all factors). The Area Under the Curve of these variables for NAFLD is 0.88. However, cause-effect analyses showed that only body mass index, gender and triglycerides directly contributed to NAFLD development. Over an 11-year follow-up period, 12.6%, 37.7% and 14.2% of male NAFLD patients and 11.6%, 44.7% and 22.6% of female NAFLD patients developed diabetes, hypertension and hyperuricemia, respectively. Except one male patient who developed cirrhosis, no NAFLD patients progressed into severe liver disease.

Conclusion Diabetes, hypertension and hyperuricemia are the main clinical outcomes of NAFLD. Eleven years of NAFLD are not sufficient to cause severe liver disease. Age and obesity are direct risk factors for NAFLD. BMI, gender and triglycerides are three parameters directly reflecting the occurrence of NAFLD.

(Words: 250)

Strengths and limitations of this study

- This study dynamically follows up NAFLD prevalence in an eastern Chinese community for eleven years.
- The study adopted First order Markov models to evaluate the cause-effect link between NAFLD and risk factors.
- The relatively low sensitivity of ultrasound for the detection of liver fat might underestimate the incidence of NAFLD in this cohort.

- Given that the current study is a single-center observation, multiple-center studies are required to confirm the conclusions in the future.
- The study population is a highly select, relatively homogenous group of well-educated professionals in privileged social positions and permanent employment. Thus, the conclusions might not be transferable to the general Chinese population.

Introduction

Non-alcoholic fatty liver disease (NAFLD) is one of the most common causes of chronic liver disease globally [1]. The global prevalence of NAFLD is currently around 25% [2,3]. NAFLD is predicted to become the most frequent indication for liver transplantation by 2030 in Western countries [4]. An analysis based on 18 million patients in four European cohorts showed that NAFLD and non-alcoholic steatohepatitis (NASH) increase the risk of end-stage liver diseases, e.g., cirrhosis and hepatocellular carcinoma (HCC) [5]. Of note, NAFLD is not only a disease restricted to the liver, but also affects extra-hepatic organs. NAFLD is tightly associated with the occurrence of type 2 diabetes mellitus (T2DM), cardiovascular (CVD) and cardiac diseases, and chronic kidney disease (CKD) [4].

In China, the incidence of NAFLD has been increasing over the last two decades. A recent meta-analysis, based on 392 studies between 2008 and 2018, showed the national incidence of NAFLD in China to be at 29.2% [6]. In Shanghai, the adult incidence of NAFLD has increased from 14.04% in 1995 to 43.65% in 2015 [2]. Being a vast country, Chinese living in different areas vary widely in lifestyle and economic status. Thus, the epidemiology, natural history and clinical outcomes of NAFLD in different areas of the country are worth further investigation.

It is well accepted that viral hepatitis is a major reason for progressive chronic liver diseases, e.g., fibrosis, cirrhosis and ultimately, hepatocellular carcinoma (HCC). With regard to NAFLD, incidence and severity of associated chronic liver disease outcomes has not been monitored in large Chinese cohorts yet – especially over an extended time span. The current study therefore describes the prevalence of NAFLD in a large Eastern Chinese community over eleven years (2006 - 2016). We focused on three questions: (1) What is the annual incidence of NAFLD? (2) What are the risk factors for NAFLD? And (3) What are the most frequent extra- and intrahepatic clinical outcomes of NAFLD in this selected population?

Methods

Patient and public involvement

No patients were involved in this study

Design and participants

In this retrospective study we analyzed the "annual health examination database" of the Zhenhai Lianhua Hospital from 2006 to 2016. This hospital is affiliated to Sinopec Zhenhai Refining &

Chemical Company. Supported by the company, all employees were offered the opportunity to go to this hospital for an annual health examination. Over a period of 11 years, a total 13,032 employees received health examinations. From 2006 to 2016, 11689, 11706, 11584, 9521, 9592, 9725, 9710, 9869, 9718, 9702 and 9706 persons received health examinations (**Figure 1**). To describe the longitudinal NAFLD occurrence in this cohort, we excluded subjects with the following conditions: (1) viral hepatitis B and C infection, which were identified by blood virus measurements (HBV-DNA and HCV-RNA), and (2) alcoholic liver disease, which was defined as previously described [7,8]. NAFLD was defined as the presence of hepatic steatosis, determined by ultrasonography.

The study protocol was approved by the Ethics Committee of Zhenhai Lianhua Hospital ([2016]001). Informed consent was obtained from all subjects.

Measures

Supplementary Table 1 shows all parameters measured in the annual health examinations. Ultrasonography was performed by the same three experienced doctors (L.C., F. L., and J.Y.) with an Ultrasonograph B, GE, Voluson 730 pro. Blood biochemistry and HBV serum levels were measured by an Olympus AU640 autoanalyzer (Olympus, Kobe, Japan) and an ImmunoAssay Analyzer VitrosECI (JOHNSON-JOHNSON, USA), respectively. All methods were carried out in accordance with relevant guidelines and regulations.

Statistical analysis

For population characteristics, variables were described as means and standard deviation (SD) or proportions as appropriate. Student's t-test or nonparametric test was used to analyze differences between two groups as mentioned. Chi-square test was used to verify the differences of nominative variables between two groups. Multivariate analysis of risk factors for NAFLD was performed using logistic regression analysis. Combined receiver operating characteristic curve (ROC) and area under curve (AUC) analyses were used to evaluate the diagnostic performance of biomarkers based on the logistic regression model. Multivariate Cox regression model was performed to calculate hazard ratios of variables to identify independent prognostic variables. First order Markov models were used to analyze the cause-effect link between NAFLD and risk factors. L1 penalized logistic regression was applied to select predictive predictors. R package "glmnet" contains functions to select predictors using L1 penalized logistic regression. Statistical analyses were performed using SPSS 22.0 and R version 3.5.3.

P-values that were less than 0.05 were considered statistically significant. Figures were generated by R package such as 'forestplot', 'ROCR', 'bnlearn', or 'survival'.

Results

Prevalence of NAFLD from 2006 to 2016

We retrospectively analyzed 9786, 9852, 9827, 8026, 8225, 8309, 8311, 8552, 8442, 8463 and 8436 persons who received health examinations from 2006 to 2016. **Supplementary Table 2** shows the eleven-year annual NAFLD incidence in this population. In 2006, NAFLD was diagnosed in 17.2% of persons, and gradually increased over the examination period to 19% (2007), 22% (2008), 22.4% (2009), 22.7% (2010), 23.4% (2011), 24.5% (2012), 25.4% (2013), 27.9% (2014), 30.8% (2015), and 32.4% (2016), respectively (**Supplementary Table 2**). Both males and females demonstrated continuously increasing NAFLD prevalence (**Supplementary Table 2**). Compared to females, male Chinese demonstrated significantly higher NAFLD prevalence, e.g., in 2006, the prevalence of NAFLD in males and females was 20.5% and 9.8%, respectively. Eleven years later, the prevalence had increased to 37% in males and 22.2% in females (**Supplementary Table 2**). Noteworthy, the prevalence of NAFLD in male and female was correlating with age. The peak prevalence of NAFLD in men emerged in those aged between 40 and 60 years. In 2006, the prevalence of NAFLD in men aged between 40 - 50 and 50 - 60 years was 24.2% and 24.6%, respectively. In 2016, prevalence reached 42.8% and 46.6% (**Supplementary Table 2**) for men. Distinct from males, the peak NAFLD prevalence in females emerged at an age above 60 years. In 2006, the prevalence of NAFLD in women older than 60 and 70 years was 25.6% and 22.9%, respectively. Eleven years later, these values had increased to 53.4% and 30.9% (**Supplementary Table 2**).

Among the observed population, 5606 persons received annual health examinations for 11 years, and thus prevalence of NAFLD was analyzed in these individuals. As shown in **Table 1**, the prevalence of NAFLD increased from 17% in 2006 to 35.2% in 2016. The highest prevalence rates for NAFLD in 3795 men emerged in those older than 40 years. In 2006, the NAFLD prevalence in males aged between 40 - 50, 50 – 60, and 60 – 70 years was 24.2%, 25.4% and 19.8%, respectively. In 2016, these values reached 43.3%, 46.8% and 42.3% (**Table 1**). Different from males, the peak NAFLD prevalence in 1811 females emerged at an age of more than 60 years. In 2006, the prevalence of NAFLD in women older than 60 or 70 years

was 22.1% and 28%, respectively. Eleven years later, these values had increased to 31.8% and 28.6% (**Table 1**).

BMI and NAFLD incidence

Given the tight link between obesity and NAFLD, we paid special attention to the population with high BMI. We focused on the 5606 persons with complete follow-up and analyzed the prevalence of NAFLD in those with BMI >25. In total, out of the 5606 persons, 2445 presented with a BMI of >25. The prevalence of NAFLD in this overweight subpopulation was far higher than in the general population. In 2006, 45.2% of individuals (n=1104; male *vs.* female: 47.3% *vs.* 37.1%) with BMI >25 were suffering from NAFLD (**Supplementary Table 3**). In 2016, values reached 67.1% (n=1414; male *vs.* female: 69% *vs.* 59.2%, **Supplementary Table 3**). Impressively, the NAFLD prevalence in both genders was very high at any age, even in those below the age of 30 years. In 2006, among 213 overweight men, younger than 30 years, 52.6% were also diagnosed for NAFLD (**Supplementary Table 3**). This number increased to 63% in 2016 (**Supplementary Table 3**). In 2006, there were 15 overweight women aged less than 30 years. Among them, 3 presented as NAFLD (20%). In 2016, 7 out of 16 overweight women aged less than 30 years were identified. The NALFD prevalence had increased to 43.8% (**Supplementary Table 3**). In those older than 40 years, NAFLD prevalence increased from 36.6 – 45.4% in 2006 to 53 – 65.6% in 2016 (**Supplementary Table 3**).

Risk factors relevant to NAFLD occurrence

Next, we analyzed risk factors relevant to NAFLD occurrence. Logistic regression analysis was performed on 26 parameters, including gender, age, BMI, albumin, white globulin ratio (WGR), white blood cell count (WBC), low density lipoprotein (LDL), triglycerides (TG), high density lipoprotein (HDL), glutamyl transpeptidase (GLT), alanine transaminase (ALT), aspartate transaminase (AST), creatinine (Cr), alkaline phosphatase (ALP), blood urea nitrogen (Bun), uric acid (UA), blood glucose (Glu), systolic blood pressure (SBp), diastolic blood pressure (DBp), Blood sedimentation (ESR), hemoglobin (HGB), platelets (PLT), Apolipoprotein A1 (ApoA1), Apolipoprotein B2 (ApoB), total bilirubin (TB), total protein (TP). We performed variable selection by penalized Logistic regression using R package glmnet. Cross validation selected 16 variables as potential predictors. These were BMI, albumin, WBC, TG, HDL, GLT, ALT, Cr, UA, Glu, SBp, DBp, ESR, HGB, PLT and ApoB (**Supplementary Table 4**). The corresponding forest-plot is shown in **Figure 2A**. Among these variables, ApoB and BMI displayed the most robust positive correlation with NAFLD occurrence, while HDL had a

strong negative correlation with NAFLD incidence (**Supplementary Table 4**). The AUC of these variables for NAFLD is 0.88 (see ROC curve in **Figure 2B**). We further performed a time dependent Cox regression to calculate the hazard ratios of these parameters for NAFLD occurrence. Cox regression confirmed that the 16 parameters were significantly relevant to NAFLD incidence (**Supplementary Table 5** and **Figure 2C**). Furthermore, ApoB and HDL were the most robust positive and negative risk factors for NAFLD (**Figure 2C**).

Cause-effect link between risk factors and NAFLD occurrence

Although the aforementioned parameters were regarded as "risk factors" according to statistical models, it did not necessarily mean that all of them contributed to NAFLD occurrence. Based on 11 years of longitudinal data, it was possible to construct a dynamic Bayesian network to identify the risk factors most relevant to NAFLD occurrence. As shown in **Figure 3**, these parameters constituted a complicated, but clear intercross paradigm. Only three parameters, BMI, gender and TG directly pointed to NAFLD. In addition, ApoB impacted the incidence of NAFLD through contributing to TG. Furthermore, LDL can indirectly contribute to NAFLD through influencing ApoB. Very impressively, the dynamic Bayesian network pointed out that NAFLD directly leads to alterations of seven parameters: ALT, DBp, SBp, TP, albumin, HGB and UA. Intriguingly, age, GLT, AST, Cr, and BUN did not interact with any other parameter in our model, indicating that these factors correlate by incidence, but there is no causal interaction.

Outcome of NAFLD

Subsequently, we examined clinical outcomes of NAFLD over the eleven years. **Table 2** summarizes the incidence of intra- and extrahepatic diseases of 696 NAFLD and 222 NASH patients during the follow-up period. Among the total NAFLD and NASH population, only 1 male NAFLD patient developed liver cirrhosis within the eleven years. However, this time span witnessed significantly increased extrahepatic diseases, including diabetes, hypertension and hyperuricemia. In the NAFLD population, there were 64 (12.6%) men and 22 (11.6%) women, 191 (37.7%) men and 85 (44.7%) women, 72 (14.2%) men and 43 (22.6%) women who developed into type II diabetes, hypertension and hyperuricemia, respectively (**Table 2**).

Given that 2006 is the starting point of data collection, patients diagnosed for NAFLD in this year had by no means just manifested their disease, but rather patients had possibly developed NAFLD several years prior to inclusion. To clarify the exact clinical outcomes of NAFLD over one decade, we focused on the following two cohorts of individuals with annual health

examinations for 11 years: (1) patients who were diagnosed as non-NAFLD in 2006, but were NAFLD in 2007 (new NAFLD cohort); and (2) who were non-NAFLD in both 2006 and 2007 (non-NAFLD cohort). As shown in **Table 2**, 185 new NAFLD cases (138 men and 47 women) and 4547 non-NAFLD (2786 men and 1761 women) persons were found in 2007. Between 2007 and 2016, neither NAFLD nor non-NAFLD individuals developed liver cirrhosis or cancer. However, the one-decade follow-up reveals different prevalences of diabetes, hypertension and hyperuricemia: In NAFLD patients, there were 14 (10.1%) men and 5 (10.6%) women, 47 (34.1%) men and 21 (44.7%) women, 34 (24.6%) men and 8 (17%) women who developed type II diabetes, hypertension and hyperuricemia, respectively (**Table 2**). In non-NAFLD individuals, 157 (5.6%) men and 54 (3.4%) women, 259 (9.3%) men and 324 (18.4%), 284 (10.2%) men and 84 women (4.8%) developed type II diabetes, hypertension and hyperuricemia, respectively (**Table 2**). For all three diseases, statistically significant differences were determined between the two cohorts of population (all P <0.05, **Table 2**). These results suggest that diabetes, hypertension and hyperuricemia are the main clinical outcomes of NAFLD.

Discussion

This eleven-year follow-up retrospective study reports the following: (1) NAFLD prevalence has substantially increased in the examined Eastern Chinese population. (2) The prevalence of NAFLD differs by gender and age. Middle-aged men and elderly women are the two populations at highest risk for NAFLD. (3) Gender, BMI and triglycerides are the parameters directly associated with NAFLD occurrence. Regardless of gender and age, persons with high BMI (≥25) have a high risk for NAFLD development. (4) NAFLD directly leads to alterations of seven clinical parameters: ALT, DBp, SBp, TP, albumin, HGB and UA. (5) Within 11 years, a significant part of the NAFLD population develops three clinically relevant diseases: type 2 diabetes mellitus, hypertension and hyperuricemia. (6) Within 11 years, NAFLD does not cause severe liver disease, such as cirrhosis or HCC, in patients.

The most impressive observation of the current study is that among 918 diseased persons (696 NAFLD and 222 NASH), no patient progressed towards HCC and only 1 male NAFLD patient developed liver cirrhosis within the 11 years. Furthermore, among 185 new NAFLD cases diagnosed in 2007, none developed liver cirrhosis or liver cancer. Liver cirrhosis and HCC are commonly regarded as the most severe and costly clinical outcomes of NAFLD [9]. In USA and Europe, it is estimated that 10-15% of NAFLD patients develop advanced fibrosis [10]. In China,

a study with biopsy-proven NAFLD revealed 1.97%-2.97% cirrhosis prevalence [11]. In addition, NAFLD is regarded as the third-most common cause of cancer-related death worldwide [12]. In a study based on 4949 US patients with HCC, 701 patients had NAFLD [13]. It was estimated that the cumulative incidence of HCC among patients with NAFLD and cirrhosis ranges from 2.4% to 12.8% over a median follow-up period of 3.2 to 7.2 years [14] (Global Health Observatory) data. Mortality and global health estimates were obtained from: http://www.who.int/gho/mortality_burden_disease/en/. Last accessed on 1/7/2020.). Given that the above conclusions were based on cross-sectional investigations and statistical models, it has been unknown to date over which period a NAFLD patient develops liver cirrhosis or HCC (personal risk assessment). Our 11-year follow-up provides therefore a valuable and comprehensive dataset. In this study, most patients were diagnosed with NAFLD when they received a routine health examination. Before the examination, these people did not have any symptoms or signs of NAFLD. Therefore, they belong to NAFLD patients at a very early stage (although for 2006, the duration of pre-existing NAFLD cannot be determined). Except for a single person, no serious liver problems were observed within this time period. These data suggest that for the vast majority of patients with early stage NAFLD, 11 years are not sufficient to develop liver cirrhosis or cancer. Nasr et al followed up 129 NAFLD patients with varying fibrosis stages on two occasions (mean time 13.7 and 9.3 years). Liver biopsy analyses showed that 9.3% of patients developed end-stage liver disease and 34% advanced fibrosis [15]. The NAFLD patients observed by Nasr et al. actually belonged to the NASH category, because they suffered from fibrosis and elevated ALT and/or AST levels. As our study was based on examinations of healthy, liver biopsy is not justifiable. Very likely, the current cohort included a portion of NASH patients. They also did not show significant progression towards cirrhosis or HCC was monitored.

In contrast to hepatic complications, patients with NAFLD showed a significant risk for the development of extrahepatic diseases, including diabetes, hypertension and hyperuricemia. In 696 NAFLD patients, eleven years witnessed the development of 86 cases (12.4%) type II diabetes, 276 (40%) cases of hypertension and 115 patients with (16.5%) hyperuricemia, respectively. Interestingly, in 222 NASH patients, the prevalence of these three diseases was 12 (5.4%), 46 (20.7%), and 33 (14.9%) only. In general, men had a higher probability to develop these diseases than women. These results are consistent with previous reports from USA and Europe [16-18]. Whether NAFLD is associated with the risk of severe heart or brain diseases such as acute myocardial infarction (AMI) and stroke is worth further investigation. A recent

matched cohort study analyzed databases from four European countries, which included 17.7 million patients with NAFLD or NASH [19]. These patients had a mean follow-up of 2.1 to 5.5 years. The study showed that the diagnosis of NAFLD appears not to be associated with AMI or stroke risk after adjustment for established cardiovascular risk factors. Nevertheless, the authors mentioned that cardiovascular risk assessment in adults with a diagnosis of NAFLD is important [19]. Follow-up for 5 years might be not sufficient to reach a conclusion for this issue.

An important issue is the cause-and-effect relationship between NAFLD and its clinical outcomes such as diabetes, hypertension and hyperuricemia. A dynamic Bayesian network in the current study provides direct evidence on this issue: NAFLD directly results in alterations of several parameters, including DBp, SBp and UA, suggesting that NAFLD directly contributes to the occurrence of hypertension and hyperuricemia. The underlying mechanisms require further investigation.

The current dynamic Bayesian network analysis does not confirm a direct cause-and-effect relationship between NAFLD and type 2 diabetes mellitus. There are plenty of studies showing the close relationship between type 2 diabetes and NALFD [20]. Pathophysiologically, insulin resistance is a key event in both NAFLD and diabetes progression [20]. However, genome-wide association studies have not yet identified the exact impact of insulin resistance on the variants associated with NAFLD severity [20,21]. Clarification of the cause-and-effect relationship between NAFLD and diabetes requires further long-term follow-up studies.

To date, there are a large number of studies investigating risk factors for NAFLD [22]. These studies tried to identify single, or multiple combined biomarkers to predict NAFLD occurrence. Given that most studies were based on cross-sectional designs, or with only short follow-up periods, it is difficult to clarify the causality between the proposed predictors and NAFLD morbidity. Our eleven-year dataset provides a chance to shed led on this issue. Here, the dynamic causal relationships between variables, including risk parameters and clinical outcomes, were identified by a first order Markov model, which was displayed by a dynamic Bayes network. The dynamic Bayes model discriminates causal relationship through time sequence. When a variable change is closely related to a previous variance alteration, a causal relationship between the two variables is assumed. Based on Logistic and Cox regression and dynamic Bayesian network analyses, we confirmed three direct risk factors for NAFLD occurrence: gender, BMI and TG. These findings are supported by the following data: (1) Men have higher NAFLD prevalence than women in this population (37% *versus* 22.2% in 2016);

(2) In overweight people with a BMI >25, NAFLD prevalence reached 69% in males and 59.2% in females. Given that triglycerides are a major energy source, but are leading to obesity, it is not surprising that this parameter directly reflects the risk for NAFLD development. These findings provide robust evidence supporting the use of BMI to monitor or predict NAFLD.

Conclusion

This 11-year follow-up study documents the rapid increase in NAFLD prevalence in an Eastern Chinese population. In contrast to previous reports, our observation does not observe that one decade of NAFLD is sufficient to lead to severe hepatic clinical outcomes. It is worthy to note that our population were biased because they are on the well-off, well-educated side of the Chinese people, while previous studies were often based on hospital populations, who suffered from negative selection bias and thus came up with higher estimates. In addition, given there are differences in NAFLD profiles between Eastern and Western populations, it would be interesting to know the natural development of NAFLD in a Western population. A key point for clarifying the true history of NAFLD is to follow a population starting from the early phases of the disease. Consistent with previous studies, NAFLD is tightly associated with multiple extra-hepatic diseases relevant to the metabolic syndrome. In the future, follow-up of the current cohort for another one and two decades will provide further valuable data to clarify the extended natural history of NAFLD. Last but not least, a large portion of the men and women in this study were educated above the average and have a position in the company that gave them the availability of better food choices as well as regular sport. On the other hand, the relatively low sensitivity of ultrasound for the detection of liver fat might underestimate the incidence of NAFLD in this cohort.

References

[1] Younossi ZM, Tampi R, Priyadarshini M, et al. Burden of Illness and Economic Model for Patients With Nonalcoholic Steatohepatitis in the United States. Hepatology, 2019, 69: 564-572. http://doi.org/10.1002/hep.30254

[2] Younossi ZM, Loomba R, Anstee QM, et al. Diagnostic modalities for nonalcoholic fatty liver disease, nonalcoholic steatohepatitis, and associated fibrosis. Hepatology, 2018, 68: 349-360. http://doi.org/10.1002/hep.29721

[3] Fan JG, Kim SU, Wong VW. New trends on obesity and NAFLD in Asia. J Hepatol, 2017, 67: 862-873. http://doi.org/10.1016/j.jhep.2017.06.003

[4] Byrne CD, Targher G. NAFLD: a multisystem disease. J Hepatol, 2015, 62: S47-64. http://doi.org/10.1016/j.jhep.2014.12.012

[5] Alexander M, Loomis AK, van der Lei J, et al. Risks and clinical predictors of cirrhosis and hepatocellular carcinoma diagnoses in adults with diagnosed NAFLD: real-world study of 18 million patients in four European cohorts. BMC Med, 2019, 17: 95. http://doi.org/10.1186/s12916-019-1321-x

[6] Zhou F, Zhou J, Wang W, et al. Unexpected Rapid Increase in the Burden of NAFLD in China From 2008 to 2018: A Systematic Review and Meta-Analysis. Hepatology, 2019, 70: 1119-1133. http://doi.org/10.1002/hep.30702

[7] EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. J Hepatol, 2016, 64: 1388-1402. http://doi.org/10.1016/j.jhep.2015.11.004

[8] Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of nonalcoholic fatty liver disease: Practice guidance from the American Association for the Study of Liver Diseases. Hepatology, 2018, 67: 328-357. http://doi.org/10.1002/hep.29367

[9] Younossi Z, Tacke F, Arrese M, et al. Global Perspectives on Nonalcoholic Fatty Liver Disease and Nonalcoholic Steatohepatitis. Hepatology, 2019, 69: 2672-2682. http://doi.org/10.1002/hep.30251

[10] Younossi ZM, Koenig AB, Abdelatif D, et al. Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. Hepatology, 2016, 64: 73-84. http://doi.org/10.1002/hep.28431

[11] Shen F, Zheng RD, Shi JP, et al. Impact of skin capsular distance on the performance of controlled attenuation parameter in patients with chronic liver disease. Liver Int, 2015, 35: 2392-2400. http://doi.org/10.1111/liv.12809

[12] Ascha MS, Hanouneh IA, Lopez R, et al. The incidence and risk factors of hepatocellular carcinoma in patients with nonalcoholic steatohepatitis. Hepatology, 2010, 51: 1972-1978. http://doi.org/10.1002/hep.23527

[13] Younossi ZM, Otgonsuren M, Henry L, et al. Association of nonalcoholic fatty liver disease (NAFLD) with hepatocellular carcinoma (HCC) in the United States from 2004 to 2009. Hepatology, 2015, 62: 1723-1730. http://doi.org/10.1002/hep.28123

[14] White DL, Kanwal F, El-Serag HB. Association between nonalcoholic fatty liver disease and risk for hepatocellular cancer, based on systematic review. Clin Gastroenterol Hepatol, 2012, 10: 1342-1359.e1342. http://doi.org/10.1016/j.cgh.2012.10.001

[15] Nasr P, Ignatova S, Kechagias S, et al. Natural history of nonalcoholic fatty liver disease: A prospective follow-up study with serial biopsies. Hepatol Commun, 2018, 2: 199-210. http://doi.org/10.1002/hep4.1134

[16] Estes C, Razavi H, Loomba R, et al. Modeling the epidemic of nonalcoholic fatty liver disease demonstrates an exponential increase in burden of disease. Hepatology, 2018, 67: 123-133. http://doi.org/10.1002/hep.29466

[17] Williamson RM, Price JF, Glancy S, et al. Prevalence of and risk factors for hepatic steatosis and nonalcoholic Fatty liver disease in people with type 2 diabetes: the

Edinburgh Type 2 Diabetes Study. Diabetes Care, 2011, 34: 1139-1144. http://doi.org/10.2337/dc10-2229

[18] Targher G, Bertolini L, Padovani R, et al. Prevalence of nonalcoholic fatty liver disease and its association with cardiovascular disease among type 2 diabetic patients. Diabetes Care, 2007, 30: 1212-1218. http://doi.org/10.2337/dc06-2247

[19] Alexander M, Loomis AK, van der Lei J, et al. Non-alcoholic fatty liver disease and risk of incident acute myocardial infarction and stroke: findings from matched cohort study of 18 million European adults. Bmj, 2019, 367: l5367. http://doi.org/10.1136/bmj.l5367

[20] Tilg H, Moschen AR, Roden M. NAFLD and diabetes mellitus. Nat Rev Gastroenterol Hepatol, 2017, 14: 32-42. http://doi.org/10.1038/nrgastro.2016.147

[21] Zhou Y, Llauradó G, Orešič M, et al. Circulating triacylglycerol signatures and insulin sensitivity in NAFLD associated with the E167K variant in TM6SF2. J Hepatol, 2015, 62: 657-663. http://doi.org/10.1016/j.jhep.2014.10.010

[22] Miyake T, Kumagi T, Furukawa S, et al. Non-alcoholic fatty liver disease: factors associated with its presence and onset. J Gastroenterol Hepatol, 2013, 28 Suppl 4: 71-78. http://doi.org/10.1111/jgh.12251

Figure legends

Figure 1. Flow chart depicting the enrollment of a population with non-alcoholic fatty liver disease (NAFLD) for follow-up in Ningbo Zhenhai Lianhua Hospital, China.

Figure 2. Penalized logistic regression and Cox regression analysis were performed for risk factors and hazard ratios of non-alcoholic fatty liver disease (NAFLD). The following parameters were available from 5606 participants: Gender, age, BMI, albumin, white globulin ratio (WGR), white blood cell (WBC), low density lipoprotein (LDL), triglycerides (TG), high density lipoprotein (HDL), glutamyl transpeptidase (GLT), alanine transaminase (ALT), aspartate transaminase (AST), creatinine (Cr), alkaline phosphatase (ALP), blood urea nitrogen (Bun), Uric Acid (UA), blood glucose (Glu), systolic blood pressure (SBp), diastolic blood pressure (DBp), blood sedimentation (ESR), hemoglobin (HGB), platelets (PLT), Apolipoprotein A1 (ApoA1), Apolipoprotein B2 (ApoB), total bilirubin (TB), total protein (TP). Cross validation selected 16 variables to be potential predictors. The corresponding forest-plot is shown in (**A**). The AUC of these above 16 variables for NAFLD is 0.88 **(B)**. Cox regression confirmed that the 16 variables were relevant for NAFLD incidence, including BMI, albumin, WBC, TG, HDL, GLT, ALT, Cr, UA, Glu, SBp, DBp, ESR, HGB, PLT and ApoB. The corresponding forest-plot is shown (**C**).

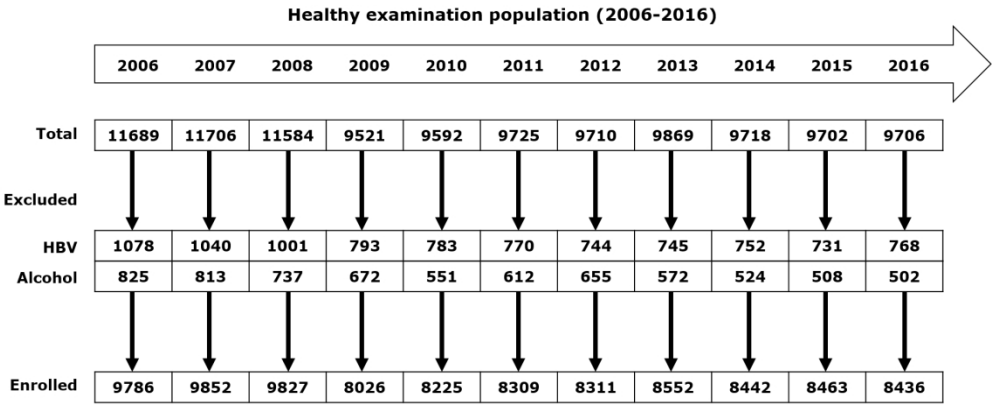
Figure 3. Dynamic Bayesian network analyses were performed to show the cause-effect link between non-alcoholic fatty liver disease (NAFLD) and its potential risk factors. Three variables, BMI, gender and trigylcerides directly pointed to NAFLD. ApoB impacted on the incidence of NAFLD through TG abundance. LDL indirectly contributed to NAFLD through ApoB. NAFLD directly led to alterations of seven clinical parameters: ALT, DBp, SBp, TP, albumin, HGB and UA.

Table 1. Prevalence of NAFLD in 5606 persons with 11-year follow-up (2006-2016)

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Total (n)	5606	5606	5606	5606	5606	5606	5606	5606	5606	5606	5606
NAFLD (n)	951	1068	1247	1322	1378	1466	1524	1574	1700	1883	1976
(%)	17	19.1	22.2	23.6	24.6	26.2	27.2	28.1	30.3	33.6	35.2
Male (n)	3795	3795	3795	3795	3795	3795	3795	3795	3795	3795	3795
NAFLD (n)	776	871	1014	1075	1117	1206	1244	1283	1387	1517	1569
(%)	20.4	23	26.7	28.3	29.4	31.8	32.8	33.8	36.5	40	41.3
≤ 30ys (n)	668	466	366	297	246	204	142	67	14	4	0
NAFLD (n)	97	71	57	46	35	35	30	17	3	1	0
(%)	14.5	15.2	15.6	15.5	14.2	17.2	21.1	25.4	21.4	25	0
>30, ≤ 40ys (n)	1212	1315	1342	1258	1180	1091	1034	981	896	778	668
NAFLD (n)	231	286	349	349	338	330	330	331	331	301	261
(%)	19.1	21.7	26	27.7	28.6	30.2	31.9	33.7	36.9	38.7	39.1
>40, ≤ 50ys (n)	873	863	842	928	988	1022	1097	1128	1170	1201	1212
NAFLD (n)	211	215	239	278	314	353	388	411	454	524	525
(%)	24.2	24.9	28.4	30	31.8	34.5	35.4	36.4	38.8	43.6	43.3
>50, ≤ 60ys (n)	562	623	696	741	776	826	796	798	820	851	873
NAFLD (n)	143	184	231	262	276	314	292	286	320	363	409
(%)	25.4	29.5	33.2	35.4	35.6	38	36.7	35.8	39	42.7	46.8
>60, ≤ 70ys (n)	368	382	373	348	346	356	398	465	496	521	562
NAFLD (n)	73	85	97	95	96	108	125	156	178	210	238
(%)	19.8	22.3	26	27.3	27.7	30.3	31.4	33.5	35.9	40.3	42.3
>70ys	112	146	176	223	259	296	328	356	399	440	480
NAFLD (n)	21	30	41	45	58	66	79	82	101	118	136
(%)	18.8	20.5	23.3	20.2	22.4	22.3	24.1	23	25.3	26.8	28.3
Female (n)	1811	1811	1811	1811	1811	1811	1811	1811	1811	1811	1811
NAFLD (n)	175	197	233	247	261	260	280	291	313	366	407
(%)	9,7	10,9	12,9	13,6	14,4	14,4	15,5	16,1	17,3	20,2	22,5
≤ 30ys (n)	85	44	22	13	9	9	5	2	0	0	0
NAFLD (n)	3	1	1	1	0	0	0	0	0	0	0
(%)	3.5	2.3	4.5	7.7	0	0	0	0	0	0	0
>30, ≤ 40ys (n)	582	578	541	480	403	334	277	213	173	124	85
NAFLD (n)	17	21	25	22	24	18	18	13	15	14	12
(%)	2.9	3.6	4.6	4.6	6	5.4	6.5	6.1	8.7	11.3	14.1
>40, ≤ 50ys (n)	485	469	482	496	541	580	612	638	617	608	582
NAFLD (n)	31	31	35	41	47	54	61	66	65	81	86
(%)	6.4	6.6	7.3	8.3	8.7	9.3	10	10.3	10.5	13.3	14.8
>50, ≤ 60ys (n)	365	400	422	450	461	469	456	452	467	476	485
NAFLD (n)	56	67	85	86	90	88	88	83	88	98	109
(%)	15.3	16.8	20.1	19.1	19.5	18.8	19.3	18.4	18.8	20.6	22.5
>60, ≤ 70ys (n)	244	260	262	267	266	266	280	301	314	337	365
NAFLD (n)	54	62	61	66	69	66	70	81	82	104	116
(%)	22.1	23.8	23.3	24.7	25.9	24.8	25	26.9	26.1	30.9	31.8
>70ys (n)	50	60	82	105	131	153	181	205	240	266	294
NAFLD (n)	14	15	26	31	31	34	43	48	63	69	84
(%)	28	25	31.7	29.5	23.7	22.2	23.8	23.4	26.3	25.9	28.6

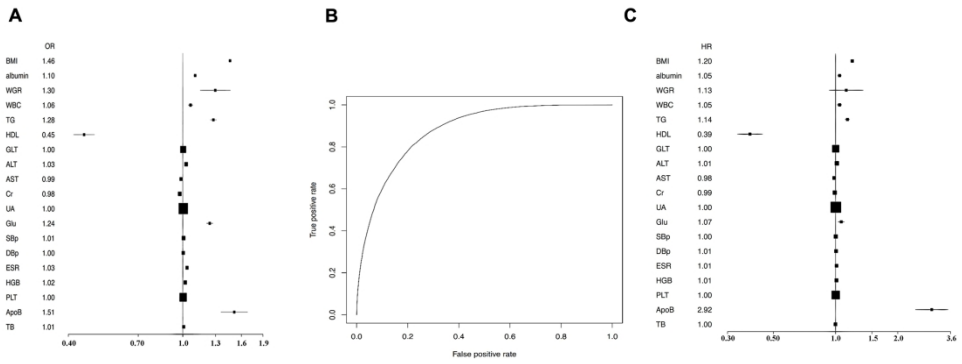
Table 2. Clinical outcome of NAFLD patients

2006 - 2016									
		Cirrhosis	HCC	Diabetes n (%)		Hypertension n (%)		Hyperuricemia n (%)	
Male (n=506)		0	0	64 (12.6)		191 (37.7)		72 (14.2)	
Female (n=190)		0	0	22 (11.6)		85 (44.7)		43 (22.6)	
2007 – 2016 (outcome of new NFLAD)									
		Cirrhosis	HCC	Diabetes		Hypertension		Hyperuricemia	
				n (%)	*P*-value	n (%)	*P*-value	n (%)	*P*-value
Male	NAFLD n=138	0	0	14 (10.1)	0.028	47(34.1)	< 0.001	34 (24.6)	< 0.001
	Non-NAFLD n=2786	0	0	157 (5.6)		259 (9.3)		284 (10.2)	
Female	NAFLD n=47	0	0	5 (10.6)	0.014	21 (44.7)	< 0.001	8 (17)	< 0.001
	Non-NAFLD n=1761	0	0	54 (3.1)		324 (18.4)		84 (4.8)	



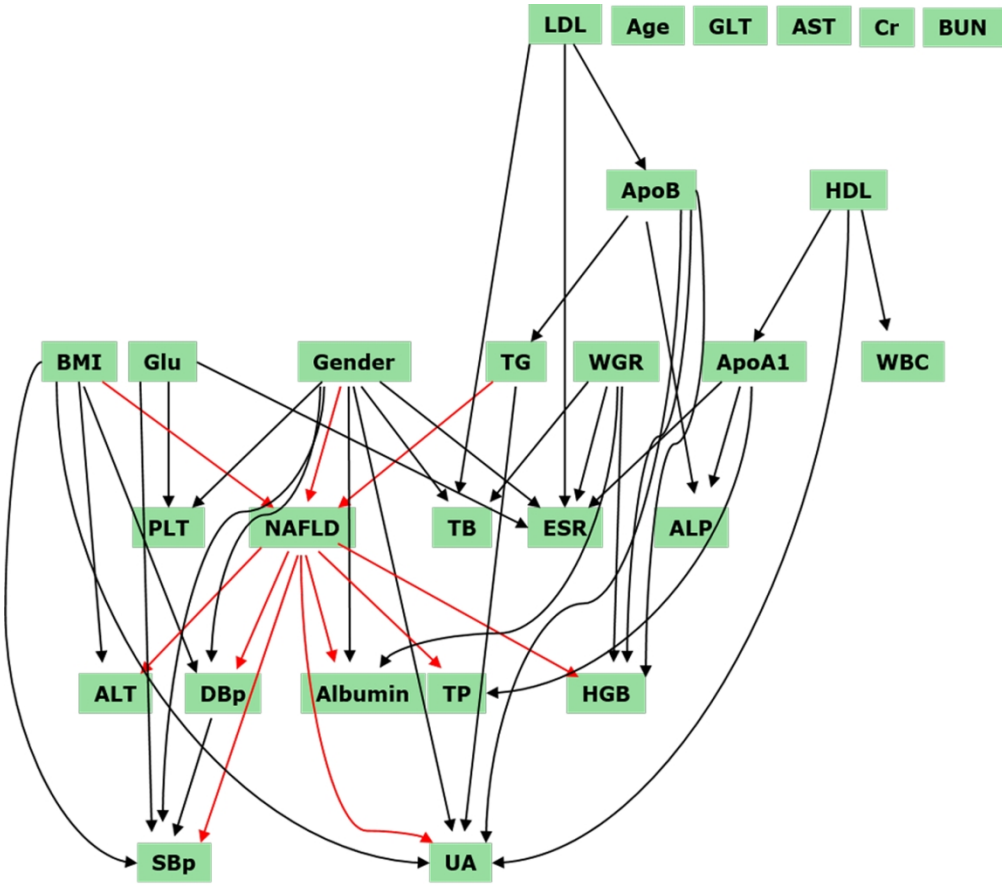
Flow chart depicting the enrollment of a population with non-alcoholic fatty liver disease (NAFLD) for follow-up in Ningbo Zhenhai Lianhua Hospital, China.

335x136mm (150 x 150 DPI)



Penalized logistic regression and Cox regression analysis were performed for risk factors and hazard ratios of non-alcoholic fatty liver disease (NAFLD). The following parameters were available from 5606 participants: Gender, age, BMI, albumin, white globulin ratio (WGR), white blood cell (WBC), low density lipoprotein (LDL), triglycerides (TG), high density lipoprotein (HDL), glutamyl transpeptidase (GLT), alanine transaminase (ALT), aspartate transaminase (AST), creatinine (Cr), alkaline phosphatase (ALP), blood urea nitrogen (Bun), Uric Acid (UA), blood glucose (Glu), systolic blood pressure (SBp), diastolic blood pressure (DBp), blood sedimentation (ESR), hemoglobin (HGB), platelets (PLT), Apolipoprotein A1 (ApoA1), Apolipoprotein B2 (ApoB), total bilirubin (TB), total protein (TP). Cross validation selected 16 variables to be potential predictors. The corresponding forest-plot is shown in (A). The AUC of these above 16 variables for NAFLD is 0.88 (B). Cox regression confirmed that the 16 variables were relevant for NAFLD incidence, including BMI, albumin, WBC, TG, HDL, GLT, ALT, Cr, UA, Glu, SBp, DBp, ESR, HGB, PLT and ApoB. The corresponding forest-plot is shown (C).

344x124mm (300 x 300 DPI)



Dynamic Bayesian network analyses were performed to show the cause-effect link between non-alcoholic fatty liver disease (NAFLD) and its potential risk factors. Three variables, BMI, gender and trigylcerides directly pointed to NAFLD. ApoB impacted on the incidence of NAFLD through TG abundance. LDL indirectly contributed to NAFLD through ApoB. NAFLD directly led to alterations of seven clinical parameters: ALT, DBp, SBp, TP, albumin, HGB and UA.

522x461mm (96 x 96 DPI)

Nonalcoholic fatty liver disease increased risk of metabolic diseases, but not severe liver disease: an eleven-year follow-up study

Xiaoping Tang, Yanyan Shi, Juan Du, Keming Hu, Tingting Zhou, Lan Chen, Yanming Zhang, Fujun Li, Huier Zhang, Roman Liebe, Christoph Meyer, Steven Dooley, Zhongwei Zhu, Hong-Lei Weng, Jinzhu Jia, Tong Huang

Supplementary Table 1. Parameters measured in the annual health examinations.

Age
Albumin
AKP (alkaline phosphatase)
ALT (alanine transaminase)
ApoA1 (Apolipoprotein A1)
ApoB (Apolipoprotein B2)
AST (aspartate transaminase)
BMI (Body mass index)
BLRV (whole blood low shear reduced viscosity)
BLRI (relative index of whole blood low shear)
BHRV (whole blood high shear reduced viscosity)
BHRI (relative index of whole blood high shear)
BVV200 (Whole blood viscosity value)
BUN (blood urea nitrogen)
BUS (ultrasound prompt)
CRP (high sensitive C-reactive protein)
Cr (creatinine)
CA (carotid atherosclerosis)
DBIL (Direct bilirubin)
DBp (diastolic blood pressure)
DM (type II diabetes)
ESR (Blood sedimentation)
ESRKV (Blood sedimentation equation K value)
Gender
GLT (glutamyl transpeptidase)
Glucose
HBP (Hypertension)
HBX (red blood cell deformation index TK)
HCT (Hematocrit)
HCY (Homocysteine)
HDL (high density lipoprotein C)
Height
HGB (hemoglobin)
LDL (low density lipoprotein C)
LVH (left ventricular hypertrophy)
MPV (mean platelet volume)
NAFLD (non-alcoholic fatty liver disease)
PhyExa (physical examination results)
PV (plasma viscosity)
PDW (Platelet distribution width)
PLT (platelet)
PCT (prothrombin consumption time)
RBC (red blood cell count)
SBp (systolic blood pressure)
TB (Total Bilirubin)
TC (Total cholesterol)
TG (Triglyceride)

TP (total protein)
UA (uric acid)
Waist
Weight
WGR (white globulin ratio)
WBC (white blood cell count)

Supplementary Table 2. Prevalence of NAFLD in an eastern Chinese population (2006-2016)

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Total (n)	9786	9852	9827	8026	8225	8309	8311	8552	8442	8463	8436
NAFLD (n)	1687	1875	2161	1797	1871	1943	2033	2169	2356	2610	2734
(%)	17.2	19.0	22.0	22.4	22.7	23.4	24.5	25.4	27.9	30.8	32.4
Male (n)	6834	6872	6857	5468	5640	5692	5680	5891	5820	5812	5816
NAFLD (n)	1399	1555	1790	1419	1488	1562	1616	1734	1896	2081	2152
(%)	20.5	22.6	26.1	26.0	26.4	27.4	28.5	29.4	32.6	35.8	37.0
≤ 30ys (n)	1196	938	729	826	980	1092	1068	1163	1013	964	927
NAFLD (n)	155	112	100	78	83	94	108	157	171	208	200
(%)	13.0	11.9	13.7	9.4	8.5	8.6	10.1	13.5	16.9	21.6	21.6
>30, ≤ 40ys (n)	2144	2292	2353	1504	1395	1248	1150	1096	1028	970	953
NAFLD (n)	419	502	590	412	386	375	358	367	372	357	344
(%)	19.5	21.9	25.1	27.4	27.7	30.0	31.1	33.5	36.2	36.8	36.1
>40, ≤ 50ys (n)	1480	1465	1450	1107	1183	1211	1279	1299	1339	1360	1347
NAFLD (n)	358	376	429	334	380	418	456	478	517	595	577
(%)	24.2	25.7	29.6	30.2	32.1	34.5	35.7	36.8	38.6	43.8	42.8
>50, ≤ 60ys (n)	1084	1180	1314	1033	1055	1068	1008	979	978	984	978
NAFLD (n)	267	341	415	352	364	394	361	350	385	416	456
(%)	24.6	28.9	31.6	34.1	34.5	36.9	35.8	35.8	39.4	42.3	46.6
>60, ≤ 70ys (n)	635	655	622	553	547	574	632	774	834	879	937
NAFLD (n)	137	147	165	151	159	172	202	243	292	331	382
(%)	21.6	22.4	26.5	27.3	29.1	30.0	32.0	31.4	35.0	37.7	40.8
>70ys	295	342	389	445	480	499	543	580	628	655	674
NAFLD (n)	63	77	91	92	116	109	131	139	159	174	193
(%)	21.4	22.5	23.4	20.7	24.2	21.8	24.1	24.0	25.3	26.6	28.6
Female (n)	2952	2980	2970	2558	2585	2617	2631	2661	2622	2651	2620
NAFLD (n)	288	320	371	378	383	381	417	435	460	529	582
(%)	9.8	10.7	12.5	14.8	14.8	14.6	15.8	16.3	17.5	20.0	22.2
≤ 30ys (n)	209	167	147	143	213	251	248	270	239	244	214
NAFLD (n)	5	4	4	1	0	1	2	8	9	14	13
(%)	2.4	2.4	2.7	0.7	0.0	0.4	0.8	3.0	3.8	5.7	6.1
>30, ≤ 40ys (n)	934	924	869	596	496	415	348	276	232	198	177
NAFLD (n)	33	29	36	28	28	24	23	17	19	20	20
(%)	3.5	3.1	4.1	4.7	5.6	5.8	6.6	6.2	8.2	10.1	11.3
>40, ≤ 50ys (n)	801	785	776	627	666	701	733	754	707	704	658
NAFLD (n)	54	60	63	62	69	76	85	88	84	98	99
(%)	6.7	7.6	8.1	9.9	10.4	10.8	11.6	11.7	11.9	13.9	15.0
>50, ≤ 60ys (n)	536	604	651	643	647	659	641	638	665	680	689
NAFLD (n)	78	100	133	131	132	128	132	132	131	157	171
(%)	14.6	16.6	20.4	20.4	20.4	19.4	20.6	20.7	19.7	23.1	24.8
>60, ≤ 70ys (n)	367	368	367	358	347	347	380	417	438	461	494
NAFLD (n)	94	96	92	99	95	92	106	117	123	140	159
(%)	25.6	26.1	25.1	27.7	27.4	26.5	27.9	28.1	28.1	30.4	32.2
>70ys	105	132	160	191	216	244	281	306	341	364	388
NAFLD (n)	24	31	43	57	59	60	69	73	94	100	120
(%)	22.9	23.5	26.9	29.8	27.3	24.6	24.6	23.9	27.6	27.5	30.9

Supplementary Table 3. Prevalence of NAFLD in obese persons (BMI > 25) (2006-2016)

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Total (n)	2445	2674	2749	2119	2314	2271	2218	2593	2159	2121	2106
NAFLD (n)	1104	1256	1397	1103	1231	1218	1239	1387	1309	1384	1414
(%)	45.2	47	50.8	52.1	53.2	53.6	55.9	53.5	60.6	65.3	67.1
Male (n)	1927	2163	2232	1663	1822	1813	1744	2086	1728	1716	1699
NAFLD (n)	912	1053	1180	885	1002	1001	1010	1153	1086	1147	1173
(%)	47.3	48.7	52.9	53.2	55	55.2	57.9	55.3	62.8	66.8	69
≤ 30ys (n)	213	202	174	156	190	227	220	293	224	212	200
NAFLD (n)	112	86	67	53	64	70	79	115	119	132	126
(%)	52.6	42.6	38.5	34	33.7	30.8	35.9	39.2	53.1	62.3	63
>30, ≤ 40ys (n)	549	640	689	427	428	375	346	407	310	302	291
NAFLD (n)	269	325	388	259	250	229	226	247	229	211	206
(%)	49	50.8	56.3	60.7	58.4	61.1	65.3	60.7	73.9	69.9	70.8
>40, ≤ 50ys (n)	433	456	468	353	414	423	431	517	443	437	425
NAFLD (n)	210	225	263	185	248	264	276	321	293	323	310
(%)	48.5	49.3	56.2	52.4	59.9	62.4	64	62.1	66.1	73.9	72.9
>50, ≤ 60ys (n)	351	447	491	364	390	392	353	381	312	309	303
NAFLD (n)	173	240	280	222	236	247	209	226	207	217	232
(%)	49.3	53.7	57	61	60.5	63	59.2	59.3	66.3	70.2	76.6
>60, ≤ 70ys (n)	265	274	249	213	213	209	222	291	263	262	281
NAFLD (n)	98	111	113	102	113	113	133	155	150	159	189
(%)	37	40.5	45.4	47.9	53.1	54.1	59.9	53.3	57	60.7	67.3
>70ys	116	144	161	150	187	187	172	197	176	194	199
NAFLD (n)	50	66	69	64	91	78	87	89	88	105	110
(%)	43.1	45.8	42.9	42.7	48.7	41.7	50.6	45.2	50	54.1	55.3
Female (n)	518	511	517	456	492	458	474	507	431	405	407
NAFLD (n)	192	203	217	218	229	217	229	234	223	237	241
(%)	37.1	39.7	42	47.8	46.5	47.4	48.3	46.2	51.7	58.5	59.2
≤ 30ys (n)	15	5	11	9	6	12	13	15	13	14	16
NAFLD (n)	3	1	2	0	0	1	2	5	5	5	7
(%)	20	20	18.2	0	0	8.3	15.4	33.3	38.5	35.7	43.8
>30, ≤ 40ys (n)	69	78	75	46	50	36	36	31	21	20	18
NAFLD (n)	20	18	22	18	22	14	17	9	10	11	11
(%)	29	23.1	29.3	39.1	44	38.9	47.2	29	47.6	55	61.1
>40, ≤ 50ys (n)	101	107	109	83	90	88	94	115	89	76	66
NAFLD (n)	37	41	41	34	36	39	44	48	45	46	35
(%)	36.6	38.3	37.6	41	40	44.3	46.8	41.7	50.6	60.5	53
>50, ≤ 60ys (n)	141	146	146	151	155	137	121	125	101	92	93
NAFLD (n)	49	56	63	72	72	66	56	61	53	62	61
(%)	34.8	38.4	43.2	47.7	46.5	48.2	46.3	48.8	52.5	67.4	65.6
>60, ≤ 70ys (n)	141	127	115	103	115	105	120	127	109	105	105
NAFLD (n)	64	66	58	59	65	59	68	67	57	63	66
(%)	45.4	52	50.4	57.3	56.5	56.2	56.7	52.8	52.3	60	62.9
>70ys (n)	51	48	61	64	76	80	90	94	98	98	109
NAFLD (n)	19	21	31	35	34	38	42	44	53	50	61
(%)	37.3	43.8	50.8	54.7	44.7	47.5	46.7	46.8	54.1	51	56

Supplementary Table 4. Logistic regression for risk factors of NAFLD

	Estimate	Std. Error	z value	OR	2.5%CI	97.5%CI	Pr(>|z|)
(Intercept)	-21.63	0.3151	-68.631	4.05E-10	2.18E-10	7.49E-10	< 0.001***
BMI	0.3791	0.005224	72.563	1.460966	1.446132	1.476054	< 0.001***
albumin	0.0994	0.005635	17.64	1.104513	1.092391	1.116791	< 0.001***
WGR	0.2666	0.06118	4.357	1.305464	1.157849	1.471676	< 0.001***
WBC	0.06149	0.008274	7.432	1.063423	1.046307	1.080801	< 0.001***
TG	0.2447	0.01213	20.17	1.277201	1.247335	1.308081	< 0.001***
HDL	-0.7862	0.04247	-18.511	0.45558	0.419127	0.495053	< 0.001***
GLT	0.002947	0.000358	8.242	1.002952	1.002255	1.003661	< 0.001***
ALT	0.02717	0.001056	25.738	1.027538	1.025425	1.029676	< 0.001***
AST	-0.0142	0.001834	-7.745	0.985896	0.982336	0.989418	< 0.001***
Cr	-0.02375	0.001025	-23.177	0.976528	0.974562	0.978485	< 0.001***
ALP	0.000909	0.000547	1.661	1.000909	0.999835	1.001982	0.09664
UA	0.00433	0.000176	24.542	1.004339	1.003992	1.004687	< 0.001***
Glu	0.2159	0.01197	18.044	1.241037	1.212327	1.270567	< 0.001***
SBp	0.006338	0.001014	6.252	1.006358	1.00436	1.008359	< 0.001***
DBp	0.003561	0.001546	2.303	1.003567	1.000532	1.006614	0.021265*
ESR	0.03401	0.001808	18.808	1.034593	1.030931	1.038265	< 0.001***
HGB	0.01966	0.00121	16.252	1.019853	1.017441	1.022277	< 0.001***
PLT	0.002588	0.000249	10.401	1.002592	1.002102	1.003081	< 0.001***
ApoB	0.4143	0.05413	7.654	1.513337	1.361002	1.6827	< 0.001***
TB	0.00778	0.002057	3.782	1.00781	1.003739	1.011865	< 0.001***

Supplementary Table 5. Cox regression for risk factors of NAFLD

	se(coef)	z value	coef	HR	lower .95	upper .95	Pr(>|z|)
BMI	0.005327	35.102	0.186996	1.205622	1.1930993	1.2182761	< 0.001***
Albumin	0.009655	4.656	0.04495	1.045976	1.0263691	1.0659575	< 0.001***
WGR	0.096408	1.326	0.12788	1.136417	0.9407519	1.3727778	0.184691
WBC	0.010893	4.087	0.044521	1.045527	1.0234412	1.0680884	< 0.001***
TG	0.011283	11.755	0.132633	1.141831	1.1168563	1.1673638	< 0.001***
HDL	0.071682	-13.214	-0.94719	0.387829	0.3369955	0.4463302	< 0.001***
GLT	0.000494	2.601	0.001285	1.001285	1.0003166	1.0022553	0.009302**
ALT	0.001478	9.309	0.013757	1.013853	1.0109201	1.0167934	< 0.001***
AST	0.003074	-4.965	-0.01526	0.984853	0.978937	0.990805	< 0.001***
Cr	0.001677	-5.609	-0.00941	0.990637	0.9873861	0.9938985	< 0.001***
ALP	0.000871	1.451	0.001263	1.001264	0.9995565	1.0029745	0.146888
UA	0.000277	11.14	0.003085	1.003089	1.0025451	1.0036338	< 0.001***
Glu	0.016397	3.771	0.061837	1.063788	1.0301446	1.0985309	0.000162***
SBp	0.001643	1.485	0.00244	1.002443	0.9992205	1.0056749	0.137485
DBp	0.002517	2.383	0.005996	1.006014	1.0010643	1.0109889	0.01719*
ESR	0.003054	3.429	0.010473	1.010528	1.0044965	1.0165956	0.000606***
HGB	0.002056	4.507	0.009269	1.009312	1.0052523	1.0133885	< 0.001***
PLT	0.000389	3.318	0.001291	1.001292	1.0005286	1.0020554	0.000905***
ApoB	0.092419	11.569	1.069192	2.913026	2.4303962	3.4914963	< 0.001***
TB	0.003316	-0.216	-0.00072	0.999283	0.9928097	1.0057994	0.828863

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	3
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	3-4
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	5
Objectives	3	State specific objectives, including any prespecified hypotheses	5
Methods			
Study design	4	Present key elements of study design early in the paper	5-6
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5-6
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	5-6
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	6
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	
Bias	9	Describe any efforts to address potential sources of bias	6
Study size	10	Explain how the study size was arrived at	6
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	6
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	6
		(*b*) Describe any methods used to examine subgroups and interactions	6
		(*c*) Explain how missing data were addressed	6
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	6
		(*e*) Describe any sensitivity analyses	6

Continued on next page

Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	
		(b) Give reasons for non-participation at each stage	
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	
		(b) Indicate number of participants with missing data for each variable of interest	
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	
		Cross-sectional study—Report numbers of outcome events or summary measures	
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	7-10
		(*b*) Report category boundaries when continuous variables were categorized	7-10
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	7-10
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	7-10
Discussion			
Key results	18	Summarise key results with reference to study objectives	10
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	11-13
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	11-13
Generalisability	21	Discuss the generalisability (external validity) of the study results	13
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	2

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.