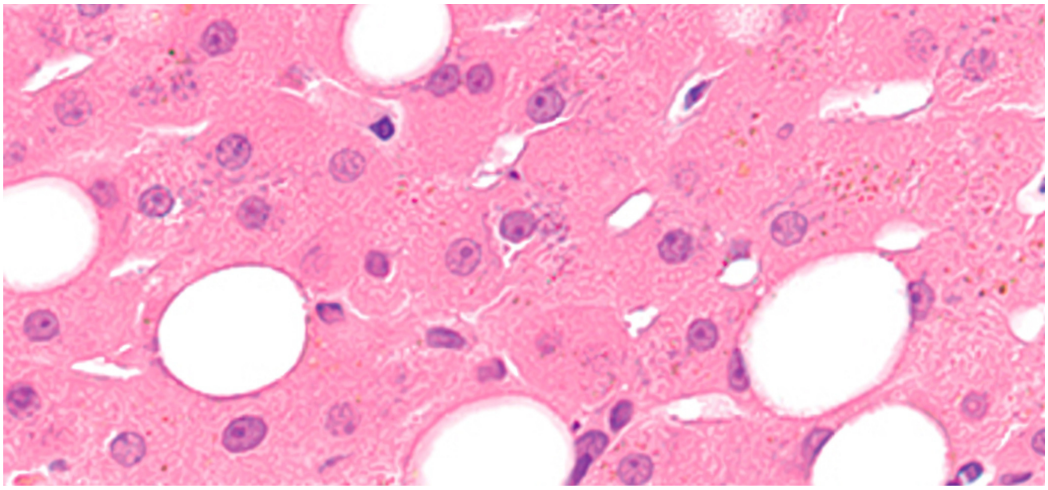

External validation of biomarkers of fatty liver in the general population: the Bagnacavallo study

Francesco Giuseppe Foschi[1], Fabio Conti[1], Marco Domenicali[2], Pierluigi Giacomoni[1], Alberto Borghi[2], Vittoria Bevilacqua[1], Lucia Napoli[1], Dante Berardinelli[1], Mattia Altini[1], Alessandro Cucchetti[2], Giorgio Ercolani[2], Andrea Casadei-Gardini[3], Stefano Bellentani[4], Amalia Gastaldelli[5], Claudio Tiribelli[4], Giorgio Bedogni[4]

1 Azienda Unità Sanitaria Locale della Romagna

2 University of Bologna

3 Università degli Studi di Modena e Reggio Emilia

4 Italian Liver Foundation

5 Italian National Research Council

Funding: The present analysis was sponsored by a research grant from Gilead (Milan, Italy) to the Italian Liver Foundation. The Sponsor had no role in the design and conduct of the study; in the collection, analysis, and interpretation of the data; or in the preparation and review of the manuscript.

Potential competing interests: Giorgio BEDOGNI is a co-founder of Qeios.

Abstract

Objective: We externally validated the fatty liver index (FLI), the lipid accumulation product (LAP), the hepatic steatosis index (HSI), and the Zhejiang University index (ZJU) for the diagnosis of fatty liver (FL) and non-alcoholic fatty liver disease (NAFLD) in the general population.

Subjects and Methods: The validation was performed on 2159 citizens of the town of Bagnacavallo (Ravenna, Italy). Calibration was evaluated by calculating the calibration slope and intercept and by inspecting calibration plots; discrimination was evaluated using the c-statistic.

Results: The average calibration slope was 1 and the average intercept was 0 for all combinations of outcomes and

biomarkers. As for FL, the c-statistic was 0.85 for FLI, 0.83 for ZJU, 0.82 for HSI, and 0.80 for LAP. As for NAFLD, the c-statistic was 0.77 for FLI, 0.76 for ZJU, 0.75 for HSI, and 0.74 for LAP. All the biomarkers were strongly correlated with each other.

Conclusion: FLI, LAP, HSI, and ZJU can be used to diagnose FL in the Bagnacavallo population, even if FLI has the highest discriminative ability. The same biomarkers perform similarly for the diagnosis of NAFLD even if FLI has a small advantage as discrimination is concerned.

Definitions

Fatty Liver Index
Defined by National Cancer Institute

Fatty liver (liver steatosis)
Defined by EASL–EASD–EASO

Non-alcoholic fatty liver disease (NAFLD)
Defined by EASL–EASD–EASO

Francesco Giuseppe Foschi and Giorgio Bedogni contributed equally to the present work. The Bagnacavallo Study Group includes Pietro Andreone, Anna Chiara Dall'Aglio, Mauro Bernardi, Lauro Bucchi, Francesca Dazzani, Fabio Falcini, Arianna Lanzi, Alessandra Ravaioli, Margherita Rimini, Giulia Rovesti, Gaia Saini, Giuseppe Francesco Stefanini.

Corresponding Author: Dr. Giorgio Bedogni, Clinical Epidemiology Unit, Liver Research Center, Building Q, AREA Science Park, Strada Statale 14 km 163.5, 34012 Basovizza, Trieste, Italy; Email: giorgiobedogni@gmail.com

Keywords: cross-sectional study, diagnostic techniques and procedures, validation study, fatty liver, non-alcoholic fatty liver disease.

Short title: Validation of biomarkers of fatty liver in the general population.

ABBREVIATIONS

95%CI = 95% confidence interval
AFLD = alcoholic fatty liver disease
ALT = alanine transaminase
AST = aspartate transaminase
FL = fatty liver
GGT = gamma-glutamyltransferase
FLI = fatty liver index

HBV = hepatitis B virus

HCV = hepatitis C virus

HSI = hepatic steatosis index

LAP = lipid accumulation product

LUS = liver ultrasonography

MAFLD = metabolic dysfunction-associated fatty liver disease

NAFLD = non-alcoholic fatty liver disease

ZJU = Zhejiang University index

Introduction

Fatty liver (liver steatosis), the most common liver disease worldwide, has been classified into non-alcoholic fatty liver disease (NAFLD) and alcoholic fatty liver disease (AFLD) for almost 40 years[1]. Such dichotomization has been increasingly criticized so that an international panel of experts has recently proposed to abandon the NAFLD definition, adopting instead the more comprehensive definition of metabolic dysfunction-associated fatty liver disease (MAFLD), which has the advantage of being independent of alcohol intake[2][3][4].

Independently of its etiology, FL is operationally defined as visible steatosis in more than 5% of hepatocytes at liver biopsy or as an intrahepatic triglyceride content of at least 5.6% at magnetic resonance spectroscopy or magnetic resonance imaging[5]. Liver biopsy can be performed only in selected patients followed at tertiary care centers and the use of magnetic resonance techniques is restricted to few research centers because of its cost[5]. The method most commonly used to diagnose FL in both clinical practice and epidemiological research is liver ultrasonography (LUS)[5]. Another option, suggested by current guidelines to diagnose FL when LUS is not available, is the use of surrogate biomarkers of FL[5].

As for any diagnostic test, the performance of biomarkers of FL should be externally validated in terms of calibration and discrimination[6][7][8]. However, as it happens for most diagnostic research[6][7][8], calibration is often neglected by the available validation studies of FL biomarkers, with some notable exceptions[9][10]. Calibration is nonetheless the primary requirement to perform decision-making and inform patients, and a test with high discrimination but low (or unknown) calibration is not clinically useful[6][7][8].

We performed, therefore, an external validation of biomarkers of LUS-diagnosed FL in the general population of the Bagnacavallo study by evaluating both calibration and discrimination[11][12].

Subjects and methods

Sources of data

The validation of the FL biomarkers was performed using data collected during the Bagnacavallo Study[11][12]. The study was aimed at evaluating the prevalence of and the risk factors for FL in a cross-section of the general population of a Northern Italy town and at developing a cohort of subjects from the general population where the association between FL and incident health outcomes could be studied. The study was approved by the Ethical Committee of Area Vasta Romagna - IRST (reference number 112), and all subjects gave their written informed consent.

Participants

As described in detail elsewhere[11], 3933 citizens of the town of Bagnacavallo (Ravenna, Italy) aged 30 to 60 years, were studied between October 2005 and March 2009. Altered liver enzymes were defined as alanine transaminase (ALT) > 40 U/l and/or aspartate transaminase (AST) > 37 U/l, i.e., the upper limit of normal of the laboratory. After the exclusion of subjects with hepatitis B virus (HBV) infection, hepatitis C virus (HCV) infection, and lack of LUS, the Bagnacavallo cross-sectional analysis was performed on 349 citizens with and 1810 without altered liver enzymes[11]. The same sample of 2159 (349 + 1810) citizens was analyzed here. All participants underwent a detailed clinical history and physical examination[13]. Alcohol intake was assessed by interview[11]. Weight and height were measured following international guidelines[14] and waist circumference was measured at the midpoint between the last rib and the iliac crest[15]. Body mass index (BMI) was calculated as weight (kg) / height (m)^2[16]. Performed blood tests included: 1) glucose; 2) triglycerides; 3) total cholesterol; 4) high-density lipoprotein (HDL) cholesterol; 5) low-density lipoprotein (LDL) cholesterol; 6) ALT; 7) AST; 8) GGT. Systolic and diastolic blood pressure was measured using a sphygmomanometer following international guidelines. (The recommended method of measurement of blood pressure remained the same during the study period). The metabolic syndrome (MS) was diagnosed using the harmonized international definition[17].

Outcomes

The main outcome of the validation was FL diagnosed by LUS; the secondary outcome was NAFLD diagnosed by the same method. We focused on LUS not because it is the gold-standard method but because it is the most common option in practice, and we wanted to control the error attributable to the use of different standards.

LUS was performed by five experienced physicians using the same methodology of the Dionysos Nutrition & Liver Study[13]. In detail, normal liver was defined as the absence of liver steatosis or other liver abnormalities. Light FL was defined as the presence of slight bright liver or hepatorenal echo contrast without intrahepatic vessels blurring and no deep attenuation; moderate FL as the presence of mild bright liver or hepatorenal echo contrast without intrahepatic vessel blurring and with deep attenuation; and severe FL as diffusely severe bright liver or hepatorenal echo contrast, with intrahepatic vessels blurring (no visible borders) and deep attenuation without visibility of the diaphragm. For the present analysis, FL was coded as any degree of FL (0 = no; 1 = yes).

NAFLD was defined as FL associated with ethanol intake ≤ 2 alcohol units (20 g) / day in women and ≤ 3 alcohol units (30 g) / day in men testing negative for hepatitis B surface antigen and anti-HCV antibodies and not under treatment with steatogenic drugs[5]. Alcoholic fatty liver disease (AFLD) was defined as FL associated with ethanol intake ≥ 2 (20 g) alcohol units/day in women and ≥ 3 alcohol units (30 g) /day in men testing negative for hepatitis B surface antigen and anti-HCV antibodies and not under treatment with steatogenic drugs[5]. For the present analysis, NAFLD was coded as any degree of FL (0 = normal liver or AFLD; 1 = NAFLD).

Predictors

We identified five non-patented FL biomarkers developed using LUS as the reference method for potential inclusion into the study: fatty liver index (FLI)[18], lipid accumulation product (LAP)[15], hepatic steatosis index (HSI)[19], Zhejiang University index (ZJU)[20], and index of NASH (ION)[21].

FLI is suggested by the European Association for the Study of the Liver (EASL) as biomarker of liver steatosis[5]. Other biomarkers suggested by EASL are SteatoTest[22], which is based on a proprietary formula and could not be validated here, and the NAFLD-liver fat score[23], which was developed using magnetic resonance spectroscopy as the reference method and was therefore not considered here. We were also unable to calculate NAFLD-LFS because insulin, which is a required predictor of NAFLD-LFS, was available only in 1415 (66%) of our 2159 subjects. For the same reason and because of the unavailability of hip circumference, we could not to calculate the ION index, which requires both insulin and the waist-to-hip ratio. We could have imputed the missing values of insulin[12], but we did not do that because insulin is known to be a key predictor of FL[18] and missingness of key predictors should be avoided when developing or validating prediction models[7].

FLI and LAP were developed to predict FL while HSI and ZJU were developed to predict NAFLD. All biomarkers were developed, using LUS as the reference method, in cross-sections of individuals from the general population (FLI, LAP) or health-care facilities (HSI, ZJU) by matching individuals with FL or NAFLD to individuals without it. The formulae for calculating the biomarkers are given in **Appendix 1**.

Sample size

We did not perform any formal sample size calculation but were quite confident that with 896/2159 (42%) cases of FL and 567/2159 (26%) cases of NAFLD we could attain a precise assessment of the performance of the biomarkers[11]. At least 200 events and non-events are in fact required for reasonable external validation of model performance[6][7].

Missing data

There were no missing data.

Statistical analysis

Most continuous variables were not Gaussian-distributed, and all are reported as median (50th percentile) and interquartile range (25th and 75th percentiles). Discrete variables are reported as the number and proportion of subjects with the characteristic of interest. Calibration was evaluated by applying Van Calster's three-level hierarchy[8][24]. Level 1 of this hierarchy is "mean calibration" or "calibration-in-the-large", which compares the observed event rate with the average predicted risk. Level 2 is "weak calibration", which consists of a logistic calibration analysis testing whether the calibration slope is 1 and the calibration intercept is 0 and is aimed at revealing systematic overestimation or underestimation of risk. Level 3 is "moderate calibration", which evaluates whether the predicted risks correspond to the observed event rates using a calibration plot. Such a graph plots the predicted (expected) outcome probabilities (x-axis) against the observed outcome frequencies (y-axis). As suggested by TRIPOD[6], we performed the calibration using tenths of the predicted risk and superimposed a line obtained by locally weighted scatterplot smoothing[6]. A well-calibrated model shows predictions lying or around the 45° line of the calibration plot. Discrimination was evaluated using Harrell's c-statistic[25]. Statistical analysis was performed using Stata 16.1 (Stata Corporation, College Station, TX, USA) with the *pmcalplot* module[26], and R 4.0.3 (R Core Team 2020, Vienna, Austria) with the *val.prob.ci.2* function[8]. R code was run from within Stata using the *rcall* package[27].

Results

Study population

The measurements of the 2159 citizens who took part in the study are given in **Table 1** and are described in greater detail elsewhere[11][12]. FL was diagnosed in 896 (42%, 95%CI 39 to 44%) and NAFLD in 567 (26, %24 to 28%) of them.

Table 1 - Measurements of the study subjects.

	N = 2159
Altered liver enzymes	349 (16.2%)
Male sex	1079 (50.0%)
Age (years)	49 (41; 56)
Body mass index (kg/m^2)	25.5 (23.0; 29.0)
Fatty liver	896 (41.5%)
Fatty liver classification	
Normal liver	1263 (58.5%)
Non-alcoholic fatty liver disease	567 (26.3%)
Alcoholic fatty liver disease	329 (15.2%)
Waist circumference (cm)	101.0 (94.0; 108.0)
Glucose (mg/dl)	89 (83; 97)
Triglycerides (mg/dl)	102 (71; 153)
Total cholesterol (mg/dl)	209 (185; 235)
HDL cholesterol (mg/dl)	59 (49; 71)
LDL cholesterol (mg/dl)	128 (105; 152)
Systolic blood pressure (mm Hg)	130 (120; 140)
Diastolic blood pressure (mm Hg)	80 (80; 90)
Metabolic syndrome	615 (28.5%)
Alanine transaminase (U/l)	22 (16; 32)
Aspartate transaminase (U/l)	22 (18; 26)
Gamma-glutamyltransferase (U/l)	19 (13; 32)
Alcohol intake (units/day)	2 (0; 4)
Fatty liver index (FLI)	46 (21; 76)
Lipid accumulation product (LAP)	44 (28; 75)
Hepatic steatosis index (HSI)	39 (35; 44)
Zhejiang University index (ZJU)	36 (33; 41)

Continuous variables are reported as median (50th percentile) and interquartile range (25th and 75th percentile). Discrete variables are reported as the number and proportion of subjects with the characteristic of interest.

Diagnosis of FL

The average expected rate of FL (42%) equaled the average observed rate (42%) for all biomarkers, showing a satisfactory mean calibration (**Table 2**).

Table 2 – Calibration and discrimination of the fatty liver index, lipid accumulation product, hepatic steatosis index and Zhejiang University index at diagnosing fatty liver and non-alcoholic fatty liver disease.

	Fatty liver			
	FLI	LAP	HSI	ZJU
Expected event rate*	0.42 (0.40 to 0.43)	0.42 (0.40 to 0.43)	0.42 (0.40 to 0.43)	0.42 (0.40 to 0.43)
Calibration intercept	0.00 (-0.11 to 0.11)	0.00 (-0.10 to 0.10)	0.00 (-0.10 to 0.10)	0.00 (-0.10 to 0.10)
Calibration slope	1.00 (0.92 to 1.08)	1.00 (0.89 to 1.11)	1.00 (0.91 to 1.09)	1.00 (0.91 to 1.09)
C-statistic	0.85 (0.84 to 0.87)	0.80 (0.78 to 0.82)	0.82 (0.80 to 0.83)	0.83 (0.81 to 0.85)
	Non-alcoholic fatty liver disease			
	FLI	LAP	HSI	ZJU
Expected event rate**	0.26 (0.25 to 0.28)	0.26 (0.24 to 0.28)	0.26 (0.25 to 0.28)	0.26 (0.25 to 0.28)
Calibration intercept	0.00 (-0.11 to 0.11)	0.00 (-0.10 to 0.10)	0.00 (-0.10 to 0.10)	0.00 (-0.10 to 0.10)
Calibration slope	1.00 (0.89 to 1.11)	1.00 (0.84 to 1.16)	1.00 (0.88 to 1.12)	1.00 (0.88 to 1.12)
C-statistic	0.77 (0.75 to 0.79)	0.74 (0.71 to 0.76)	0.75 (0.72 to 0.77)	0.76 (0.74 to 0.78)

Values are averages and 95% confidence intervals
* *vs.* observed event rate of 0.42 (0.39 to 0.44)
** *vs.* observed event rate of 0.26 (0.24 to 0.28)

Abbreviations: FLI = fatty liver index; LAP = lipid accumulation product; HSI = hepatic steatosis index; ZJU = Zhejiang University index.

At logistic calibration, the average calibration slope was 1 and the average intercept was 0 for all biomarkers, showing a satisfactory weak calibration (**Table 2**). Lastly, the examination of calibration plots (**Figure 1**) showed an acceptable profile of moderate calibration for all predictors. FLI had the highest (0.85) c-statistic, followed by ZJU (0.83), HSI (0.82), and LAP (0.80).

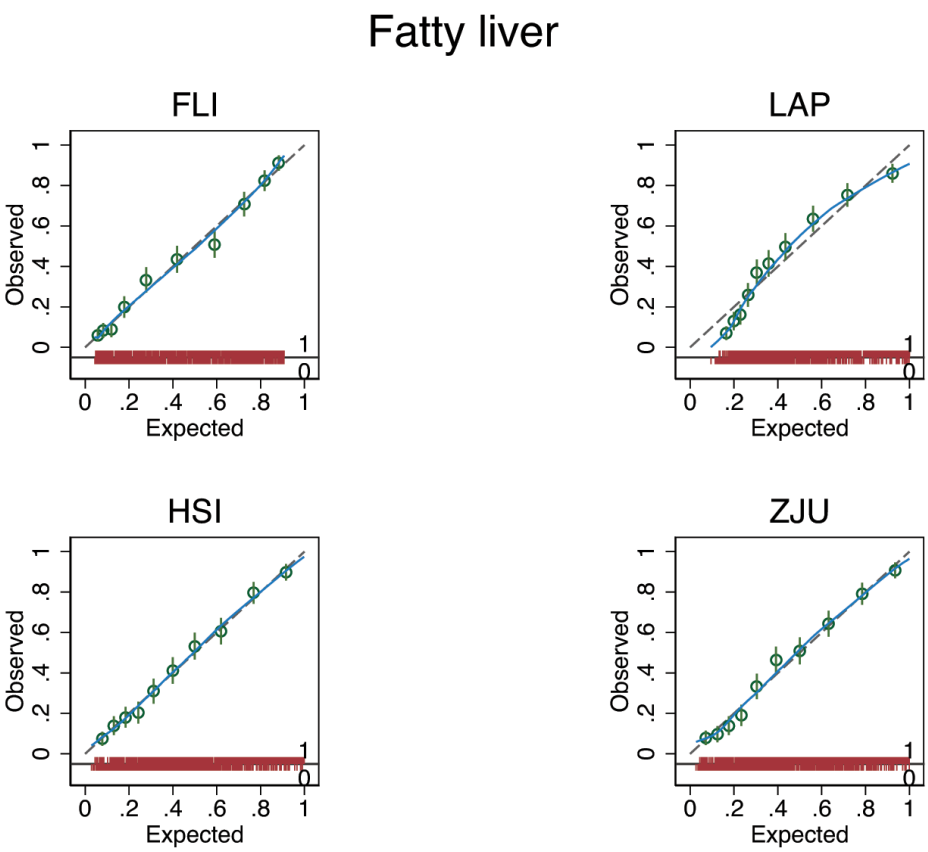

Figure 1: Calibration plots for the diagnosis of fatty liver. The expected (predicted) risk is divided into 10 equally sized groups (tenths). The green dots and spikes on the diagonal line are average risks and 95% confidence intervals. The dotted line is the reference line of calibration. The blue line connecting the green values is obtained by locally weighted scatterplot smoothing. The red spike plot at the bottom gives the distribution of fatty liver (0 = no; 1 = yes). Abbreviations: FLI = fatty liver index; LAP = lipid accumulation product; HSI = hepatic steatosis index; ZJU = Zhejiang University index.

Diagnosis of NAFLD

The expected rate of NAFLD (26%) equaled the observed rate (26%) for all biomarkers, showing a satisfactory mean calibration (**Table 2**). At logistic calibration, the average calibration slope was 1 and the average intercept was 0 for all biomarkers, showing a satisfactory weak calibration (**Table 2**). Lastly, the examination of calibration plots showed an acceptable profile of moderate calibration for all predictors (**Figure 2**). FLI had the highest (0.77) c-statistic, followed by ZJU (0.76), HSI (0.75), and LAP (0.74).

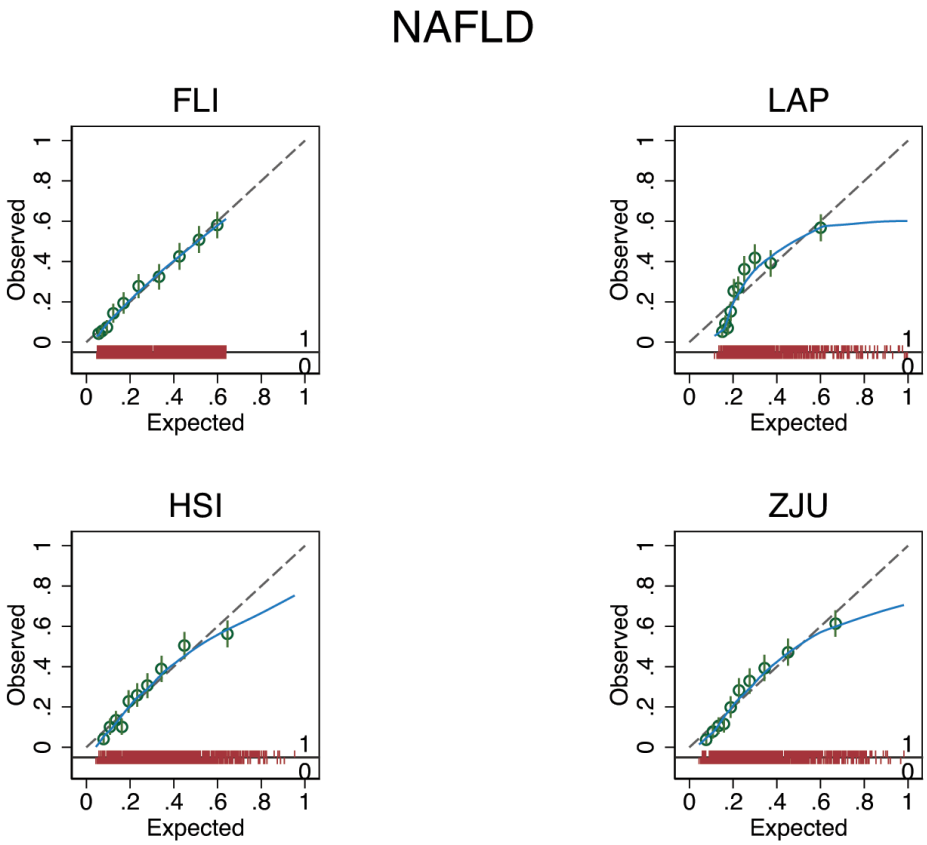

Figure 2: Calibration plots for the diagnosis of non-alcoholic fatty liver disease. The expected (predicted) risk is divided into 10 equally sized groups (tenths). The green dots and spikes on the diagonal line are average risks and 95% confidence intervals. The dotted line is the reference line of calibration. The blue line connecting the green values is obtained by locally weighted scatterplot smoothing. The red spike plot at the bottom gives the distribution of fatty liver (0 = no; 1 = yes). Abbreviations: NAFLD = non-alcoholic fatty liver disease; FLI = fatty liver index; LAP = lipid accumulation product; HSI = hepatic steatosis index; ZJU = Zhejiang University index.

Association between biomarkers

Further analysis revealed a strong association between all biomarkers (**Figure 3**), partially explained by the use of the same or highly correlated predictors (**Appendix 1**).

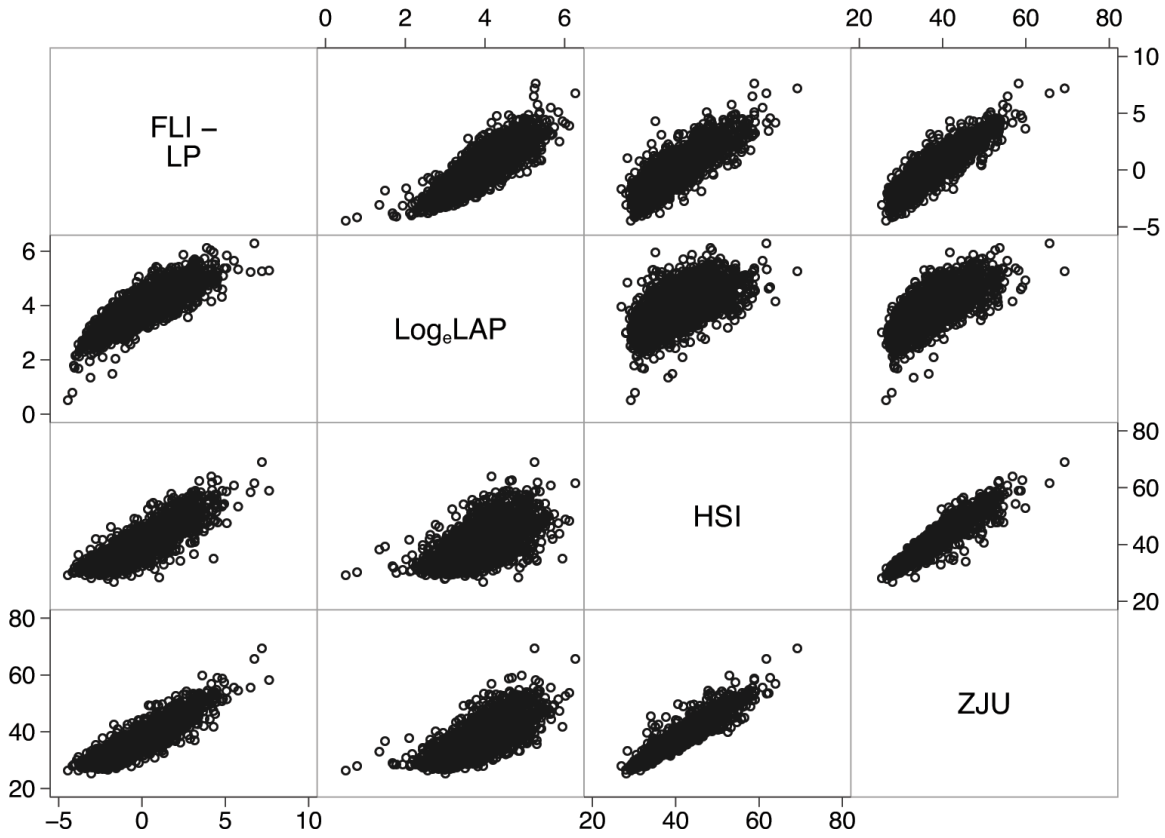

Figure 3: Correlation matrix showing strong associations between all biomarkers. Abbreviations: FLI-LP = fatty liver index-linear predictor (see **Appendix 1**); Log$_e$LAP = natural logarithm of the lipid accumulation product; HSI = hepatic steatosis index; ZJU = Zhejiang index.

For instance, the linear predictor of FLI explained 72% of the variance of HSI, 81% of the variance of ZJU, and 51% of the variance of log$_e$-transformed LAP. Moreover, ZJU explained 89% of the variance of HSI. The similar performance of these biomarkers at diagnosing FL and NAFLD (**Table 1**) is thus likely to be partially explained by their underlying mutual association.

Discussion

In the present study, we took advantage of the Bagnacavallo cross-sectional study of liver disease[11] to externally validate FLI[18], LAP[15], HSI[19], and ZJU[20] for the diagnosis of FL and NAFLD in the general population. All biomarkers showed an acceptable mean, weak, and moderate calibration for the diagnosis of both FL (**Figure 1** and **Table 2**) and NAFLD (**Figure 2** and **Table 2**)[8] (8).

We hypothesized that FLI would perform better than LAP, HSI, and ZJU at diagnosing FL and possibly NAFLD in the present population. (We had some reservations about NAFLD because FLI was purposely developed to predict FL.) Our

hypothesis was based on the fact that FLI was developed in the general population of a town (Campogalliano, Modena, Italy) similar to that studied here (Bagnacavallo, Ravenna, Italy)[18]. We nonetheless expected, in line with what happens for most diagnostic tests[7], that FLI had to be recalibrated for proper use in the Bagnacavallo population[9][10]. We were thus surprised to find that FLI had a satisfactory mean, weak and moderate calibration, and that it could be applied without modification to the Bagnacavallo population for the diagnosis of both FL and NAFLD. We were even more surprised to find that biomarkers (LAP, HSI and ZJU) developed in different populations (US, Korea and China) showed a satisfactory profile of mean, weak, and moderate calibration.

The strengths of the present study are that it was performed in a representative sample of the general population, that it enrolled a high number of subjects, and that it had a high observed event rate for both FL and NAFLD. A sample size of at least 200 subjects with and without the outcome of interest is presently suggested for proper validation of a diagnostic test[7]. With its 896 citizens with and 1263 without FL and 567 citizens with and 1592 without NAFLD, the Bagnacavallo Study is thus in an excellent position to serve as a platform to externally validate biomarkers of FL.

A limitation of the present study is the unavailability of some predictors needed to calculate the ION index[21], which was one of the biomarkers that we identified as theoretically suitable for validation in this population. ION employs insulin, which was available only in a subsample of subjects, and hip circumference, which was not measured in the Bagnacavallo study. Moreover, the partial availability of insulin and the unavailability of c-reactive protein impeded us to diagnose MAFLD and to evaluate the performance of the biomarkers at diagnosing this newly proposed entity, which is expected to attract much attention in coming years[2][3]. Furthermore, even if we choose to include studies that used LUS to diagnose FL to reduce the error attributable to different diagnostic methods, LUS is known to offer an accurate assessment of FL only starting from an intrahepatic triglyceride content of 10%[28].

External calibration is more important than discrimination at establishing the value of a test for a given diagnosis[6][7][8], but this fact is not taken into account by most diagnostic studies of FL biomarkers, with some notable exceptions[9][10]. This is not to say that discrimination is irrelevant as, in the presence of an acceptable calibration, the greater discrimination is the better. Another problem of most diagnostic studies is that they compare an externally derived predictor with an internally derived one and more often than not declare the latter superior to the former[6]. This is, however, largely expected on both theoretical and empirical grounds and is one of the primary reasons why external validation of diagnostic models is so important[6].

The similar performance of the biomarkers at diagnosing FL and NAFLD in the present study is likely to be partially explained by their underlying mutual association (**Figure 3**). This finding, which awaits replication in other populations, suggests that the same set of predictors may be employed to develop a common algorithm for the prediction of FL, by re-estimating some or all regression coefficients, and updating the model with new predictors if they can increase its performance[7].

Conclusion

In conclusion, we found that FLI, LAP, ZJU, and HSI can be satisfactorily used to diagnose FL and NAFLD in the Bagnacavallo population, even if FLI has the highest discriminative ability. These biomarkers are strongly associated and this is likely to partially explain their similar performance. Further studies are needed to evaluate the use of these biomarkers for the diagnosis of MAFLD[29], the diagnostic entity which is going to replace NAFLD[2][3][4].

Appendix 1

1.1 Predictors employed by the biomarkers

	Biomarkers			
	FLI	LAP	HSI	ZJU
Triglycerides	✓	✓		✓
BMI	✓		✓	✓
GGT	✓			
Waist	✓	✓		
ALT:AST			✓	✓
T2DM			✓	
Sex			✓	
Glucose				✓

Abbreviations: BMI = body mass index; GGT = gamma-glutamyltransferase; ALT = alanine transaminase; AST = aspartate transaminase; T2DM = type 2 diabetes mellitus.

1.2 Multivariable prediction models employed by the biomarkers

Abbreviations and units of measurements

altsalt = alanine transaminase (U/l) / aspartate transaminase (U/l)

bmi = body mass index (kg/m2)

exp = exponential operator

female = female sex (1 = female; 0 = male)

ggt = gamma-glutamyltransferase (U/l)

gmmol = glucose (mmol/l)

$\log_e$ = natural logarithm

fli_lp = fatty liver index - linear predictor

t2dm = type 2 diabetes mellitus (1 = yes; 0 = no);

tg = triglycerides (mg/dl)

tgmmol = triglycerides (mmol/l)

wc = waist circumference (cm)

Fatty liver index (FLI)

fli_lp = 0.953*$\log_e$(tg) + 0.139*bmi + 0.718* $\log_e$(ggt) + 0.053*wc - 15.745

FLI = [(exp(fli_lp) / (1 + exp(fli_lp)]*100

Lipid accumulation product (LAP)

LAP = (wc-k)*tgmmol

k=65 if sex==male or k=58 if sex==female

Hepatic steatosis index (HSI)

HSI = 8*altast + bmi + 2*t2dm + 2*female

Zhejiang University index (ZJU)

ZJU = bmi + gmmol + tgmmol + 3*altast +2*female

References

1. ^*Yasser Fouad, Imam Waked, Steven Bollipo, Ahmed Gomaa, Yousef Ajlouni, Dina Attia. (2020). What's in a name? Renaming 'NAFLD' to 'MAFLD'. Liver Int, vol. 40 (6), 1254-1261. doi:10.1111/liv.14478.*
2. a, b, c*Mohammed Eslam, Philip N. Newsome, Shiv K. Sarin, Quentin M. Anstee, Giovanni Targher, Manuel Romero-Gomez. (2020). A new definition for metabolic dysfunction-associated fatty liver disease: An international expert consensus statement. Journal of Hepatology, vol. 73 (1), 202-209. doi:10.1016/j.jhep.2020.03.039.*
3. a, b, c*Mohammed Eslam, Arun J. Sanyal, Jacob George, Arun Sanyal, Brent Neuschwander-Tetri, Claudio Tiribelli. (2020). MAFLD: A Consensus-Driven Proposed Nomenclature for Metabolic Associated Fatty Liver Disease. Gastroenterology, vol. 158 (7), 1999-2014.e1. doi:10.1053/j.gastro.2019.11.312.*
4. a, b*Vlad Ratziu, Mary Rinella, Ulrich Beuers, Rohit Loomba, Quentin M. Anstee, Stephen Harrison. (2020). The times they are a-changin' (for NAFLD as well). Journal of Hepatology. doi:10.1016/j.jhep.2020.08.028.*
5. a, b, c, d, e, f, g*(2016). EASL–EASD–EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. Journal of Hepatology, vol. 64 (6), 1388-1402. doi:10.1016/j.jhep.2015.11.004.*
6. a, b, c, d, e, f, g, h, i*Karel G.M. Moons, Douglas G. Altman, Johannes B. Reitsma, John P.A. Ioannidis, Petra Macaskill, Ewout W. Steyerberg. (2015). Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. Ann Intern Med, vol. 162 (1), W1. doi:10.7326/m14-0698.*
7. a, b, c, d, e, f, g, h, i*Steyerberg EW. (2019). Clinical prediction models. Cham, Switzerland: Springer International.*
8. a, b, c, d, e, f, g*Ben Van Calster, Daan Nieboer, Yvonne Vergouwe, Bavo De Cock, Michael J. Pencina, Ewout W.*

Steyerberg. (2016). A calibration hierarchy for risk models was defined: from utopia to empirical data. Journal of Clinical Epidemiology, vol. 74 , 167-176. doi:10.1016/j.jclinepi.2015.12.005.

9. [a, b, c] *Ya-Nan Shen, Ming-Xing Yu, Qian Gao, Yan-Yan Li, Jian-Jun Huang, Chen-Ming Sun. (2017). External validation of non-invasive prediction models for identifying ultrasonography-diagnosed fatty liver disease in a Chinese population. Medicine, vol. 96 (30), e7610. doi:10.1097/md.0000000000007610.*
10. [a, b, c] *Peter J Meffert, Sebastian E Baumeister, Markus M Lerch, Julia Mayerle, Wolfgang Kratzer, Henry Völzke. (2014). Development, External Validation, and Comparative Assessment of a New Diagnostic Score for Hepatic Steatosis. American Journal of Gastroenterology, vol. 109 (9), 1404-1414. doi:10.1038/ajg.2014.155.*
11. [a, b, c, d, e, f, g, h] *Francesco Giuseppe Foschi, Giorgio Bedogni, Marco Domenicali, Pierluigi Giacomoni, Anna Chiara Dall'Aglio, Francesca Dazzani. (2018). Prevalence of and risk factors for fatty liver in the general population of Northern Italy: the Bagnacavallo Study. BMC Gastroenterol, vol. 18 (1). doi:10.1186/s12876-018-0906-8.*
12. [a, b, c, d] *Francesco Giuseppe Foschi, Marco Domenicali, Pierluigi Giacomoni, Anna Chiara Dall'Aglio, Fabio Conti, Alberto Borghi. (2020). Is there an association between commonly employed biomarkers of liver fibrosis and liver stiffness in the general population?. Annals of Hepatology, vol. 19 (4), 380-387. doi:10.1016/j.aohep.2020.04.003.*
13. [a, b] *Giorgio Bedogni, Lucia Miglioli, Flora Masutti, Claudio Tiribelli, Giulio Marchesini, Stefano Bellentani. (2005). Prevalence of and risk factors for nonalcoholic fatty liver disease: The Dionysos nutrition and liver study. Hepatology, vol. 42 (1), 44-52. doi:10.1002/hep.20734.*
14. [^] *Lohman TG, Roche AF, Martorell R. (1991). Anthropometric standardization reference manual. Champaign, Illinois: Human Kinetics Books.*
15. [a, b, c] *Giorgio Bedogni, Henry S Kahn, Stefano Bellentani, Claudio Tiribelli. (2010). A simple index of lipid overaccumulation is a good marker of liver steatosis. BMC Gastroenterol, vol. 10 (1). doi:10.1186/1471-230x-10-98.*
16. [^] *National Institutes of Health. (1998). Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults. The Evidence Report. Obes Res, vol. 6 Suppl 2:51S-209S . National Institutes of Health.*
17. [^] *K.G.M.M. Alberti, Robert H. Eckel, Scott M. Grundy, Paul Z. Zimmet, James I. Cleeman, Karen A. Donato. (2009). Harmonizing the Metabolic Syndrome. Circulation, vol. 120 (16), 1640-1645. doi:10.1161/circulationaha.109.192644.*
18. [a, b, c, d] *Giorgio Bedogni, Stefano Bellentani, Lucia Miglioli, Flora Masutti, Marilena Passalacqua, Anna Castiglione. (2006). The Fatty Liver Index: a simple and accurate predictor of hepatic steatosis in the general population. BMC Gastroenterol, vol. 6 (1). doi:10.1186/1471-230x-6-33.*
19. [a, b] *Jeong-Hoon Lee, Donghee Kim, Hwa Jung Kim, Chang-Hoon Lee, Jong In Yang, Won Kim. (2010). Hepatic steatosis index: A simple screening tool reflecting nonalcoholic fatty liver disease. Digestive and Liver Disease, vol. 42 (7), 503-508. doi:10.1016/j.dld.2009.08.002.*
20. [a, b] *Jinghua Wang, Chengfu Xu, Yunhao Xun, Zhenya Lu, Junping Shi, Chaohui Yu. (2015). ZJU index: a novel model for predicting nonalcoholic fatty liver disease in a Chinese population. Sci Rep, vol. 5 (1). doi:10.1038/srep16494.*
21. [a, b] *Munkhzul Otgonsuren, Michael J. Estep, Nayeem Hossain, Elena Younossi, Spencer Frost, Linda Henry. (2014). A single non-invasive model to diagnose non-alcoholic fatty liver disease (NAFLD) and non-alcoholic steatohepatitis (NASH). J Gastroenterol Hepatol, vol. 29 (12), 2006-2013. doi:10.1111/jgh.12665.*
22. [^] *Thierry Poynard, Vlad Ratziu, Sylvie Naveau, Dominique Thabut, Frederic Charlotte, Djamila Messous. (2005).*

Comparative Hepatology. Comp Hepatol, vol. 4 (1), 10. doi:10.1186/1476-5926-4-10.

23. ^Anna Kotronen, Markku Peltonen, Antti Hakkarainen, Ksenia Sevastianova, Robert Bergholm, Lina M. Johansson. (2009). *Prediction of Non-Alcoholic Fatty Liver Disease and Liver Fat Using Metabolic and Genetic Factors.* Gastroenterology, vol. 137 (3), 865-872. doi:10.1053/j.gastro.2009.06.005.
24. ^Richard J. Stevens, Katrina K. Poppe. (2020). *Validation of clinical prediction models: what does the "calibration slope" really measure?.* Journal of Clinical Epidemiology, vol. 118 , 93-99. doi:10.1016/j.jclinepi.2019.09.016.
25. ^Harrell F. (2016). Regression Modeling Strategies. Cham, Switzerland: Springer.
26. ^Ensor J, Snell K, IE., Martin E, C.. (2018). *PMCALPLOT: Stata module to produce calibration plot of prediction model performance.* Statistical Software Components.
27. ^E. F. Haghish. (2019). *Seamless interactive language interfacing between R and Stata.* The Stata Journal, vol. 19 (1), 61-82. doi:10.1177/1536867x19830891.
28. ^Ruben Hernaez, Mariana Lazo, Susanne Bonekamp, Ihab Kamel, Frederick L. Brancati, Eliseo Guallar. (2011). *Diagnostic accuracy and reliability of ultrasonography for the detection of fatty liver: A meta-analysis.* Hepatology, vol. 54 (3), 1082-1090. doi:10.1002/hep.24452.
29. ^Zhou Xu, Hao Li, Shen Tian, Juan Wu, Xin Li, Zi-li Liu. (2020). *Blood biomarkers for the diagnosis of hepatic steatosis in metabolic dysfunction-associated fatty liver disease.* Journal of Hepatology, vol. 73 (5), 1264-1265. doi:10.1016/j.jhep.2020.06.003.