

BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email info.bmjopen@bmj.com

BMJ Open

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

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

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



SCHOLARONE™
Manuscripts



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 Hu¹, Tingting Zhou³, Lan Chen⁴, Yanming Zhang⁵, Fujun Li⁴, Huier Zhang⁵, Roman Liebe^{6,7}, Christoph Meyer⁸, Steven Dooley⁸, Zhongwei Zhu⁹, Hong-Lei Weng⁸, 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

1
2
3
4
5 Jinzhu Jia, PhD.

6 Department of Biostatistics, School of Public Health, Peking University Health Science
7 Center
8

9 Address: No. 38, Xueyuan Road, Haidian District, Beijing, 100191, P.R. China

10 Email: jzjia@math.pku.edu.cn

11 Tel: +86-15801049187
12
13
14
15
16

17 **Acknowledgements:**

18
19 This project is supported by the National Key R&D Program of China No. 2017YFC0908103
20 (X.T., and J.D.). We are grateful to Drs. Changxi Chen, Jingyi Yuan and Yongjun Chen for
21 support and discussion.
22
23
24
25

26 **Author contributions:**

27
28 Conception and design: Xiaoping Tang, Yanyan Shi, Juan Du, Hong-Lei Weng, Jinzhu Jia
29 and Tong Huang
30
31

32 Ultrasonography: Lan Chen, Fujun Li

33 Blood assays: Yanming Zhang and Huier Zhang

34
35 Other examinations and data collection: Xiaoping Tang, Keming Hu, Tingting Zhou, Juan Du,
36
37 Zhongwei Zhu and Tong Huang
38
39

40 Statistical analyses: Yanyan Shi and Jinzhu Jia

41
42 Drafting the article: Hong-Lei Weng
43
44

45
46 Reviewing and editing the article critically: Xiaoping Tang, Yanyan Shi, Christoph Meyer,
47
48 Roman Liebe, Steven Dooley, Hong-Lei Weng, Jinzhu Jia and Tong Huang
49
50
51
52
53
54
55
56
57
58
59
60

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 is 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 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

1
2
3 NAFLD and risk factors. L1 penalized logistic regression was applied to select predictive
4 predictors. R package “glmnet” contains functions to select predictors using L1 penalized
5 logistic regression. Statistical analyses were performed using SPSS 22.0 and R version 3.5.3.
6
7
8 *P*-values that were less than 0.05 were considered statistically significant. Figures were
9 generated by R package such as ‘forestplot’, ‘ROCR’, ‘bnlearn’, or ‘survival’.

15 Results

17 *Prevalence of NAFLD from 2006 to 2016*

19 We retrospectively analyzed 9786, 9852, 9827, 8026, 8225, 8309, 8311, 8552, 8442, 8463 and
20 8436 persons who received health examinations from 2006 to 2016. **Supplementary Table 2**
21 shows the eleven-year annual NAFLD prevalence in this population. In 2006, NAFLD was
22 diagnosed in 17.2% of persons, and gradually increased over the examination period to 19%
23 (2007), 22% (2008), 22.4% (2009), 22.7% (2010), 23.4% (2011), 24.5% (2012), 25.4% (2013),
24 27.9% (2014), 30.8% (2015), and 32.4% (2016), respectively (**Supplementary Table 2**). Both
25 males and females demonstrated continuously increasing NAFLD prevalence (**Supplementary**
26 **Table 2**). Compared to females, male Chinese demonstrated significantly higher NAFLD
27 prevalence, e.g., in 2006, the prevalence of NAFLD in males and females was 20.5% and 9.8%,
28 respectively. Eleven years later, the prevalence had increased to 37% in males and 22.2% in
29 females (**Supplementary Table 2**). Noteworthy, the prevalence of NAFLD in male and female
30 was correlating with age. The peak prevalence of NAFLD in men emerged in those aged
31 between 40 and 60 years. In 2006, the prevalence of NAFLD in men aged between 40 - 50 and
32 50 - 60 years was 24.2% and 24.6%, respectively. In 2016, prevalence reached 42.8% and
33 46.6% (**Supplementary Table 2**) for men. Distinct from males, the peak NAFLD prevalence
34 in females emerged at an age above 60 years. In 2006, the prevalence of NAFLD in women
35 older than 60 and 70 years was 25.6% and 22.9%, respectively. Eleven years later, these values
36 had increased to 53.4% and 30.9% (**Supplementary Table 2**).

51
52 Among the observed population, 5606 persons received annual health examinations for 11
53 years, and thus prevalence of NAFLD was analyzed in these individuals. As shown in **Table 1**,
54 the prevalence of NAFLD increased from 17% in 2006 to 35.2% in 2016. The highest
55 prevalence rates for NAFLD in 3795 men emerged in those older than 40 years. In 2006, the
56 NAFLD prevalence in males aged between 40 - 50, 50 - 60, and 60 - 70 years was 24.2%,
57 25.4% and 19.8%, respectively. In 2016, these values reached 43.3%, 46.8% and 42.3% (**Table**
58
59
60

1
2
3 1). Different from males, the peak NAFLD prevalence in 1811 females emerged at an age of
4 more than 60 years. In 2006, the prevalence of NAFLD in women older than 60 or 70 years
5 was 22.1% and 28%, respectively. Eleven years later, these values had increased to 31.8% and
6 28.6% (**Table 1**).
7
8
9

10 *BMI and NAFLD incidence*

11
12 Given the tight link between obesity and NAFLD, we paid special attention to the population
13 with high BMI. We focused on the 5606 persons with complete follow-up and analyzed the
14 prevalence of NAFLD in those with BMI >25. In total, out of the 5606 persons, 2445 presented
15 with a BMI of >25. The prevalence of NAFLD in this overweighted population was far higher
16 than in the general population. In 2006, 45.2% of individuals (n=1104; male vs. female: 47.3%
17 vs. 37.1%) with BMI >25 were suffering from NAFLD (**Supplementary Table 3**). In 2016,
18 values reached 67.1% (n=1414; male vs. female: 69% vs. 59.2%, **Supplementary Table 3**).
19 Impressively, the NAFLD prevalence in both genders was very high at any age, even in those
20 below the age of 30 years. In 2006, among 213 overweighted men, younger than 30 years,
21 52.6% were also diagnosed for NAFLD (**Supplementary Table 3**). This number increased to
22 63% in 2016 (**Supplementary Table 3**). In 15 (in 2006) and 16 (in 2016) overweighted women
23 aged less than 30 years, the NAFLD prevalence was 20% and 43.8%, respectively
24 (**Supplementary Table 3**). In those older than 40 years, NAFLD prevalence increased from
25 36.6 – 45.4% in 2006 to 53 – 65.6% (**Supplementary Table 3**).
26
27
28
29
30
31
32
33
34
35
36
37
38

39 *Risk factors relevant to NAFLD occurrence*

40
41 Next, we analyzed risk factors relevant to NAFLD occurrence. Logistic regression analysis was
42 performed on 26 parameters, including gender, age, BMI, albumin, white globulin ratio (WGR),
43 white blood cell count (WBC), low density lipoprotein (LDL), triglycerides (TG), high density
44 lipoprotein (HDL), glutamyl transpeptidase (GLT), alanine transaminase (ALT), aspartate
45 transaminase (AST), creatinine (Cr), alkaline phosphatase (ALP), blood urea nitrogen (Bun),
46 uric acid (UA), blood glucose (Glu), systolic blood pressure (SBp), diastolic blood pressure
47 (DBp), Blood sedimentation (ESR), hemoglobin (HGB), platelets (PLT), Apolipoprotein A1
48 (ApoA1), Apolipoprotein B2 (ApoB), total bilirubin (TB), total protein (TP). We performed
49 variable selection by penalized Logistic regression using R package glmnet. Cross validation
50 selected 16 variables as potential predictors. These were BMI, albumin, WBC, TG, HDL, GLT,
51 ALT, Cr, UA, Glu, SBp, DBp, ESR, HGB, PLT and ApoB (**Supplementary Table 4**). The
52 corresponding forest-plot is shown in **Figure 2A**. Among these variables, ApoB and BMI
53
54
55
56
57
58
59
60

1
2
3 displayed the most robust positive correlation with NAFLD occurrence, while HDL had a
4 strong negative correlation with NAFLD incidence (**Supplementary Table 4**). The AUC of
5 these variables for NAFLD is 0.88 (see ROC curve in **Figure 2B**). We further performed a time
6 dependent Cox regression to calculate the hazard ratios of these parameters for NAFLD
7 occurrence. Cox regression confirmed that the 16 parameters were significantly relevant to
8 NAFLD incidence (**Supplementary Table 5** and **Figure 2C**). Furthermore, ApoB and HDL
9 were the most robust positive and negative risk factors for NAFLD (**Figure 2C**).

16 17 *Cause-effect link between risk factors and NAFLD occurrence*

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

39 40 41 *Outcome of NAFLD*

42 Subsequently, we examined clinical outcomes of NAFLD over the eleven years. **Table 2**
43 summarizes the incidence of intra- and extrahepatic diseases of 696 NAFLD and 222 NASH
44 patients during the follow-up period. Among the total NAFLD and NASH population, only 1
45 male NAFLD patient developed liver cirrhosis within the eleven years. However, this time span
46 witnessed significantly increased extrahepatic diseases, including diabetes, hypertension and
47 hyperuricemia. In the NAFLD population, there were 64 (12.6%) men and 22 (11.6%) women,
48 191 (37.7%) men and 85 (44.7%) women, 72 (14.2%) men and 43 (22.6%) women who
49 developed into type II diabetes, hypertension and hyperuricemia, respectively (**Table 2**).

50 Given that 2006 is the starting point of data collection, patients diagnosed for NAFLD in this
51 year did not mean that their disease had just manifested, but rather patients could have suffered
52 for more years. To clarify the exact clinical outcomes of NAFLD over one decade, we focused
53
54
55
56
57
58
59
60

1
2
3 on the following two cohorts of individuals with annual health examinations for 11 years: (1)
4 patients who were diagnosed as non-NAFLD in 2006, but were NAFLD in 2007 (new NAFLD
5 cohort); and (2) who were non-NAFLD in both 2006 and 2007 (non-NAFLD cohort). As shown
6 in **Table 2**, 185 new NAFLD cases (138 men and 47 women) and 4547 non-NAFLD (2786
7 men and 1761 women) persons were found in 2007. Between 2007 and 2016, neither NAFLD
8 nor non-NAFLD individuals developed liver cirrhosis or cancer. However, the one-decade
9 follow-up shows different prevalence of diabetes, hypertension and hyperuricemia. In NAFLD
10 patients, there were 14 (10.1%) men and 5 (10.6%) women, 47 (34.1%) men and 21 (44.7%)
11 women, 34 (24.6%) men and 8 (17%) women who developed type II diabetes, hypertension
12 and hyperuricemia, respectively (**Table 2**). In non-NAFLD individuals, 157 (5.6%) men and
13 54 (3.4%) women, 259 (9.3%) men and 324 (18.4%), 284 (10.2%) men and 84 women (4.8%)
14 developed into type II diabetes, hypertension and hyperuricemia, respectively (**Table 2**). For
15 all three diseases, statistically significant differences were determined between the two cohorts
16 of population (all $P < 0.05$, **Table 2**). These results suggest that diabetes, hypertension and
17 hyperuricemia are the main clinical outcomes of NAFLD.
18
19
20
21
22
23
24
25
26
27
28
29
30
31

32 Discussion

33 This eleven-year follow-up retrospective study reports the following: (1) NAFLD prevalence
34 has substantially increased in the examined Eastern Chinese population. (2) The prevalence of
35 NAFLD differs by gender and age. Middle-aged men and elderly women are the two
36 populations at highest risk for NAFLD. (3) Gender, BMI and triglycerides are the parameters
37 directly associated with NAFLD occurrence. Regardless of gender and age, persons with high
38 BMI (≥ 25) have a high risk for NAFLD development. (4) NAFLD directly leads to alterations
39 of seven clinical parameters: ALT, DBp, SBp, TP, albumin, HGB and UA. (5) Within 11 years,
40 a significant part of the NAFLD population develops three clinically relevant diseases: type 2
41 diabetes mellitus, hypertension and hyperuricemia. (6) Within 11 years, NAFLD does not cause
42 severe liver disease, such as cirrhosis or HCC, in patients.
43
44
45
46
47
48
49
50
51

52 The most impressive observation of the current study is that among 918 diseased persons (696
53 NAFLD and 222 NASH), no patient progressed towards HCC and only 1 male NAFLD patient
54 developed liver cirrhosis within the 11 years. Furthermore, among 185 new NAFLD cases
55 diagnosed in 2007, none developed liver cirrhosis or liver cancer. Liver cirrhosis and HCC are
56 commonly regarded as the most severe and costly clinical outcomes of NAFLD [9]. In USA and
57 Europe, it is estimated that 10-15% of NAFLD patients develop advanced fibrosis [10]. In China,
58
59
60

1
2
3 a study with biopsy-proven NAFLD revealed 1.97%-2.97% cirrhosis prevalence [11]. In
4 addition, NAFLD is regarded as the third-most common cause of cancer-related death
5 worldwide [12]. In a study based on 4949 US patients with HCC, 701 patients had NAFLD [13].
6
7 It was estimated that the cumulative incidence of HCC among patients with NAFLD and
8 cirrhosis ranges from 2.4% to 12.8% over a median follow-up period of 3.2 to 7.2 years [14]
9
10 (Global Health Observatory) data. Mortality and global health estimates were obtained from:
11 http://www.who.int/gho/mortality_burden_disease/en/. Last accessed on 1/7/2020.). Given that
12 the above conclusions were based on cross-sectional investigations and statistical models, it has
13 been unknown to date over which period a NAFLD patient develops liver cirrhosis or HCC
14 (personal risk assessment). Our 11-year follow-up provides therefore a valuable and
15 comprehensive dataset. In this study, most patients were diagnosed with NAFLD when they
16 received a routine health examination. Before the examination, these people did not have any
17 symptoms or signs of NAFLD. Therefore, they belong to NAFLD patients at a very early stage
18 (although for 2006, the duration of pre-existing NAFLD cannot be determined). Except for a
19 single person, no serious liver problems were observed within this time period. These data
20 suggest that for the vast majority of patients with early stage NAFLD, 11 years are not sufficient
21 to develop liver cirrhosis or cancer. Nasr et al followed up 129 NAFLD patients with varying
22 fibrosis stages on two occasions (mean time 13.7 and 9.3 years). Liver biopsy analyses showed
23 that 9.3% of patients developed end-stage liver disease and 34% advanced fibrosis [15]. The
24 NAFLD patients observed by Nasr et al. actually belonged to the NASH category, because they
25 suffered from fibrosis and elevated ALT and/or AST levels. As a study based on healthy
26 examination, liver biopsy is impossible for such a study. Very likely, the current cohort included
27 a portion of NASH patients. They also did not show significant progression towards cirrhosis
28 or HCC was monitored.

29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46 In contrast to hepatic complications, patients with NAFLD showed a significant risk for the
47 development of extrahepatic diseases, including diabetes, hypertension and hyperuricemia. In
48 696 NAFLD patients, eleven years witnessed the development of 86 cases (12.4%) type II
49 diabetes, 276 (40%) cases of hypertension and 115 patients with (16.5%) hyperuricemia,
50 respectively. Interestingly, in 222 NASH patients, the prevalence of these three diseases was
51 12 (5.4%), 46 (20.7%), and 33 (14.9%) only. In general, men had a higher probability to develop
52 these diseases than women. These results are consistent with previous reports from USA and
53 Europe [16-18].
54
55
56
57
58
59
60

1
2
3 To date, there are a large number of studies investigating risk factors for NAFLD [19]. These
4 studies tried to identify single, or multiple combined biomarkers to predict NAFLD occurrence.
5 Given that most studies were based on cross-sectional designs, or with only short follow-up
6 periods, it is difficult to clarify the causality between the proposed predictors and NAFLD
7 morbidity. Our eleven-year dataset provides a chance to shed led on this issue. Here, the
8 dynamic causal relationships between variables, including risk parameters and clinical
9 outcomes, were identified by a first order Markov model, which was displayed by a dynamic
10 Bayes network. The dynamic Bayes model discriminates causal relationship through time
11 sequence. When a variable change is closely related to a previous variance alteration, a causal
12 relationship between the two variables is assumed. Based on Logistic and Cox regression and
13 dynamic Bayesian network analyses, we confirmed three direct risk factors for NAFLD
14 occurrence: gender, BMI and TG. These findings are supported by the following data: (1) Men
15 have higher NAFLD prevalence than women in this population (37% *versus* 22.2% in 2016);
16 (2) In overweighted people with a BMI >25, NAFLD prevalence reached 69% in males and
17 59.2% in females. Given that triglycerides are a major energy source, but are leading to obesity,
18 it is not surprising that this parameter directly reflects the risk for NAFLD development. These
19 findings provide robust evidence supporting the use of BMI to monitor or predict NAFLD.
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35

36 **Conclusion**

37 This 11-year follow-up study documents the rapid increase in NAFLD prevalence in an Eastern
38 Chinese population. In contrast to previous reports, our observation does not observe that one
39 decade of NAFLD is sufficient to lead to severe hepatic clinical outcomes. It is worthy to note
40 that our population was biased towards physically fit and active people in full employment,
41 while previous studies were often based on hospital populations, who suffered from negative
42 selection bias and thus came up with higher estimates. In addition, given there are differences
43 in NAFLD profiles between Eastern and Western populations, it would be interesting to know
44 the natural development of NAFLD in a Western population. A key point for clarifying the true
45 history of NAFLD is to follow a population starting from the early phases of the disease.
46 Consistent with previous studies, NAFLD does indeed induce multiple extra-hepatic diseases
47 relevant to the metabolic syndrome. In the future, follow-up of the current cohort for another
48 one and two decades will provide further valuable data to clarify the extended natural history
49 of NAFLD. Last but not least, a large portion of the men and women in this study were educated
50 above the average and have a position in the company that gave them the availability of better
51
52
53
54
55
56
57
58
59
60

1
2
3 food choices as well as regular sport. On the other hand, the relatively low sensitivity of
4 ultrasound for the detection of liver fat might underestimate the incidence of NAFLD in this
5 cohort.
6
7
8
9
10

11 **References**

- 12
13
14 [1] Younossi ZM, Tampi R, Priyadarshini M, et al. Burden of Illness and Economic Model
15 for Patients With Nonalcoholic Steatohepatitis in the United States. *Hepatology*, 2019,
16 69: 564-572.
17
18 [2] Younossi ZM, Loomba R, Anstee QM, et al. Diagnostic modalities for nonalcoholic
19 fatty liver disease, nonalcoholic steatohepatitis, and associated fibrosis. *Hepatology*,
20 2018, 68: 349-360.
21
22 [3] Fan JG, Kim SU, Wong VW. New trends on obesity and NAFLD in Asia. *J Hepatol*,
23 2017, 67: 862-873.
24
25 [4] Byrne CD, Targher G. NAFLD: a multisystem disease. *J Hepatol*, 2015, 62: S47-64.
26
27 [5] Alexander M, Loomis AK, van der Lei J, et al. Risks and clinical predictors of cirrhosis
28 and hepatocellular carcinoma diagnoses in adults with diagnosed NAFLD: real-world
29 study of 18 million patients in four European cohorts. *BMC Med*, 2019, 17: 95.
30
31 [6] Zhou F, Zhou J, Wang W, et al. Unexpected Rapid Increase in the Burden of NAFLD
32 in China From 2008 to 2018: A Systematic Review and Meta-Analysis. *Hepatology*,
33 2019, 70: 1119-1133.
34
35 [7] EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic
36 fatty liver disease. *J Hepatol*, 2016, 64: 1388-1402.
37
38 [8] Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of
39 nonalcoholic fatty liver disease: Practice guidance from the American Association for
40 the Study of Liver Diseases. *Hepatology*, 2018, 67: 328-357.
41
42 [9] Younossi Z, Tacke F, Arrese M, et al. Global Perspectives on Nonalcoholic Fatty Liver
43 Disease and Nonalcoholic Steatohepatitis. *Hepatology*, 2019, 69: 2672-2682.
44
45 [10] Younossi ZM, Koenig AB, Abdelatif D, et al. Global epidemiology of nonalcoholic
46 fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes.
47 *Hepatology*, 2016, 64: 73-84.
48
49 [11] Shen F, Zheng RD, Shi JP, et al. Impact of skin capsular distance on the performance
50 of controlled attenuation parameter in patients with chronic liver disease. *Liver Int*,
51 2015, 35: 2392-2400.
52
53 [12] Ascha MS, Hanouneh IA, Lopez R, et al. The incidence and risk factors of
54 hepatocellular carcinoma in patients with nonalcoholic steatohepatitis. *Hepatology*,
55 2010, 51: 1972-1978.
56
57 [13] Younossi ZM, Otgonsuren M, Henry L, et al. Association of nonalcoholic fatty liver
58 disease (NAFLD) with hepatocellular carcinoma (HCC) in the United States from 2004
59 to 2009. *Hepatology*, 2015, 62: 1723-1730.
60

- 1
2
3
4 [14] White DL, Kanwal F, El-Serag HB. Association between nonalcoholic fatty liver
5 disease and risk for hepatocellular cancer, based on systematic review. Clin
6 Gastroenterol Hepatol, 2012, 10: 1342-1359.e1342.
7
8 [15] Nasr P, Ignatova S, Kechagias S, et al. Natural history of nonalcoholic fatty liver
9 disease: A prospective follow-up study with serial biopsies. Hepatol Commun, 2018, 2:
10 199-210.
11
12 [16] Estes C, Razavi H, Loomba R, et al. Modeling the epidemic of nonalcoholic fatty liver
13 disease demonstrates an exponential increase in burden of disease. Hepatology, 2018,
14 67: 123-133.
15
16 [17] Williamson RM, Price JF, Glancy S, et al. Prevalence of and risk factors for hepatic
17 steatosis and nonalcoholic Fatty liver disease in people with type 2 diabetes: the
18 Edinburgh Type 2 Diabetes Study. Diabetes Care, 2011, 34: 1139-1144.
19
20 [18] Targher G, Bertolini L, Padovani R, et al. Prevalence of nonalcoholic fatty liver disease
21 and its association with cardiovascular disease among type 2 diabetic patients. Diabetes
22 Care, 2007, 30: 1212-1218.
23
24 [19] Miyake T, Kumagi T, Furukawa S, et al. Non-alcoholic fatty liver disease: factors
25 associated with its presence and onset. J Gastroenterol Hepatol, 2013, 28 Suppl 4: 71-
26 78.
27
28
29
30

31 **Figure legends**

32 **Figure 1.** Flow chart depicting the enrollment of a population with non-alcoholic fatty liver
33 disease (NAFLD) for follow-up in Ningbo Zhenhai Lianhua Hospital, China.
34
35
36

37 **Figure 2.** Penalized logistic regression and Cox regression analysis were performed for risk
38 factors and hazard ratios of non-alcoholic fatty liver disease (NAFLD). The following
39 parameters were available from 5606 participants: Gender, age, BMI, albumin, white globulin
40 ratio (WGR), white blood cell (WBC), low density lipoprotein (LDL), triglycerides (TG), high
41 density lipoprotein (HDL), glutamyl transpeptidase (GLT), alanine transaminase (ALT),
42 aspartate transaminase (AST), creatinine (Cr), alkaline phosphatase (ALP), blood urea nitrogen
43 (Bun), Uric Acid (UA), blood glucose (Glu), systolic blood pressure (SBp), diastolic blood
44 pressure (DBp), blood sedimentation (ESR), hemoglobin (HGB), platelets (PLT),
45 Apolipoprotein A1 (ApoA1), Apolipoprotein B2 (ApoB), total bilirubin (TB), total protein
46 (TP). Cross validation selected 16 variables to be potential predictors. The corresponding
47 forest-plot is shown in (A). The AUC of these above 16 variables for NAFLD is 0.88 (B). Cox
48 regression confirmed that the 16 variables were relevant for NAFLD incidence, including BMI,
49 albumin, WBC, TG, HDL, GLT, ALT, Cr, UA, Glu, SBp, DBp, ESR, HGB, PLT and ApoB.
50 The corresponding forest-plot is shown (C).
51
52
53
54
55
56
57
58
59
60

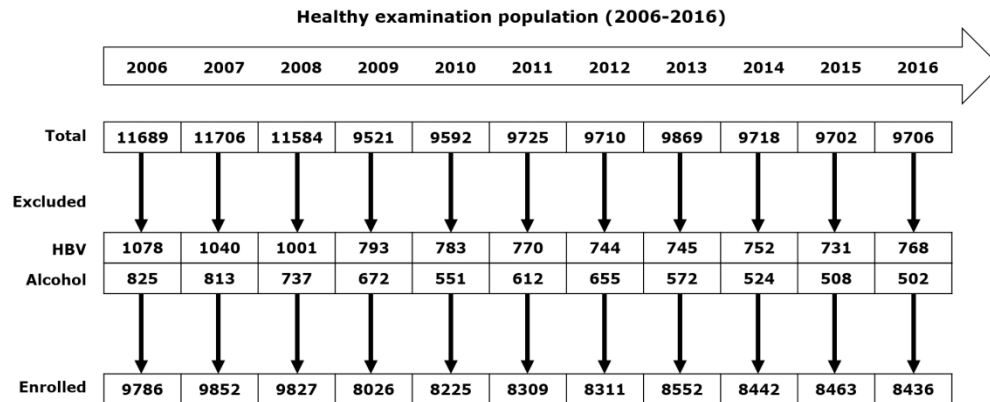
1
2
3
4
5 **Figure 3.** Dynamic Bayesian network analyses were performed to show the cause-effect link
6 between non-alcoholic fatty liver disease (NAFLD) and its potential risk factors. Three
7 variables, BMI, gender and triglycerides directly pointed to NAFLD. ApoB impacted on the
8 incidence of NAFLD through TG abundance. LDL indirectly contributed to NAFLD through
9 ApoB. NAFLD directly led to alterations of seven clinical parameters: ALT, DBp, SBp, TP,
10 albumin, HGB and UA.
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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											
(n)	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

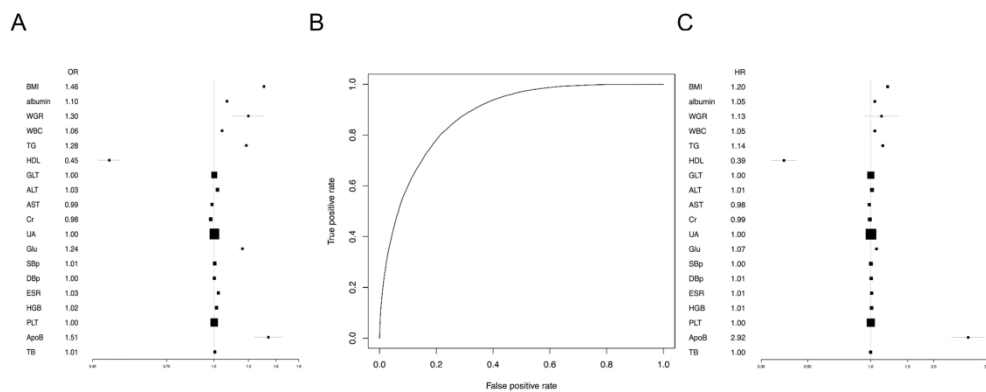
Table 2. Clinical outcome of NAFLD patients
2006 - 2016

		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)	

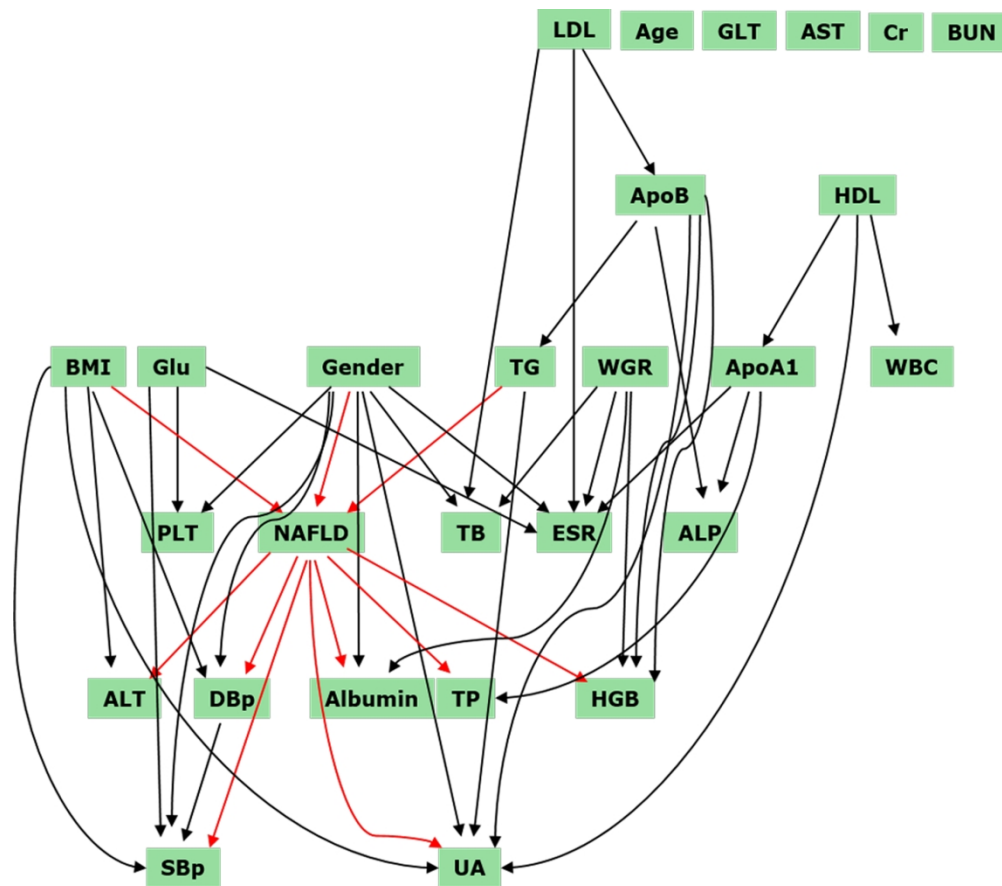


Flow chart depicting the enrollment of a population with non-alcoholic fatty liver disease (NAFLD) for follow-up in Ningbo Zhenhai Lianhua Hospital, China.

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 triglycerides 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)

1
2
3 **Nonalcoholic fatty liver disease increased risk of metabolic diseases, but not severe liver**
4 **disease: an eleven-year follow-up study**
5

6 Xiaoping Tang, Yanyan Shi, Juan Du, Keming Hu, Tingting Zhou, Lan Chen, Yanming
7 Zhang, Fujun Li, Huier Zhang, Roman Liebe, Christoph Meyer, Steven Dooley, Zhongwei
8 Zhu, Hong-Lei Weng, Jinzhu Jia, Tong Huang
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

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)

1	
2	
3	TP (total protein)
4	UA (uric acid)
5	Waist
6	Weight
7	WGR (white globulin ratio)
8	WBC (white blood cell count)
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49	
50	
51	
52	
53	
54	
55	
56	
57	
58	
59	
60	

For peer review only

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 (n)	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 (n)	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 (n)	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	0.9939885	< 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) <i>Cohort study</i> —Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up <i>Case-control study</i> —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 <i>Cross-sectional study</i> —Give the eligibility criteria, and the sources and methods of selection of participants (b) <i>Cohort study</i> —For matched studies, give matching criteria and number of exposed and unexposed <i>Case-control study</i> —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) <i>Cohort study</i> —If applicable, explain how loss to follow-up was addressed <i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed <i>Cross-sectional study</i> —If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses	

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Continued on next page

For peer review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60**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) <i>Cohort study</i> —Summarise follow-up time (eg, average and total amount)	
Outcome data	15 *	<i>Cohort study</i> —Report numbers of outcome events or summary measures over time <i>Case-control study</i> —Report numbers in each exposure category, or summary measures of exposure <i>Cross-sectional study</i> —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:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2021-054891.R1
Article Type:	Original research
Date Submitted by the Author:	17-Mar-2022
Complete List of Authors:	<p>Tang, Xiaoping ; Ningbo Zhenhai District Lianhua Hospital , Department of Prevention and Health Care</p> <p>Shi, Yanyan; Peking University Third Hospital, Research Center of Clinical Epidemiology</p> <p>Du, Juan ; Ningbo Zhenhai District Lianhua Hospital , Department of Internal Medicine</p> <p>Hu, Keming; Ningbo Zhenhai District Lianhua Hospital, Department of Prevention and Health Care</p> <p>Zhou, Tingting; Ningbo Zhenhai District Lianhua Hospital , Department of Internal Medicine</p> <p>Chen, Lan; Ningbo Zhenhai District Lianhua Hospital , Department of Radiology</p> <p>Zhang, Yanming; Ningbo Zhenhai District Lianhua Hospital , Center Laboratory</p> <p>Li, Fujun; Ningbo Zhenhai District Lianhua Hospital , Department of Radiology</p> <p>Zhang, Huier ; Ningbo Zhenhai District Lianhua Hospital , Center Laboratory</p> <p>Liebe, Roman ; Heinrich Heine University Düsseldorf, Clinic of Gastroenterology, Hepatology and Infectious Diseases,; Saarland University, Department of Medicine II, Saarland University Medical Center</p> <p>Meyer, Christoph ; Heidelberg University, Department of Medicine II, Medical Faculty Mannheim</p> <p>Dooley, Steven ; Heidelberg University, Department of Medicine II, Medical Faculty Mannheim</p> <p>Zhu, Zhongwei; Ningbo Zhenhai District Lianhua Hospital , Department of Surgery</p> <p>Weng, Hong-Lei ; Heidelberg University, Department of Medicine II, Medical Faculty Mannheim</p> <p>JIA, Jinzhu; Peking University, Department of Biostatistics, School of Public Health; Peking University, Center for Statistical Science</p> <p>Huang, Tong; Ningbo City First Hospital; Ningbo Zhenhai District Lianhua Hospital , Department of Prevention and Health Care</p>
Primary Subject Heading:	Infectious diseases
Secondary Subject Heading:	Infectious diseases
Keywords:	Epidemiology < INFECTIOUS DISEASES, Hypertension < CARDIOLOGY,

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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

SCHOLARONE™
Manuscripts



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 Hu¹, Tingting Zhou³, Lan Chen⁴, Yanming Zhang⁵, Fujun Li⁴, Huier Zhang⁵, Roman Liebe^{6,7}, Christoph Meyer⁸, Steven Dooley⁸, Zhongwei Zhu⁹, Hong-Lei Weng⁸, 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

1
2
3 35
4
5 36 Jinzhu Jia, PhD.
6
7 37 Department of Biostatistics, School of Public Health, Peking University Health Science
8
9 38 Center
10
11 39 Address: No. 38, Xueyuan Road, Haidian District, Beijing, 100191, P.R. China
12
13 40 Email: jzjia@math.pku.edu.cn
14
15 41 Tel: +86-15801049187
16
17 42

17 43 **Acknowledgements:**

18
19 44 This project is supported by the National Key R&D Program of China No. 2017YFC0908103
20
21 45 (X.T., and J.D.). We are grateful to Drs. Changxi Chen, Jingyi Yuan and Yongjun Chen for
22
23 46 support and discussion.
24
25 47
26
27 48

1
2
3 **49 Abstract**

4 50
5 51 **Objectives** To clarify NAFLD prevalence, risk factors, and clinical outcome in an exemplary
6 52 Chinese population, a cohort of company employees was followed up for eleven years.

7 53 **Design** Retrospective cohort

8 54 **Setting** Between 2006-2016 in China

9 55 **Participants** 13032 company employees

10 56 **Results** Over eleven years, the prevalence of NAFLD increased from 17.2% to 32.4% (males
11 57 20.5% to 37% *versus* females 9.8% to 22.2%). Male peak prevalence was between 40 and 60
12 58 years of age, whereas highest prevalence in women was at an age of 60 years and older. Logistic
13 59 and Cox regression revealed 16 risk factors, including BMI, albumin, WBC, TG, HDL, GLT,
14 60 ALT, Cr, UA, Glu, SBp, DBp, ESR, HGB, PLT and ApoB. The AUC of these variables for
15 61 NAFLD is 0.88. However, cause-effect analyses showed that only BMI, gender and
16 62 triglycerides directly contributed to NAFLD development. Over an 11-year follow-up period,
17 63 12.6%, 37.7% and 14.2% of male NAFLD patients and 11.6%, 44.7% and 22.6% of female
18 64 NAFLD patients developed diabetes, hypertension and hyperuricemia, respectively. Except one
19 65 male patient who developed cirrhosis, no NAFLD patients progressed into severe liver disease.

20 66 **Conclusion** Diabetes, hypertension and hyperuricemia are the main clinical outcomes of
21 67 NAFLD. Eleven years of NAFLD are not sufficient to cause severe liver disease. Age and
22 68 obesity are direct risk factors for NAFLD. BMI, gender and triglycerides are three parameters
23 69 directly reflecting the occurrence of NAFLD.

24 70

25 71 **(Words: 221)**

26 72

27 73 **Strengths and limitations of this study**

- 28 74 ● This study dynamically follows up NAFLD prevalence in an eastern Chinese community
29 75 for eleven years.
- 30 76 ● The study adopted First order Markov models to evaluate the cause-effect link between
31 77 NAFLD and risk factors.
- 32 78 ● The relatively low sensitivity of ultrasound for the detection of liver fat might
33 79 underestimate the incidence of NAFLD in this cohort.
- 34 80 ● Given that the current study is a single-center observation, multiple-center studies are
35 81 required to confirm the conclusions in the future.

- 1
2
3 82 ● The study population is a highly select, relatively homogenous group of well-educated
4 professionals in privileged social positions and permanent employment. Thus, the
5 83
6 84 conclusions might not be transferable to the general Chinese population.
7
8 85
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

86 Introduction

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

96
97 In China, the incidence of NAFLD has been increasing over the last two decades. A recent
98 meta-analysis, based on 392 studies between 2008 and 2018, showed the national incidence of
99 NAFLD in China to be at 29.2% [6]. In Shanghai, the adult incidence of NAFLD has increased
100 from 14.04% in 1995 to 43.65% in 2015 [2]. Being a vast country, Chinese living in different
101 areas vary widely in lifestyle and economic status. Thus, the epidemiology, natural history and
102 clinical outcomes of NAFLD in different areas of the country are worth further investigation.

103
104 It is well accepted that viral hepatitis is a major reason for progressive chronic liver diseases,
105 e.g., fibrosis, cirrhosis and ultimately, hepatocellular carcinoma (HCC). With regard to
106 NAFLD, incidence and severity of associated chronic liver disease outcomes has not been
107 monitored in large Chinese cohorts yet – especially over an extended time span. The current
108 study therefore describes the prevalence of NAFLD in a large Eastern Chinese community over
109 eleven years (2006 - 2016). We focused on three aspects: (1) annual prevalence of NAFLD, (2)
110 risk factors of NAFLD, and (3) intra- and extra-hepatic clinical outcomes of NAFLD.

113 Methods

114 *Patient and public involvement*

115 No patients were involved in this study

116 *Design and participants*

117
118 In this retrospective study we analyzed the "annual health examination database" of the Zhenhai
119 Lianhua Hospital from 2006 to 2016. This hospital is affiliated to Sinopec Zhenhai Refining &
120 Chemical Company. Supported by the company, all employees were offered the opportunity to
121

1
2
3 122 go to this hospital for an annual health examination. During 11 years, a total 13,032 employees
4
5 123 received health examinations. From 2006 to 2016, 11689, 11706, 11584, 9521, 9592, 9725,
6
7 124 9710, 9869, 9718, 9702 and 9706 persons received health examinations (**Figure 1**). To describe
8
9 125 the longitudinal NAFLD occurrence in this cohort, we excluded subjects with the following
10
11 126 conditions: (1) viral hepatitis B and C infection, which were identified by blood virus
12
13 127 measurements (HBV-DNA and HCV-RNA), and (2) alcoholic liver disease, which was defined
14
15 128 as previously described [7,8]. NAFLD was defined as the presence of hepatic steatosis,
16
17 129 determined by ultrasonography.

17 130 The study protocol was approved by the Ethics Committee of Zhenhai Lianhua Hospital
18
19 131 ([2016]001). Informed consent was obtained from all subjects.
20
21 132

22 133

23 133 *Measures*

24 134 **Supplementary Table 1** shows all parameters measured in the annual health examinations.
25
26 135 Ultrasonography was performed by the same three experienced doctors (L.C., F. L., and J.Y.)
27
28 136 with an Ultrasonograph B, GE, Voluson 730 pro. Blood biochemistry and HBV serum levels
29
30 137 were measured by an Olympus AU640 autoanalyzer (Olympus, Kobe, Japan) and an
31
32 138 ImmunoAssay Analyzer VitrosECI (JOHNSON-JOHNSON, USA), respectively. All methods
33
34 139 were carried out in accordance with relevant guidelines and regulations.
35
36 140

37 141 *Statistical analysis*

38 142 For population characteristics, variables were described as means and standard deviation (SD)
39
40 143 or proportions as appropriate. Student's t-test or nonparametric test was used to analyze
41
42 144 differences between two groups as mentioned. Chi-square test was used to verify the differences
43
44 145 of nominative variables between two groups. Multivariate analysis of risk factors for NAFLD
45
46 146 was performed using logistic regression analysis. Combined receiver operating characteristic
47
48 147 curve (ROC) and area under curve (AUC) analyses were used to evaluate the diagnostic
49
50 148 performance of biomarkers based on the logistic regression model. Multivariate Cox regression
51
52 149 model was performed to calculate hazard ratios of variables to identify independent prognostic
53
54 150 variables. First order Markov models were used to analyze the cause-effect link between
55
56 151 NAFLD and risk factors. L1 penalized logistic regression was applied to select predictive
57
58 152 predictors. R package "glmnet" contains functions to select predictors using L1 penalized
59
60 153 logistic regression. Statistical analyses were performed using SPSS 22.0 and R version 3.5.3.
154
155 154 *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'.

1
2
3 156
4
5 157
6
7 158
8 159
9 160
10
11 161
12 162
13
14 163
15
16 164
17 165
18 166
19 167
20 168
21 169
22 170
23 171
24 172
25 173
26 174
27 175
28 176
29 177
30 178
31 179
32
33 180
34 181
35 182
36 183
37 184
38 185
39 186
40 187
41 188
42 189
43 190

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**).

191 *BMI and NAFLD incidence*

192 Given the tight link between obesity and NAFLD, we paid special attention to the population
193 with high BMI. We focused on the 5606 persons with complete follow-up and analyzed the
194 prevalence of NAFLD in those with BMI >25. In total, out of the 5606 persons, 2445 presented
195 with a BMI of >25. The prevalence of NAFLD in this overweight subpopulation was far higher
196 than in the general population. In 2006, 45.2% of individuals (n=1104; male vs. female: 47.3%
197 vs. 37.1%) with BMI >25 were suffering from NAFLD (**Supplementary Table 3**). In 2016,
198 values reached 67.1% (n=1414; male vs. female: 69% vs. 59.2%, **Supplementary Table 3**).
199 Impressively, the NAFLD prevalence in both genders was very high at any age, even in those
200 below the age of 30 years. In 2006, among 213 overweight men, younger than 30 years, 52.6%
201 were also diagnosed for NAFLD (**Supplementary Table 3**). This number increased to 63% in
202 2016 (**Supplementary Table 3**). In 2006, there were 15 overweight women aged less than 30
203 years. Among them, 3 presented as NAFLD (20%). In 2016, 7 out of 16 overweight women
204 aged less than 30 years were identified. The NAFLD prevalence had increased to 43.8%
205 (**Supplementary Table 3**). In those older than 40 years, NAFLD prevalence increased from
206 36.6 – 45.4% in 2006 to 53 – 65.6% (**Supplementary Table 3**).

208 *Risk factors relevant to NAFLD occurrence*

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

1
2
3 225 occurrence. Cox regression confirmed that the 16 parameters were significantly relevant to
4 226 NAFLD incidence (**Supplementary Table 5** and **Figure 2C**). Furthermore, ApoB and HDL
5 227 were the most robust positive and negative risk factors for NAFLD (**Figure 2C**).
6
7
8
9

10
11

12 229 *Cause-effect link between risk factors and NAFLD occurrence*

13 230 Although the aforementioned parameters were regarded as “risk factors” according to statistical
14 231 models, it did not necessarily mean that all of them contributed to NAFLD occurrence. Based
15 232 on 11 years of longitudinal data, it was possible to construct a dynamic Bayesian network to
16 233 identify the risk factors most relevant to NAFLD occurrence. As shown in **Figure 3**, these
17 234 parameters constituted a complicated, but clear intercross paradigm. Only three parameters,
18 235 BMI, gender and TG directly pointed to NAFLD. In addition, ApoB impacted the incidence of
19 236 NAFLD through contributing to TG. Furthermore, LDL can indirectly contribute to NAFLD
20 237 through influencing ApoB. Very impressively, the dynamic Bayesian network pointed out that
21 238 NAFLD directly leads to alterations of seven parameters: ALT, DBp, SBp, TP, albumin, HGB
22 239 and UA. Intriguingly, age, GLT, AST, Cr, and BUN did not interact with any other parameter
23 240 in our model, indicating that these factors by incidence correlate, but not any causal interaction
24 241 is existing.
25
26
27
28
29
30
31
32
33

34

35 243 *Outcome of NAFLD*

36 244 Subsequently, we examined clinical outcomes of NAFLD over the eleven years. **Table 2**
37 245 summarizes the incidence of intra- and extrahepatic diseases of 696 NAFLD and 222 NASH
38 246 patients during the follow-up period. Among the total NAFLD and NASH population, only 1
39 247 male NAFLD patient developed liver cirrhosis within the eleven years. However, this time span
40 248 witnessed significantly increased extrahepatic diseases, including diabetes, hypertension and
41 249 hyperuricemia. In the NAFLD population, there were 64 (12.6%) men and 22 (11.6%) women,
42 250 191 (37.7%) men and 85 (44.7%) women, 72 (14.2%) men and 43 (22.6%) women who
43 251 developed into type II diabetes, hypertension and hyperuricemia, respectively (**Table 2**).
44
45
46
47
48
49

50 252 Given that 2006 is the starting point of data collection, patients diagnosed for NAFLD in this
51 253 year had by no means just manifested their disease, but rather patients had possibly developed
52 254 NAFLD several years prior to inclusion. To clarify the exact clinical outcomes of NAFLD over
53 255 one decade, we focused on the following two cohorts of individuals with annual health
54 256 examinations for 11 years: (1) patients who were diagnosed as non-NAFLD in 2006, but were
55 257 NAFLD in 2007 (new NAFLD cohort); and (2) who were non-NAFLD in both 2006 and 2007
56 258 (non-NAFLD cohort). As shown in **Table 2**, 185 new NAFLD cases (138 men and 47 women)

1
2
3 259 and 4547 non-NAFLD (2786 men and 1761 women) persons were found in 2007. Between
4
5 260 2007 and 2016, neither NAFLD nor non-NAFLD individuals developed liver cirrhosis or
6
7 261 cancer. However, the one-decade follow-up reveals different prevalences of diabetes,
8
9 262 hypertension and hyperuricemia: In NAFLD patients, there were 14 (10.1%) men and 5 (10.6%)
10
11 263 women, 47 (34.1%) men and 21 (44.7%) women, 34 (24.6%) men and 8 (17%) women who
12
13 264 developed type II diabetes, hypertension and hyperuricemia, respectively (**Table 2**). In non-
14
15 265 NAFLD individuals, 157 (5.6%) men and 54 (3.4%) women, 259 (9.3%) men and 324 (18.4%),
16
17 266 284 (10.2%) men and 84 women (4.8%) developed type II diabetes, hypertension and
18
19 267 hyperuricemia, respectively (**Table 2**). For all three diseases, statistically significant differences
20
21 268 were determined between the two cohorts of population (all $P < 0.05$, **Table 2**). These results
22
23 269 suggest that diabetes, hypertension and hyperuricemia are the main clinical outcomes of
24
25 270 NAFLD.
26
27 271

272 273 **Discussion**

274 This eleven-year follow-up retrospective study reports the following: (1) NAFLD prevalence
275 has substantially increased in the examined Eastern Chinese population. (2) The prevalence of
276 NAFLD differs by gender and age. Middle-aged men and elderly women are the two
277 populations at highest risk for NAFLD. (3) Gender, BMI and triglycerides are the parameters
278 directly associated with NAFLD occurrence. Regardless of gender and age, persons with high
279 BMI (≥ 25) have a high risk for NAFLD development. (4) NAFLD directly leads to alterations
280 of seven clinical parameters: ALT, DBp, SBp, TP, albumin, HGB and UA. (5) Within 11 years,
281 a significant part of the NAFLD population develops three clinically relevant diseases: type 2
282 diabetes mellitus, hypertension and hyperuricemia. (6) Within 11 years, NAFLD does not cause
283 severe liver disease, such as cirrhosis or HCC, in patients.
284

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

1
2
3 294 It was estimated that the cumulative incidence of HCC among patients with NAFLD and
4
5 295 cirrhosis ranges from 2.4% to 12.8% over a median follow-up period of 3.2 to 7.2 years ^[14]
6
7 296 (Global Health Observatory) data. Mortality and global health estimates were obtained from:
8
9 297 http://www.who.int/gho/mortality_burden_disease/en/. Last accessed on 1/7/2020.). Given that
10
11 298 the above conclusions were based on cross-sectional investigations and statistical models, it has
12
13 299 been unknown to date over which period a NAFLD patient develops liver cirrhosis or HCC
14
15 300 (personal risk assessment). Our 11-year follow-up provides therefore a valuable and
16
17 301 comprehensive dataset. In this study, most patients were diagnosed with NAFLD when they
18
19 302 received a routine health examination. Before the examination, these people did not have any
20
21 303 symptoms or signs of NAFLD. Therefore, they belong to NAFLD patients at a very early stage
22
23 304 (although for 2006, the duration of pre-existing NAFLD cannot be determined). Except for a
24
25 305 single person, no serious liver problems were observed within this time period. These data
26
27 306 suggest that for the vast majority of patients with early stage NAFLD, 11 years are not sufficient
28
29 307 to develop liver cirrhosis or cancer. Nasr et al followed up 129 NAFLD patients with varying
30
31 308 fibrosis stages on two occasions (mean time 13.7 and 9.3 years). Liver biopsy analyses showed
32
33 309 that 9.3% of patients developed end-stage liver disease and 34% advanced fibrosis ^[15]. The
34
35 310 NAFLD patients observed by Nasr et al. actually belonged to the NASH category, because they
36
37 311 suffered from fibrosis and elevated ALT and/or AST levels. As our study was based on
38
39 312 examinations of healthy, liver biopsy is not justifiable. Very likely, the current cohort included
40
41 313 a portion of NASH patients. They also did not show significant progression towards cirrhosis
42
43 314 or HCC was monitored.

44
45 315
46
47 316 In contrast to hepatic complications, patients with NAFLD showed a significant risk for the
48
49 317 development of extrahepatic diseases, including diabetes, hypertension and hyperuricemia. In
50
51 318 696 NAFLD patients, eleven years witnessed the development of 86 cases (12.4%) type II
52
53 319 diabetes, 276 (40%) cases of hypertension and 115 patients with (16.5%) hyperuricemia,
54
55 320 respectively. Interestingly, in 222 NASH patients, the prevalence of these three diseases was
56
57 321 12 (5.4%), 46 (20.7%), and 33 (14.9%) only. In general, men had a higher probability to develop
58
59 322 these diseases than women. These results are consistent with previous reports from USA and
60
61 323 Europe ^[16-18]. Whether NAFLD is associated with the risk of severe heart or brain diseases
62
63 324 such as acute myocardial infarction (AMI) and stroke is worth further investigation. A recent
64
65 325 matched cohort study analyzed databases from four European countries, which included 17.7
66
67 326 million patients with NAFLD or NASH ^[19]. These patients had a mean follow-up of 2.1 to 5.5
68
69 327 years. The study showed that the diagnosis of NAFLD appears not to be associated with AMI

1
2
3 328 or stroke risk after adjustment for established cardiovascular risk factors. Nevertheless, the
4
5 329 authors mentioned that cardiovascular risk assessment in adults with a diagnosis of NAFLD is
6
7 330 important [19]. Follow-up for 5 years might be not sufficient to reach a conclusion for this issue.
8
9 331

10 332 An important issue is the cause-and-effect relationship between NAFLD and its clinical
11
12 333 outcomes such as diabetes, hypertension and hyperuricemia. A dynamic Bayesian network in
13
14 334 the current study provides direct evidence on this issue: NAFLD directly results in alterations
15
16 335 of several parameters, including DBp, SBp and UA, suggesting that NAFLD directly
17
18 336 contributes to the occurrence of hypertension and hyperuricemia. The underlying mechanisms
19
20 337 require further investigation.

21 338 The current dynamic Bayesian network analysis does not confirm a direct cause-and-
22
23 339 effect relationship between NAFLD and type 2 diabetes mellitus. There are plenty of studies
24
25 340 showing the close relationship between type 2 diabetes and NAFLD [20]. Pathophysiologically,
26
27 341 insulin resistance is a key event in both NAFLD and diabetes progression [20]. However,
28
29 342 genome-wide association studies have not yet identified the exact impact of insulin resistance
30
31 343 on the variants associated with NAFLD severity [20,21]. Clarification of the cause-and-effect
32
33 344 relationship between NAFLD and diabetes requires further long-term follow-up studies.

34 345
35 346 To date, there are a large number of studies investigating risk factors for NAFLD [22]. These
36
37 347 studies tried to identify single, or multiple combined biomarkers to predict NAFLD occurrence.
38
39 348 Given that most studies were based on cross-sectional designs, or with only short follow-up
40
41 349 periods, it is difficult to clarify the causality between the proposed predictors and NAFLD
42
43 350 morbidity. Our eleven-year dataset provides a chance to shed led on this issue. Here, the
44
45 351 dynamic causal relationships between variables, including risk parameters and clinical
46
47 352 outcomes, were identified by a first order Markov model, which was displayed by a dynamic
48
49 353 Bayes network. The dynamic Bayes model discriminates causal relationship through time
50
51 354 sequence. When a variable change is closely related to a previous variance alteration, a causal
52
53 355 relationship between the two variables is assumed. Based on Logistic and Cox regression and
54
55 356 dynamic Bayesian network analyses, we confirmed three direct risk factors for NAFLD
56
57 357 occurrence: gender, BMI and TG. These findings are supported by the following data: (1) Men
58
59 358 have higher NAFLD prevalence than women in this population (37% *versus* 22.2% in 2016);
60
359 (2) In overweight people with a BMI >25, NAFLD prevalence reached 69% in males and 59.2%
360 in females. Given that triglycerides are a major energy source, but are leading to obesity, it is

1
2
3 361 not surprising that this parameter directly reflects the risk for NAFLD development. These
4
5 362 findings provide robust evidence supporting the use of BMI to monitor or predict NAFLD.

6 363

7
8 364

9
10 365 **Conclusion**

11 366 This 11-year follow-up study documents the rapid increase in NAFLD prevalence in an Eastern
12
13 367 Chinese population. In contrast to previous reports, our observation does not observe that one
14
15 368 decade of NAFLD is sufficient to lead to severe hepatic clinical outcomes. It is worthy to note
16
17 369 that our population were biased because they are on the well-off, well-educated side of the
18
19 370 Chinese people, while previous studies were often based on hospital populations, who suffered
20
21 371 from negative selection bias and thus came up with higher estimates. In addition, given there
22
23 372 are differences in NAFLD profiles between Eastern and Western populations, it would be
24
25 373 interesting to know the natural development of NAFLD in a Western population. A key point
26
27 374 for clarifying the true history of NAFLD is to follow a population starting from the early phases
28
29 375 of the disease. Consistent with previous studies, NAFLD is tightly associated with multiple
30
31 376 extra-hepatic diseases relevant to the metabolic syndrome. In the future, follow-up of the
32
33 377 current cohort for another one and two decades will provide further valuable data to clarify the
34
35 378 extended natural history of NAFLD. Last but not least, a large portion of the men and women
36
37 379 in this study were educated above the average and have a position in the company that gave
38
39 380 them the availability of better food choices as well as regular sport. On the other hand, the
40
41 381 relatively low sensitivity of ultrasound for the detection of liver fat might underestimate the
42
43 382 incidence of NAFLD in this cohort.

44 383

45 384 **Contributorship:**

46 385 Conception and design: Xiaoping Tang, Yanyan Shi, Juan Du, Hong-Lei Weng, Jinzhu Jia
47
48 386 and Tong Huang

49 387 Ultrasonography: Lan Chen, Fujun Li

50 388 Blood assays: Yanming Zhang and Huier Zhang

51
52 389 Other examinations and data collection: Xiaoping Tang, Keming Hu, Tingting Zhou, Juan Du,
53
54 390 Zhongwei Zhu and Tong Huang

55
56 391 Statistical analyses: Yanyan Shi and Jinzhu Jia

57
58 392 Drafting the article: Hong-Lei Weng

1
2
3 393 Reviewing and editing the article critically: Xiaoping Tang, Yanyan Shi, Christoph Meyer,
4
5 394 Roman Liebe, Steven Dooley, Hong-Lei Weng, Jinzhu Jia and Tong Huang
6
7
8 395

9 396 **Funding Statement**

11 397 The project is supported by National Key R&D Program of China (2017YFC0908103).
12
13 398

14 399 **Competing of Interests**

16 400 The authors declare that there is no conflict of interest.
17
18 401

19 402 **Ethics approval**

21 403 The study protocol was approved by the Ethics Committee of Zhenhai Lianhua Hospital No.
22 404 2016(001). Informed consent was obtained from all subjects.
23
24 405

26 406 **Data sharing**

28 407 All data relevant to the study are included in the article or uploaded as supplementary
29 408 information.
30
31 409

32 410 **References**

- 35 411 [1] Younossi ZM, Tampi R, Priyadarshini M, et al. Burden of Illness and Economic Model
36 412 for Patients With Nonalcoholic Steatohepatitis in the United States. *Hepatology*, 2019,
37 413 69: 564-572.
38
39 414 [2] Younossi ZM, Loomba R, Anstee QM, et al. Diagnostic modalities for nonalcoholic
40 415 fatty liver disease, nonalcoholic steatohepatitis, and associated fibrosis. *Hepatology*,
41 416 2018, 68: 349-360.
42 417 [3] Fan JG, Kim SU, Wong VW. New trends on obesity and NAFLD in Asia. *J Hepatol*,
43 418 2017, 67: 862-873.
44 419 [4] Byrne CD, Targher G. NAFLD: a multisystem disease. *J Hepatol*, 2015, 62: S47-64.
45 420 [5] Alexander M, Loomis AK, van der Lei J, et al. Risks and clinical predictors of cirrhosis
46 421 and hepatocellular carcinoma diagnoses in adults with diagnosed NAFLD: real-world
47 422 study of 18 million patients in four European cohorts. *BMC Med*, 2019, 17: 95.
48
49 423 [6] Zhou F, Zhou J, Wang W, et al. Unexpected Rapid Increase in the Burden of NAFLD
50 424 in China From 2008 to 2018: A Systematic Review and Meta-Analysis. *Hepatology*,
51 425 2019, 70: 1119-1133.
52 426 [7] EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic
53 427 fatty liver disease. *J Hepatol*, 2016, 64: 1388-1402.
54
55 428 [8] Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of
56 429 nonalcoholic fatty liver disease: Practice guidance from the American Association for
57 430 the Study of Liver Diseases. *Hepatology*, 2018, 67: 328-357.
58 431 [9] Younossi Z, Tacke F, Arrese M, et al. Global Perspectives on Nonalcoholic Fatty Liver
59 432 Disease and Nonalcoholic Steatohepatitis. *Hepatology*, 2019, 69: 2672-2682.

- 1
2
3 433 [10] Younossi ZM, Koenig AB, Abdelatif D, et al. Global epidemiology of nonalcoholic
4 434 fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes.
5 435 Hepatology, 2016, 64: 73-84.
6 436 [11] Shen F, Zheng RD, Shi JP, et al. Impact of skin capsular distance on the performance
7 437 of controlled attenuation parameter in patients with chronic liver disease. Liver Int,
8 438 2015, 35: 2392-2400.
9 439 [12] Ascha MS, Hanouneh IA, Lopez R, et al. The incidence and risk factors of
10 440 hepatocellular carcinoma in patients with nonalcoholic steatohepatitis. Hepatology,
11 441 2010, 51: 1972-1978.
12 442 [13] Younossi ZM, Otgonsuren M, Henry L, et al. Association of nonalcoholic fatty liver
13 443 disease (NAFLD) with hepatocellular carcinoma (HCC) in the United States from 2004
14 444 to 2009. Hepatology, 2015, 62: 1723-1730.
15 445 [14] White DL, Kanwal F, El-Serag HB. Association between nonalcoholic fatty liver
16 446 disease and risk for hepatocellular cancer, based on systematic review. Clin
17 447 Gastroenterol Hepatol, 2012, 10: 1342-1359.e1342.
18 448 [15] Nasr P, Ignatova S, Kechagias S, et al. Natural history of nonalcoholic fatty liver
19 449 disease: A prospective follow-up study with serial biopsies. Hepatol Commun, 2018, 2:
20 450 199-210.
21 451 [16] Estes C, Razavi H, Loomba R, et al. Modeling the epidemic of nonalcoholic fatty liver
22 452 disease demonstrates an exponential increase in burden of disease. Hepatology, 2018,
23 453 67: 123-133.
24 454 [17] Williamson RM, Price JF, Glancy S, et al. Prevalence of and risk factors for hepatic
25 455 steatosis and nonalcoholic Fatty liver disease in people with type 2 diabetes: the
26 456 Edinburgh Type 2 Diabetes Study. Diabetes Care, 2011, 34: 1139-1144.
27 457 [18] Targher G, Bertolini L, Padovani R, et al. Prevalence of nonalcoholic fatty liver disease
28 458 and its association with cardiovascular disease among type 2 diabetic patients. Diabetes
29 459 Care, 2007, 30: 1212-1218.
30 460 [19] Alexander M, Loomis AK, van der Lei J, et al. Non-alcoholic fatty liver disease and risk
31 461 of incident acute myocardial infarction and stroke: findings from matched cohort study
32 462 of 18 million European adults. Bmj, 2019, 367: 15367.
33 463 [20] Tilg H, Moschen AR, Roden M. NAFLD and diabetes mellitus. Nat Rev Gastroenterol
34 464 Hepatol, 2017, 14: 32-42.
35 465 [21] Zhou Y, Llauradó G, Orešič M, et al. Circulating triacylglycerol signatures and insulin
36 466 sensitivity in NAFLD associated with the E167K variant in TM6SF2. J Hepatol, 2015,
37 467 62: 657-663.
38 468 [22] Miyake T, Kumagi T, Furukawa S, et al. Non-alcoholic fatty liver disease: factors
39 469 associated with its presence and onset. J Gastroenterol Hepatol, 2013, 28 Suppl 4: 71-
40 470 78.
41 471
42 472
43 473

Figure legends

474 **Figure 1.** Flow chart depicting the enrollment of a population with non-alcoholic fatty liver
475 disease (NAFLD) for follow-up in Ningbo Zhenhai Lianhua Hospital, China.
476

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

1
2
3 480 ratio (WGR), white blood cell (WBC), low density lipoprotein (LDL), triglycerides (TG), high
4
5 481 density lipoprotein (HDL), glutamyl transpeptidase (GLT), alanine transaminase (ALT),
6
7 482 aspartate transaminase (AST), creatinine (Cr), alkaline phosphatase (ALP), blood urea nitrogen
8
9 483 (Bun), Uric Acid (UA), blood glucose (Glu), systolic blood pressure (SBp), diastolic blood
10
11 484 pressure (DBp), blood sedimentation (ESR), hemoglobin (HGB), platelets (PLT),
12
13 485 Apolipoprotein A1 (ApoA1), Apolipoprotein B2 (ApoB), total bilirubin (TB), total protein
14
15 486 (TP). Cross validation selected 16 variables to be potential predictors. The corresponding
16
17 487 forest-plot is shown in (A). The AUC of these above 16 variables for NAFLD is 0.88 (B). Cox
18
19 488 regression confirmed that the 16 variables were relevant for NAFLD incidence, including BMI,
20
21 489 albumin, WBC, TG, HDL, GLT, ALT, Cr, UA, Glu, SBp, DBp, ESR, HGB, PLT and ApoB.
22
23 490 The corresponding forest-plot is shown (C).

24 491
25
26 492 **Figure 3.** Dynamic Bayesian network analyses were performed to show the cause-effect link
27
28 493 between non-alcoholic fatty liver disease (NAFLD) and its potential risk factors. Three
29
30 494 variables, BMI, gender and triglycerides directly pointed to NAFLD. ApoB impacted on the
31
32 495 incidence of NAFLD through TG abundance. LDL indirectly contributed to NAFLD through
33
34 496 ApoB. NAFLD directly led to alterations of seven clinical parameters: ALT, DBp, SBp, TP,
35
36 497 albumin, HGB and UA.

37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

500 **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											
(n)	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

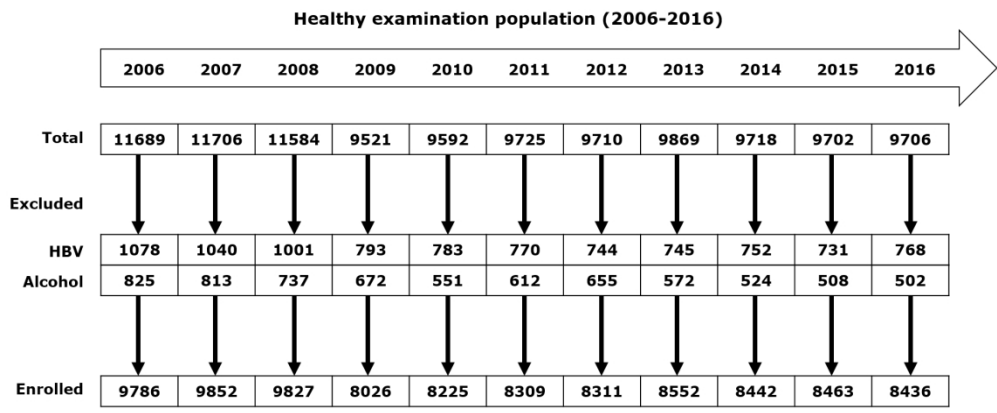
501

502
503**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)	

504
505

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



335x136mm (150 x 150 DPI)

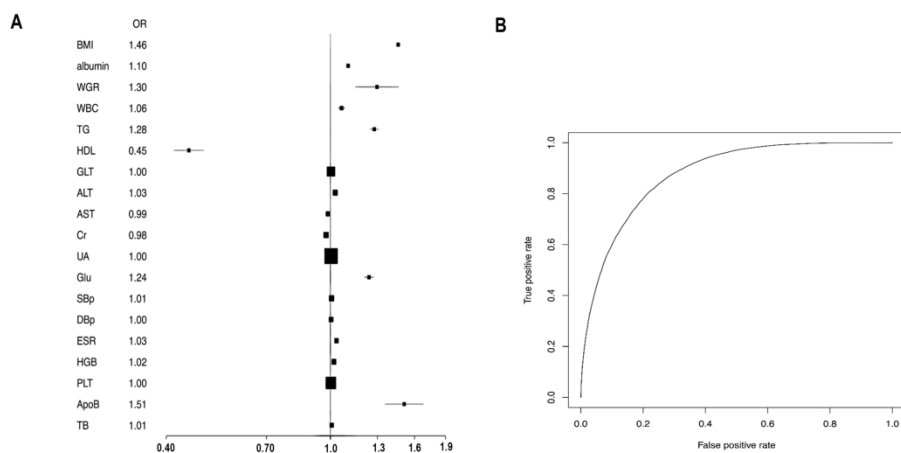
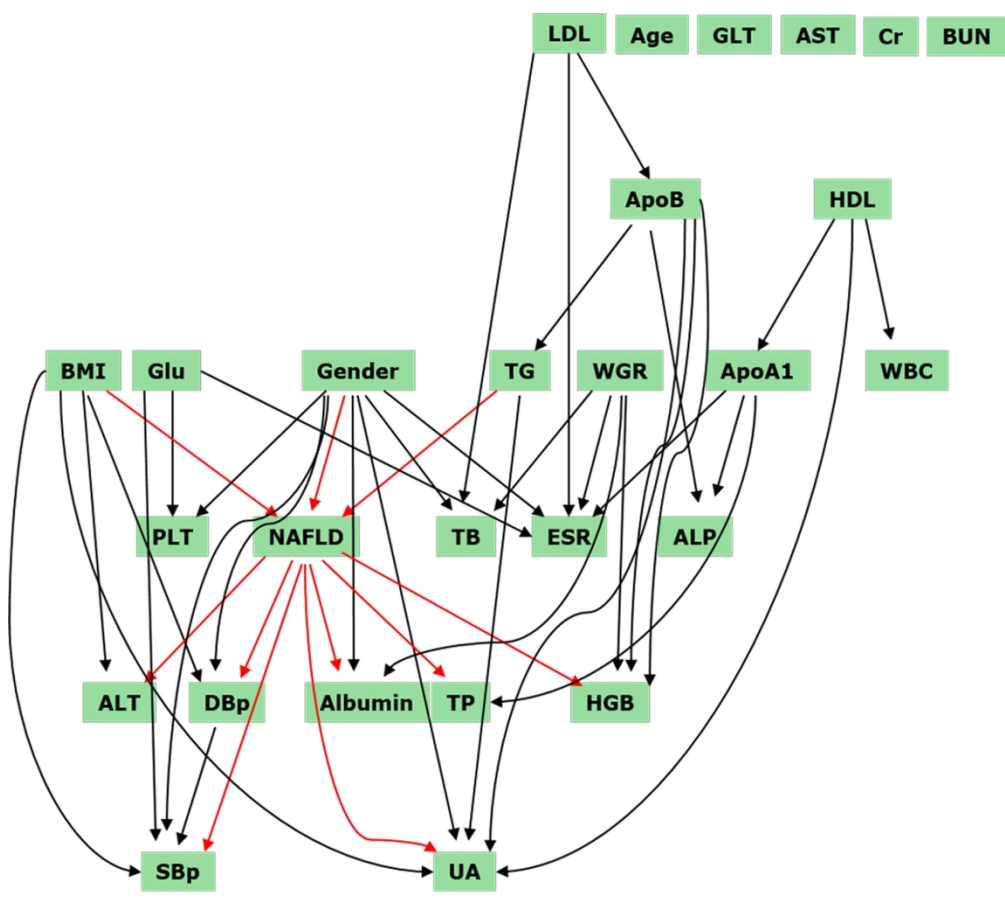


Figure 2

122x63mm (600 x 600 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



522x461mm (96 x 96 DPI)

1
2
3 **Nonalcoholic fatty liver disease increased risk of metabolic diseases, but not severe liver**
4 **disease: an eleven-year follow-up study**
5

6 Xiaoping Tang, Yanyan Shi, Juan Du, Keming Hu, Tingting Zhou, Lan Chen, Yanming
7 Zhang, Fujun Li, Huier Zhang, Roman Liebe, Christoph Meyer, Steven Dooley, Zhongwei
8 Zhu, Hong-Lei Weng, Jinzhu Jia, Tong Huang
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

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)

1	
2	
3	TP (total protein)
4	UA (uric acid)
5	Waist
6	Weight
7	
8	WGR (white globulin ratio)
9	WBC (white blood cell count)
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49	
50	
51	
52	
53	
54	
55	
56	
57	
58	
59	
60	

For peer review only

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 (n)	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 (n)	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 (n)	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	0.9939885	< 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) <i>Cohort study</i> —Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up <i>Case-control study</i> —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 <i>Cross-sectional study</i> —Give the eligibility criteria, and the sources and methods of selection of participants (b) <i>Cohort study</i> —For matched studies, give matching criteria and number of exposed and unexposed <i>Case-control study</i> —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) <i>Cohort study</i> —If applicable, explain how loss to follow-up was addressed <i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed <i>Cross-sectional study</i> —If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses	

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Continued on next page

For peer review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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) <i>Cohort study</i> —Summarise follow-up time (eg, average and total amount)
Outcome data	15 *	<i>Cohort study</i> —Report numbers of outcome events or summary measures over time <i>Case-control study</i> —Report numbers in each exposure category, or summary measures of exposure <i>Cross-sectional study</i> —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:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2021-054891.R2
Article Type:	Original research
Date Submitted by the Author:	07-May-2022
Complete List of Authors:	<p>Tang, Xiaoping ; Ningbo Zhenhai District Lianhua Hospital , Department of Prevention and Health Care</p> <p>Shi, Yanyan; Peking University Third Hospital, Research Center of Clinical Epidemiology</p> <p>Du, Juan ; Ningbo Zhenhai District Lianhua Hospital , Department of Internal Medicine</p> <p>Hu, Keming; Ningbo Zhenhai District Lianhua Hospital, Department of Prevention and Health Care</p> <p>Zhou, Tingting; Ningbo Zhenhai District Lianhua Hospital , Department of Internal Medicine</p> <p>Chen, Lan; Ningbo Zhenhai District Lianhua Hospital , Department of Radiology</p> <p>Zhang, Yanming; Ningbo Zhenhai District Lianhua Hospital , Center Laboratory</p> <p>Li, Fujun; Ningbo Zhenhai District Lianhua Hospital , Department of Radiology</p> <p>Zhang, Huier ; Ningbo Zhenhai District Lianhua Hospital , Center Laboratory</p> <p>Liebe, Roman ; Heinrich Heine University Düsseldorf, Clinic of Gastroenterology, Hepatology and Infectious Diseases,; Saarland University, Department of Medicine II, Saarland University Medical Center</p> <p>Meyer, Christoph ; Heidelberg University, Department of Medicine II, Medical Faculty Mannheim</p> <p>Dooley, Steven ; Heidelberg University, Department of Medicine II, Medical Faculty Mannheim</p> <p>Zhu, Zhongwei; Ningbo Zhenhai District Lianhua Hospital , Department of Surgery</p> <p>Weng, Hong-Lei ; Heidelberg University, Department of Medicine II, Medical Faculty Mannheim</p> <p>JIA, Jinzhu; Peking University, Department of Biostatistics, School of Public Health; Peking University, Center for Statistical Science</p> <p>Huang, Tong; Ningbo City First Hospital; Ningbo Zhenhai District Lianhua Hospital , Department of Prevention and Health Care</p>
Primary Subject Heading:	Gastroenterology and hepatology
Secondary Subject Heading:	Cardiovascular medicine, Diabetes and endocrinology
Keywords:	Epidemiology < INFECTIOUS DISEASES, Hypertension < CARDIOLOGY,

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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

SCHOLARONE™
Manuscripts



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 Hu¹, Tingting Zhou³, Lan Chen⁴, Yanming Zhang⁵, Fujun Li⁴, Huier Zhang⁵, Roman Liebe^{6,7}, Christoph Meyer⁸, Steven Dooley⁸, Zhongwei Zhu⁹, Hong-Lei Weng⁸, 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

1
2
3 35
4
5 36 Jinzhu Jia, PhD.
6
7 37 Department of Biostatistics, School of Public Health, Peking University Health Science
8
9 38 Center
10 39 Address: No. 38, Xueyuan Road, Haidian District, Beijing, 100191, P.R. China
11
12 40 Email: jzjia@math.pku.edu.cn
13
14 41 Tel: +86-15801049187
15
16 42

17 43 **Competing interests:**

18
19 44 Authors do not have conflict of interest.
20
21 45

22 46 **Funding statement:**

23
24
25 47 This project is supported by the National Key R&D Program of China No. 2017YFC0908103
26
27 48 (X.T., and J.D.).
28
29 49

30 50 **Data availability statement**

31
32 51 Data are available upon reasonable request.
33
34 52

35
36 53 **Acknowledgements:**

37
38 54 We are grateful to Drs. Changxi Chen, Jingyi Yuan and Yongjun Chen for support and
39
40 55 discussion.
41
42 56

43 57 **Author contributions:**

44
45 58 Conception and design: Xiaoping Tang, Yanyan Shi, Juan Du, Hong-Lei Weng, Jinzhu Jia
46
47
48 59 and Tong Huang

49
50 60 Ultrasonography: Lan Chen, Fujun Li

51
52 61 Blood assays: Yanming Zhang and Huier Zhang

53
54 62 Other examinations and data collection: Xiaoping Tang, Keming Hu, Tingting Zhou, Juan Du,

55
56
57 63 Zhongwei Zhu and Tong Huang

58
59 64 Statistical analyses: Yanyan Shi and Jinzhu Jia
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

65 Drafting the article: Hong-Lei Weng

66 Reviewing and editing the article critically: Xiaoping Tang, Yanyan Shi, Christoph Meyer,

67 Roman Liebe, Steven Dooley, Hong-Lei Weng, Jinzhu Jia and Tong Huang

68

For peer review only

1
2
3 **69 Abstract**

4 70
5 71 **Objectives** To clarify NAFLD prevalence, risk factors, and clinical outcome in an exemplary
6 72 Chinese population, a cohort of company employees was followed up for eleven years.

7 73 **Design** Retrospective cohort study

8 74 **Setting** Between 2006-2016 in China

9 75 **Participants** 13032 company employees

10 76 **Results** Over eleven years, the prevalence of NAFLD increased from 17.2% to 32.4% (males
11 77 20.5% to 37% *versus* females 9.8% to 22.2%). Male peak prevalence was between 40 and 60
12 78 years of age, whereas highest prevalence in women was at an age of 60 years and older. Logistic
13 79 and Cox regression revealed 16 risk factors, including body mass index, albumin, white blood
14 80 cell, triglycerides, high density lipoprotein, glutamyl transpeptidase, alanine transaminase,
15 81 creatinine, urea acid, glucose, systolic blood pressure, diastolic blood pressure, Blood
16 82 sedimentation, hemoglobin, platelet, and Apolipoprotein B2 ($P < 0.05$ for all factors). The Area
17 83 Under the Curve of these variables for NAFLD is 0.88. However, cause-effect analyses showed
18 84 that only body mass index, gender and triglycerides directly contributed to NAFLD
19 85 development. Over an 11-year follow-up period, 12.6%, 37.7% and 14.2% of male NAFLD
20 86 patients and 11.6%, 44.7% and 22.6% of female NAFLD patients developed diabetes,
21 87 hypertension and hyperuricemia, respectively. Except one male patient who developed
22 88 cirrhosis, no NAFLD patients progressed into severe liver disease.

23 89 **Conclusion** Diabetes, hypertension and hyperuricemia are the main clinical outcomes of
24 90 NAFLD. Eleven years of NAFLD are not sufficient to cause severe liver disease. Age and
25 91 obesity are direct risk factors for NAFLD. BMI, gender and triglycerides are three parameters
26 92 directly reflecting the occurrence of NAFLD.

27 93

28 94 **(Words: 250)**

29 95

30 96 **Strengths and limitations of this study**

- 31 97 ● This study dynamically follows up NAFLD prevalence in an eastern Chinese community
32 98 for eleven years.
- 33 99 ● The study adopted First order Markov models to evaluate the cause-effect link between
34 100 NAFLD and risk factors.
- 35 101 ● The relatively low sensitivity of ultrasound for the detection of liver fat might
36 102 underestimate the incidence of NAFLD in this cohort.

- 1
2
3 103 ● Given that the current study is a single-center observation, multiple-center studies are
4 required to confirm the conclusions in the future.
5 104
6 105 ● The study population is a highly select, relatively homogenous group of well-educated
7 professionals in privileged social positions and permanent employment. Thus, the
8 106 conclusions might not be transferable to the general Chinese population.
9 107
10 108
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

109 **Introduction**

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

119
120 In China, the incidence of NAFLD has been increasing over the last two decades. A recent
121 meta-analysis, based on 392 studies between 2008 and 2018, showed the national incidence of
122 NAFLD in China to be at 29.2% [6]. In Shanghai, the adult incidence of NAFLD has increased
123 from 14.04% in 1995 to 43.65% in 2015 [2]. Being a vast country, Chinese living in different
124 areas vary widely in lifestyle and economic status. Thus, the epidemiology, natural history and
125 clinical outcomes of NAFLD in different areas of the country are worth further investigation.

126
127 It is well accepted that viral hepatitis is a major reason for progressive chronic liver diseases,
128 e.g., fibrosis, cirrhosis and ultimately, hepatocellular carcinoma (HCC). With regard to
129 NAFLD, incidence and severity of associated chronic liver disease outcomes has not been
130 monitored in large Chinese cohorts yet – especially over an extended time span. The current
131 study therefore describes the prevalence of NAFLD in a large Eastern Chinese community over
132 eleven years (2006 - 2016). We focused on three questions: (1) What is the annual incidence of
133 NAFLD? (2) What are the risk factors for NAFLD? And (3) What are the most frequent extra-
134 and intrahepatic clinical outcomes of NAFLD in this selected population?

135 **Methods**

136 *Patient and public involvement*

137
138 No patients were involved in this study

139 *Design and participants*

140
141 In this retrospective study we analyzed the "annual health examination database" of the Zhenhai
142 Lianhua Hospital from 2006 to 2016. This hospital is affiliated to Sinopec Zhenhai Refining &

1
2
3 144 Chemical Company. Supported by the company, all employees were offered the opportunity to
4
5 145 go to this hospital for an annual health examination. Over a period of 11 years, a total 13,032
6
7 146 employees received health examinations. From 2006 to 2016, 11689, 11706, 11584, 9521,
8
9 147 9592, 9725, 9710, 9869, 9718, 9702 and 9706 persons received health examinations (**Figure**
10
11 148 **1**). To describe the longitudinal NAFLD occurrence in this cohort, we excluded subjects with
12
13 149 the following conditions: (1) viral hepatitis B and C infection, which were identified by blood
14
15 150 virus measurements (HBV-DNA and HCV-RNA), and (2) alcoholic liver disease, which was
16
17 151 defined as previously described [7,8]. NAFLD was defined as the presence of hepatic steatosis,
18
19 152 determined by ultrasonography.
20
21 153 The study protocol was approved by the Ethics Committee of Zhenhai Lianhua Hospital
22
23 154 ([2016]001). Informed consent was obtained from all subjects.

24 155

25 156 *Measures*

26 157 **Supplementary Table 1** shows all parameters measured in the annual health examinations.
27
28 158 Ultrasonography was performed by the same three experienced doctors (L.C., F. L., and J.Y.)
29
30 159 with an Ultrasonograph B, GE, Voluson 730 pro. Blood biochemistry and HBV serum levels
31
32 160 were measured by an Olympus AU640 autoanalyzer (Olympus, Kobe, Japan) and an
33
34 161 ImmunoAssay Analyzer VitrosECI (JOHNSON-JOHNSON, USA), respectively. All methods
35
36 162 were carried out in accordance with relevant guidelines and regulations.

37 163

38 164 *Statistical analysis*

39
40 165 For population characteristics, variables were described as means and standard deviation (SD)
41
42 166 or proportions as appropriate. Student's t-test or nonparametric test was used to analyze
43
44 167 differences between two groups as mentioned. Chi-square test was used to verify the differences
45
46 168 of nominative variables between two groups. Multivariate analysis of risk factors for NAFLD
47
48 169 was performed using logistic regression analysis. Combined receiver operating characteristic
49
50 170 curve (ROC) and area under curve (AUC) analyses were used to evaluate the diagnostic
51
52 171 performance of biomarkers based on the logistic regression model. Multivariate Cox regression
53
54 172 model was performed to calculate hazard ratios of variables to identify independent prognostic
55
56 173 variables. First order Markov models were used to analyze the cause-effect link between
57
58 174 NAFLD and risk factors. L1 penalized logistic regression was applied to select predictive
59
60 175 predictors. R package "glmnet" contains functions to select predictors using L1 penalized
176 176 logistic regression. Statistical analyses were performed using SPSS 22.0 and R version 3.5.3.

1
2
3 177 *P*-values that were less than 0.05 were considered statistically significant. Figures were
4 generated by R package such as ‘forestplot’, ‘ROCR’, ‘bnlearn’, or ‘survival’.

5 178
6 179

7 180

8 181 **Results**

9 182

10 183 *Prevalence of NAFLD from 2006 to 2016*

11 184 We retrospectively analyzed 9786, 9852, 9827, 8026, 8225, 8309, 8311, 8552, 8442, 8463 and
12 185 8436 persons who received health examinations from 2006 to 2016. **Supplementary Table 2**
13 186 shows the eleven-year annual NAFLD incidence in this population. In 2006, NAFLD was
14 187 diagnosed in 17.2% of persons, and gradually increased over the examination period to 19%
15 188 (2007), 22% (2008), 22.4% (2009), 22.7% (2010), 23.4% (2011), 24.5% (2012), 25.4% (2013),
16 189 27.9% (2014), 30.8% (2015), and 32.4% (2016), respectively (**Supplementary Table 2**). Both
17 190 males and females demonstrated continuously increasing NAFLD prevalence (**Supplementary**
18 191 **Table 2**). Compared to females, male Chinese demonstrated significantly higher NAFLD
19 192 prevalence, e.g., in 2006, the prevalence of NAFLD in males and females was 20.5% and 9.8%,
20 193 respectively. Eleven years later, the prevalence had increased to 37% in males and 22.2% in
21 194 females (**Supplementary Table 2**). Noteworthy, the prevalence of NAFLD in male and female
22 195 was correlating with age. The peak prevalence of NAFLD in men emerged in those aged
23 196 between 40 and 60 years. In 2006, the prevalence of NAFLD in men aged between 40 - 50 and
24 197 50 - 60 years was 24.2% and 24.6%, respectively. In 2016, prevalence reached 42.8% and
25 198 46.6% (**Supplementary Table 2**) for men. Distinct from males, the peak NAFLD prevalence
26 199 in females emerged at an age above 60 years. In 2006, the prevalence of NAFLD in women
27 200 older than 60 and 70 years was 25.6% and 22.9%, respectively. Eleven years later, these values
28 201 had increased to 53.4% and 30.9% (**Supplementary Table 2**).

29 202

30 203 Among the observed population, 5606 persons received annual health examinations for 11
31 204 years, and thus prevalence of NAFLD was analyzed in these individuals. As shown in **Table 1**,
32 205 the prevalence of NAFLD increased from 17% in 2006 to 35.2% in 2016. The highest
33 206 prevalence rates for NAFLD in 3795 men emerged in those older than 40 years. In 2006, the
34 207 NAFLD prevalence in males aged between 40 - 50, 50 - 60, and 60 - 70 years was 24.2%,
35 208 25.4% and 19.8%, respectively. In 2016, these values reached 43.3%, 46.8% and 42.3% (**Table**
36 209 **1**). Different from males, the peak NAFLD prevalence in 1811 females emerged at an age of
37 210 more than 60 years. In 2006, the prevalence of NAFLD in women older than 60 or 70 years

211 was 22.1% and 28%, respectively. Eleven years later, these values had increased to 31.8% and
212 28.6% (**Table 1**).

213

214 *BMI and NAFLD incidence*

215 Given the tight link between obesity and NAFLD, we paid special attention to the population
216 with high BMI. We focused on the 5606 persons with complete follow-up and analyzed the
217 prevalence of NAFLD in those with BMI >25. In total, out of the 5606 persons, 2445 presented
218 with a BMI of >25. The prevalence of NAFLD in this overweight subpopulation was far higher
219 than in the general population. In 2006, 45.2% of individuals (n=1104; male vs. female: 47.3%
220 vs. 37.1%) with BMI >25 were suffering from NAFLD (**Supplementary Table 3**). In 2016,
221 values reached 67.1% (n=1414; male vs. female: 69% vs. 59.2%, **Supplementary Table 3**).
222 Impressively, the NAFLD prevalence in both genders was very high at any age, even in those
223 below the age of 30 years. In 2006, among 213 overweight men, younger than 30 years, 52.6%
224 were also diagnosed for NAFLD (**Supplementary Table 3**). This number increased to 63% in
225 2016 (**Supplementary Table 3**). In 2006, there were 15 overweight women aged less than 30
226 years. Among them, 3 presented as NAFLD (20%). In 2016, 7 out of 16 overweight women
227 aged less than 30 years were identified. The NAFLD prevalence had increased to 43.8%
228 (**Supplementary Table 3**). In those older than 40 years, NAFLD prevalence increased from
229 36.6 – 45.4% in 2006 to 53 – 65.6% in 2016 (**Supplementary Table 3**).

230

231 *Risk factors relevant to NAFLD occurrence*

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

1
2
3 245 strong negative correlation with NAFLD incidence (**Supplementary Table 4**). The AUC of
4
5 246 these variables for NAFLD is 0.88 (see ROC curve in **Figure 2B**). We further performed a time
6
7 247 dependent Cox regression to calculate the hazard ratios of these parameters for NAFLD
8
9 248 occurrence. Cox regression confirmed that the 16 parameters were significantly relevant to
10
11 249 NAFLD incidence (**Supplementary Table 5** and **Figure 2C**). Furthermore, ApoB and HDL
12
13 250 were the most robust positive and negative risk factors for NAFLD (**Figure 2C**).
14
15 251

15 252 *Cause-effect link between risk factors and NAFLD occurrence*

16
17 253 Although the aforementioned parameters were regarded as “risk factors” according to statistical
18
19 254 models, it did not necessarily mean that all of them contributed to NAFLD occurrence. Based
20
21 255 on 11 years of longitudinal data, it was possible to construct a dynamic Bayesian network to
22
23 256 identify the risk factors most relevant to NAFLD occurrence. As shown in **Figure 3**, these
24
25 257 parameters constituted a complicated, but clear intercross paradigm. Only three parameters,
26
27 258 BMI, gender and TG directly pointed to NAFLD. In addition, ApoB impacted the incidence of
28
29 259 NAFLD through contributing to TG. Furthermore, LDL can indirectly contribute to NAFLD
30
31 260 through influencing ApoB. Very impressively, the dynamic Bayesian network pointed out that
32
33 261 NAFLD directly leads to alterations of seven parameters: ALT, DBp, SBp, TP, albumin, HGB
34
35 262 and UA. Intriguingly, age, GLT, AST, Cr, and BUN did not interact with any other parameter
36
37 263 in our model, indicating that these factors correlate by incidence, but there is no causal
38
39 264 interaction.

39 266 *Outcome of NAFLD*

40
41 267 Subsequently, we examined clinical outcomes of NAFLD over the eleven years. **Table 2**
42
43 268 summarizes the incidence of intra- and extrahepatic diseases of 696 NAFLD and 222 NASH
44
45 269 patients during the follow-up period. Among the total NAFLD and NASH population, only 1
46
47 270 male NAFLD patient developed liver cirrhosis within the eleven years. However, this time span
48
49 271 witnessed significantly increased extrahepatic diseases, including diabetes, hypertension and
50
51 272 hyperuricemia. In the NAFLD population, there were 64 (12.6%) men and 22 (11.6%) women,
52
53 273 191 (37.7%) men and 85 (44.7%) women, 72 (14.2%) men and 43 (22.6%) women who
54
55 274 developed into type II diabetes, hypertension and hyperuricemia, respectively (**Table 2**).
56
57 275 Given that 2006 is the starting point of data collection, patients diagnosed for NAFLD in this
58
59 276 year had by no means just manifested their disease, but rather patients had possibly developed
60
277 NAFLD several years prior to inclusion. To clarify the exact clinical outcomes of NAFLD over
278
one decade, we focused on the following two cohorts of individuals with annual health

1
2
3 279 examinations for 11 years: (1) patients who were diagnosed as non-NAFLD in 2006, but were
4 280 NAFLD in 2007 (new NAFLD cohort); and (2) who were non-NAFLD in both 2006 and 2007
5 281 (non-NAFLD cohort). As shown in **Table 2**, 185 new NAFLD cases (138 men and 47 women)
6 282 and 4547 non-NAFLD (2786 men and 1761 women) persons were found in 2007. Between
7 283 2007 and 2016, neither NAFLD nor non-NAFLD individuals developed liver cirrhosis or
8 284 cancer. However, the one-decade follow-up reveals different prevalences of diabetes,
9 285 hypertension and hyperuricemia: In NAFLD patients, there were 14 (10.1%) men and 5 (10.6%)
10 286 women, 47 (34.1%) men and 21 (44.7%) women, 34 (24.6%) men and 8 (17%) women who
11 287 developed type II diabetes, hypertension and hyperuricemia, respectively (**Table 2**). In non-
12 288 NAFLD individuals, 157 (5.6%) men and 54 (3.4%) women, 259 (9.3%) men and 324 (18.4%),
13 289 284 (10.2%) men and 84 women (4.8%) developed type II diabetes, hypertension and
14 290 hyperuricemia, respectively (**Table 2**). For all three diseases, statistically significant differences
15 291 were determined between the two cohorts of population (all $P < 0.05$, **Table 2**). These results
16 292 suggest that diabetes, hypertension and hyperuricemia are the main clinical outcomes of
17 293 NAFLD.
18 294
19 295

296 Discussion

297 This eleven-year follow-up retrospective study reports the following: (1) NAFLD prevalence
298 has substantially increased in the examined Eastern Chinese population. (2) The prevalence of
299 NAFLD differs by gender and age. Middle-aged men and elderly women are the two
300 populations at highest risk for NAFLD. (3) Gender, BMI and triglycerides are the parameters
301 directly associated with NAFLD occurrence. Regardless of gender and age, persons with high
302 BMI (≥ 25) have a high risk for NAFLD development. (4) NAFLD directly leads to alterations
303 of seven clinical parameters: ALT, DBp, SBp, TP, albumin, HGB and UA. (5) Within 11 years,
304 a significant part of the NAFLD population develops three clinically relevant diseases: type 2
305 diabetes mellitus, hypertension and hyperuricemia. (6) Within 11 years, NAFLD does not cause
306 severe liver disease, such as cirrhosis or HCC, in patients.
307

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

1
2
3 314 a study with biopsy-proven NAFLD revealed 1.97%-2.97% cirrhosis prevalence [11]. In
4
5 315 addition, NAFLD is regarded as the third-most common cause of cancer-related death
6
7 316 worldwide [12]. In a study based on 4949 US patients with HCC, 701 patients had NAFLD [13].
8
9 317 It was estimated that the cumulative incidence of HCC among patients with NAFLD and
10
11 318 cirrhosis ranges from 2.4% to 12.8% over a median follow-up period of 3.2 to 7.2 years [14]
12
13 319 (Global Health Observatory) data. Mortality and global health estimates were obtained from:
14
15 320 http://www.who.int/gho/mortality_burden_disease/en/. Last accessed on 1/7/2020.). Given that
16
17 321 the above conclusions were based on cross-sectional investigations and statistical models, it has
18
19 322 been unknown to date over which period a NAFLD patient develops liver cirrhosis or HCC
20
21 323 (personal risk assessment). Our 11-year follow-up provides therefore a valuable and
22
23 324 comprehensive dataset. In this study, most patients were diagnosed with NAFLD when they
24
25 325 received a routine health examination. Before the examination, these people did not have any
26
27 326 symptoms or signs of NAFLD. Therefore, they belong to NAFLD patients at a very early stage
28
29 327 (although for 2006, the duration of pre-existing NAFLD cannot be determined). Except for a
30
31 328 single person, no serious liver problems were observed within this time period. These data
32
33 329 suggest that for the vast majority of patients with early stage NAFLD, 11 years are not sufficient
34
35 330 to develop liver cirrhosis or cancer. Nasr et al followed up 129 NAFLD patients with varying
36
37 331 fibrosis stages on two occasions (mean time 13.7 and 9.3 years). Liver biopsy analyses showed
38
39 332 that 9.3% of patients developed end-stage liver disease and 34% advanced fibrosis [15]. The
40
41 333 NAFLD patients observed by Nasr et al. actually belonged to the NASH category, because they
42
43 334 suffered from fibrosis and elevated ALT and/or AST levels. As our study was based on
44
45 335 examinations of healthy, liver biopsy is not justifiable. Very likely, the current cohort included
46
47 336 a portion of NASH patients. They also did not show significant progression towards cirrhosis
48
49 337 or HCC was monitored.

50
51 338
52
53 339 In contrast to hepatic complications, patients with NAFLD showed a significant risk for the
54
55 340 development of extrahepatic diseases, including diabetes, hypertension and hyperuricemia. In
56
57 341 696 NAFLD patients, eleven years witnessed the development of 86 cases (12.4%) type II
58
59 342 diabetes, 276 (40%) cases of hypertension and 115 patients with (16.5%) hyperuricemia,
60
343 respectively. Interestingly, in 222 NASH patients, the prevalence of these three diseases was
344 12 (5.4%), 46 (20.7%), and 33 (14.9%) only. In general, men had a higher probability to develop
345 these diseases than women. These results are consistent with previous reports from USA and
346 Europe [16-18]. Whether NAFLD is associated with the risk of severe heart or brain diseases
347 such as acute myocardial infarction (AMI) and stroke is worth further investigation. A recent

1
2
3 348 matched cohort study analyzed databases from four European countries, which included 17.7
4
5 349 million patients with NAFLD or NASH [19]. These patients had a mean follow-up of 2.1 to 5.5
6
7 350 years. The study showed that the diagnosis of NAFLD appears not to be associated with AMI
8
9 351 or stroke risk after adjustment for established cardiovascular risk factors. Nevertheless, the
10 352 authors mentioned that cardiovascular risk assessment in adults with a diagnosis of NAFLD is
11
12 353 important [19]. Follow-up for 5 years might be not sufficient to reach a conclusion for this issue.
13
14 354

15 355 An important issue is the cause-and-effect relationship between NAFLD and its clinical
16
17 356 outcomes such as diabetes, hypertension and hyperuricemia. A dynamic Bayesian network in
18
19 357 the current study provides direct evidence on this issue: NAFLD directly results in alterations
20
21 358 of several parameters, including DBp, SBp and UA, suggesting that NAFLD directly
22
23 359 contributes to the occurrence of hypertension and hyperuricemia. The underlying mechanisms
24
25 360 require further investigation.

26 361 The current dynamic Bayesian network analysis does not confirm a direct cause-and-
27
28 362 effect relationship between NAFLD and type 2 diabetes mellitus. There are plenty of studies
29
30 363 showing the close relationship between type 2 diabetes and NAFLD [20]. Pathophysiologically,
31
32 364 insulin resistance is a key event in both NAFLD and diabetes progression [20]. However,
33
34 365 genome-wide association studies have not yet identified the exact impact of insulin resistance
35
36 366 on the variants associated with NAFLD severity [20,21]. Clarification of the cause-and-effect
37
38 367 relationship between NAFLD and diabetes requires further long-term follow-up studies.
39
40 368

41 369 To date, there are a large number of studies investigating risk factors for NAFLD [22]. These
42
43 370 studies tried to identify single, or multiple combined biomarkers to predict NAFLD occurrence.
44
45 371 Given that most studies were based on cross-sectional designs, or with only short follow-up
46
47 372 periods, it is difficult to clarify the causality between the proposed predictors and NAFLD
48
49 373 morbidity. Our eleven-year dataset provides a chance to shed led on this issue. Here, the
50
51 374 dynamic causal relationships between variables, including risk parameters and clinical
52
53 375 outcomes, were identified by a first order Markov model, which was displayed by a dynamic
54
55 376 Bayes network. The dynamic Bayes model discriminates causal relationship through time
56
57 377 sequence. When a variable change is closely related to a previous variance alteration, a causal
58
59 378 relationship between the two variables is assumed. Based on Logistic and Cox regression and
60
379 dynamic Bayesian network analyses, we confirmed three direct risk factors for NAFLD
380 occurrence: gender, BMI and TG. These findings are supported by the following data: (1) Men
381 have higher NAFLD prevalence than women in this population (37% *versus* 22.2% in 2016);

1
2
3 382 (2) In overweight people with a BMI >25, NAFLD prevalence reached 69% in males and 59.2%
4 383 in females. Given that triglycerides are a major energy source, but are leading to obesity, it is
5 384 not surprising that this parameter directly reflects the risk for NAFLD development. These
6 385 findings provide robust evidence supporting the use of BMI to monitor or predict NAFLD.
7
8
9

10 386

11 387

12 388 **Conclusion**

13 389 This 11-year follow-up study documents the rapid increase in NAFLD prevalence in an Eastern
14 390 Chinese population. In contrast to previous reports, our observation does not observe that one
15 391 decade of NAFLD is sufficient to lead to severe hepatic clinical outcomes. It is worthy to note
16 392 that our population were biased because they are on the well-off, well-educated side of the
17 393 Chinese people, while previous studies were often based on hospital populations, who suffered
18 394 from negative selection bias and thus came up with higher estimates. In addition, given there
19 395 are differences in NAFLD profiles between Eastern and Western populations, it would be
20 396 interesting to know the natural development of NAFLD in a Western population. A key point
21 397 for clarifying the true history of NAFLD is to follow a population starting from the early phases
22 398 of the disease. Consistent with previous studies, NAFLD is tightly associated with multiple
23 399 extra-hepatic diseases relevant to the metabolic syndrome. In the future, follow-up of the
24 400 current cohort for another one and two decades will provide further valuable data to clarify the
25 401 extended natural history of NAFLD. Last but not least, a large portion of the men and women
26 402 in this study were educated above the average and have a position in the company that gave
27 403 them the availability of better food choices as well as regular sport. On the other hand, the
28 404 relatively low sensitivity of ultrasound for the detection of liver fat might underestimate the
29 405 incidence of NAFLD in this cohort.
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

48 406

49 407

50 408 **References**

- 51 409 [1] Younossi ZM, Tampi R, Priyadarshini M, et al. Burden of Illness and Economic Model
52 410 for Patients With Nonalcoholic Steatohepatitis in the United States. *Hepatology*, 2019,
53 411 69: 564-572. <http://doi.org/10.1002/hep.30254>
- 54 412 [2] Younossi ZM, Loomba R, Anstee QM, et al. Diagnostic modalities for nonalcoholic
55 413 fatty liver disease, nonalcoholic steatohepatitis, and associated fibrosis. *Hepatology*,
56 414 2018, 68: 349-360. <http://doi.org/10.1002/hep.29721>
- 57 415 [3] Fan JG, Kim SU, Wong VW. New trends on obesity and NAFLD in Asia. *J Hepatol*,
58 416 2017, 67: 862-873. <http://doi.org/10.1016/j.jhep.2017.06.003>

- 1
2
3
4 417 [4] Byrne CD, Targher G. NAFLD: a multisystem disease. *J Hepatol*, 2015, 62: S47-64.
5 418 <http://doi.org/10.1016/j.jhep.2014.12.012>
- 6 419 [5] Alexander M, Loomis AK, van der Lei J, et al. Risks and clinical predictors of cirrhosis
7 420 and hepatocellular carcinoma diagnoses in adults with diagnosed NAFLD: real-world
8 421 study of 18 million patients in four European cohorts. *BMC Med*, 2019, 17: 95.
9 422 <http://doi.org/10.1186/s12916-019-1321-x>
- 10 423 [6] Zhou F, Zhou J, Wang W, et al. Unexpected Rapid Increase in the Burden of NAFLD
11 424 in China From 2008 to 2018: A Systematic Review and Meta-Analysis. *Hepatology*,
12 425 2019, 70: 1119-1133. <http://doi.org/10.1002/hep.30702>
- 13 426 [7] EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic
14 427 fatty liver disease. *J Hepatol*, 2016, 64: 1388-1402.
15 428 <http://doi.org/10.1016/j.jhep.2015.11.004>
- 16 429 [8] Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of
17 430 nonalcoholic fatty liver disease: Practice guidance from the American Association for
18 431 the Study of Liver Diseases. *Hepatology*, 2018, 67: 328-357.
19 432 <http://doi.org/10.1002/hep.29367>
- 20 433 [9] Younossi Z, Tacke F, Arrese M, et al. Global Perspectives on Nonalcoholic Fatty Liver
21 434 Disease and Nonalcoholic Steatohepatitis. *Hepatology*, 2019, 69: 2672-2682.
22 435 <http://doi.org/10.1002/hep.30251>
- 23 436 [10] Younossi ZM, Koenig AB, Abdelatif D, et al. Global epidemiology of nonalcoholic
24 437 fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes.
25 438 *Hepatology*, 2016, 64: 73-84. <http://doi.org/10.1002/hep.28431>
- 26 439 [11] Shen F, Zheng RD, Shi JP, et al. Impact of skin capsular distance on the performance
27 440 of controlled attenuation parameter in patients with chronic liver disease. *Liver Int*,
28 441 2015, 35: 2392-2400. <http://doi.org/10.1111/liv.12809>
- 29 442 [12] Ascha MS, Hanouneh IA, Lopez R, et al. The incidence and risk factors of
30 443 hepatocellular carcinoma in patients with nonalcoholic steatohepatitis. *Hepatology*,
31 444 2010, 51: 1972-1978. <http://doi.org/10.1002/hep.23527>
- 32 445 [13] Younossi ZM, Otgonsuren M, Henry L, et al. Association of nonalcoholic fatty liver
33 446 disease (NAFLD) with hepatocellular carcinoma (HCC) in the United States from 2004
34 447 to 2009. *Hepatology*, 2015, 62: 1723-1730. <http://doi.org/10.1002/hep.28123>
- 35 448 [14] White DL, Kanwal F, El-Serag HB. Association between nonalcoholic fatty liver
36 449 disease and risk for hepatocellular cancer, based on systematic review. *Clin*
37 450 *Gastroenterol Hepatol*, 2012, 10: 1342-1359.e1342.
38 451 <http://doi.org/10.1016/j.cgh.2012.10.001>
- 39 452 [15] Nasr P, Ignatova S, Kechagias S, et al. Natural history of nonalcoholic fatty liver
40 453 disease: A prospective follow-up study with serial biopsies. *Hepatol Commun*, 2018, 2:
41 454 199-210. <http://doi.org/10.1002/hep4.1134>
- 42 455 [16] Estes C, Razavi H, Loomba R, et al. Modeling the epidemic of nonalcoholic fatty liver
43 456 disease demonstrates an exponential increase in burden of disease. *Hepatology*, 2018,
44 457 67: 123-133. <http://doi.org/10.1002/hep.29466>
- 45 458 [17] Williamson RM, Price JF, Glancy S, et al. Prevalence of and risk factors for hepatic
46 459 steatosis and nonalcoholic Fatty liver disease in people with type 2 diabetes: the
47 460

- 1
2
3 460 Edinburgh Type 2 Diabetes Study. *Diabetes Care*, 2011, 34: 1139-1144.
4 461 <http://doi.org/10.2337/dc10-2229>
- 5
6 462 [18] Targher G, Bertolini L, Padovani R, et al. Prevalence of nonalcoholic fatty liver disease
7 463 and its association with cardiovascular disease among type 2 diabetic patients. *Diabetes*
8 464 *Care*, 2007, 30: 1212-1218. <http://doi.org/10.2337/dc06-2247>
- 9
10 465 [19] Alexander M, Loomis AK, van der Lei J, et al. Non-alcoholic fatty liver disease and risk
11 466 of incident acute myocardial infarction and stroke: findings from matched cohort study
12 467 of 18 million European adults. *Bmj*, 2019, 367: l5367. <http://doi.org/10.1136/bmj.l5367>
- 13
14 468 [20] Tilg H, Moschen AR, Roden M. NAFLD and diabetes mellitus. *Nat Rev Gastroenterol*
15 469 *Hepatol*, 2017, 14: 32-42. <http://doi.org/10.1038/nrgastro.2016.147>
- 16
17 470 [21] Zhou Y, Llauradó G, Orešič M, et al. Circulating triacylglycerol signatures and insulin
18 471 sensitivity in NAFLD associated with the E167K variant in TM6SF2. *J Hepatol*, 2015,
19 472 62: 657-663. <http://doi.org/10.1016/j.jhep.2014.10.010>
- 20
21 473 [22] Miyake T, Kumagi T, Furukawa S, et al. Non-alcoholic fatty liver disease: factors
22 474 associated with its presence and onset. *J Gastroenterol Hepatol*, 2013, 28 Suppl 4: 71-
23 475 78. <http://doi.org/10.1111/jgh.12251>
24 476
25
26 477

27 28 478 **Figure legends**

29 479 **Figure 1.** Flow chart depicting the enrollment of a population with non-alcoholic fatty liver
30 480 disease (NAFLD) for follow-up in Ningbo Zhenhai Lianhua Hospital, China.
31 481

32
33 482 **Figure 2.** Penalized logistic regression and Cox regression analysis were performed for risk
34 483 factors and hazard ratios of non-alcoholic fatty liver disease (NAFLD). The following
35 484 parameters were available from 5606 participants: Gender, age, BMI, albumin, white globulin
36 485 ratio (WGR), white blood cell (WBC), low density lipoprotein (LDL), triglycerides (TG), high
37 486 density lipoprotein (HDL), glutamyl transpeptidase (GLT), alanine transaminase (ALT),
38 487 aspartate transaminase (AST), creatinine (Cr), alkaline phosphatase (ALP), blood urea nitrogen
39 488 (Bun), Uric Acid (UA), blood glucose (Glu), systolic blood pressure (SBp), diastolic blood
40 489 pressure (DBp), blood sedimentation (ESR), hemoglobin (HGB), platelets (PLT),
41 490 Apolipoprotein A1 (ApoA1), Apolipoprotein B2 (ApoB), total bilirubin (TB), total protein
42 491 (TP). Cross validation selected 16 variables to be potential predictors. The corresponding
43 492 forest-plot is shown in (A). The AUC of these above 16 variables for NAFLD is 0.88 (B). Cox
44 493 regression confirmed that the 16 variables were relevant for NAFLD incidence, including BMI,
45 494 albumin, WBC, TG, HDL, GLT, ALT, Cr, UA, Glu, SBp, DBp, ESR, HGB, PLT and ApoB.
46 495 The corresponding forest-plot is shown (C).
47 496
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 497 **Figure 3.** Dynamic Bayesian network analyses were performed to show the cause-effect link
4
5 498 between non-alcoholic fatty liver disease (NAFLD) and its potential risk factors. Three
6
7 499 variables, BMI, gender and triglycerides directly pointed to NAFLD. ApoB impacted on the
8
9 500 incidence of NAFLD through TG abundance. LDL indirectly contributed to NAFLD through
10
11 501 ApoB. NAFLD directly led to alterations of seven clinical parameters: ALT, DBp, SBp, TP,
12
13 502 albumin, HGB and UA.
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

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

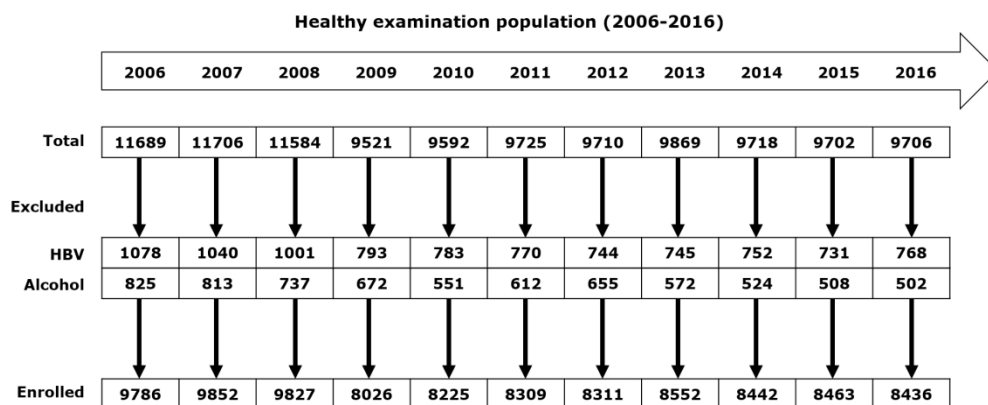
	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Total (n)	5606	5606	5606	5606	5606	5606	5606	5606	5606	5606	5606
NAFLD											
(n)	951	1068	1247	1322	1378	1466	1524	1574	1700	1883	1976
(%)	17	19.1	22.2	23.6	24.6	26.2	27.2	28.1	30.3	33.6	35.2
Male (n)	3795	3795	3795	3795	3795	3795	3795	3795	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											
(n)	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

505

506 **Table 2.** Clinical outcome of NAFLD patients
2006 - 2016

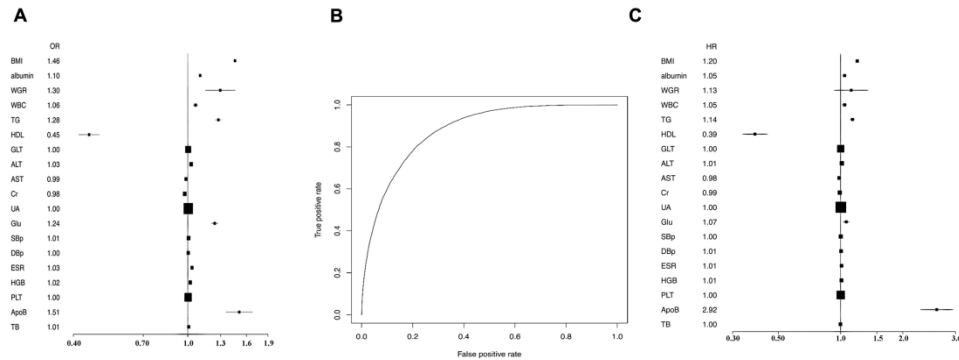
		Cirrhosis	HCC	Diabetes n (%)		Hypertension n (%)		Hyperuricemia n (%)	
Male (n=506)		0	0	64 (12.6)		191 (37.7)		72 (14.2)	
Female (n=190)		0	0	22 (11.6)		85 (44.7)		43 (22.6)	
2007 – 2016 (outcome of new NFLAD)									
		Cirrhosis	HCC	Diabetes		Hypertension		Hyperuricemia	
				n (%)	P-value	n (%)	P-value	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)	

507



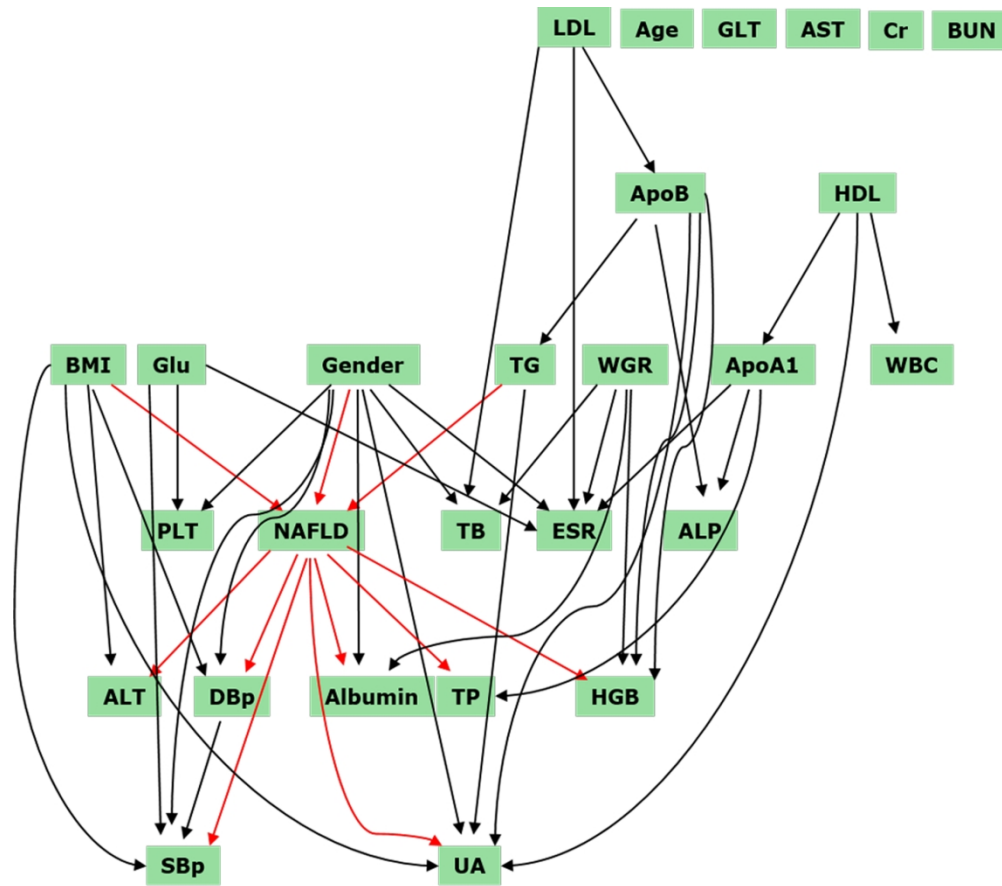
Flow chart depicting the enrollment of a population with non-alcoholic fatty liver disease (NAFLD) for follow-up in Ningbo Zhenhai Lianhua Hospital, China.

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 triglycerides 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)

1
2
3 **Nonalcoholic fatty liver disease increased risk of metabolic diseases, but not severe liver**
4 **disease: an eleven-year follow-up study**

5
6 Xiaoping Tang, Yanyan Shi, Juan Du, Keming Hu, Tingting Zhou, Lan Chen, Yanming
7 Zhang, Fujun Li, Huier Zhang, Roman Liebe, Christoph Meyer, Steven Dooley, Zhongwei
8
9 Zhu, Hong-Lei Weng, Jinzhu Jia, Tong Huang
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

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)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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

For peer review only

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 (n)	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 (n)	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 (n)	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	0.9939885	< 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) <i>Cohort study</i> —Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up <i>Case-control study</i> —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 <i>Cross-sectional study</i> —Give the eligibility criteria, and the sources and methods of selection of participants	5-6
		(b) <i>Cohort study</i> —For matched studies, give matching criteria and number of exposed and unexposed <i>Case-control study</i> —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) <i>Cohort study</i> —If applicable, explain how loss to follow-up was addressed <i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed <i>Cross-sectional study</i> —If applicable, describe analytical methods taking account of sampling strategy	6
		(e) Describe any sensitivity analyses	6

Continued on next page

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60**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) <i>Cohort study</i> —Summarise follow-up time (eg, average and total amount)	
Outcome data	15*	<i>Cohort study</i> —Report numbers of outcome events or summary measures over time <i>Case-control study</i> —Report numbers in each exposure category, or summary measures of exposure <i>Cross-sectional study</i> —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	7-10 7-10 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.