

Transaminase concentrations cannot separate NAFL and NASH in morbidly obese patients independent of histological algorithm

Peter Lemmer

Department of Internal Medicine, University Hospital Knappschaftskrankenhaus Bochum, Ruhr-University Bochum, Bochum, Germany

Nicole Selbach

Department of Internal Medicine, University Hospital Knappschaftskrankenhaus Bochum, Ruhr-University Bochum, Bochum, Germany

Theodor Baars

Department of Internal Medicine, University Hospital Knappschaftskrankenhaus Bochum, Ruhr-University Bochum, Bochum, Germany

Mustafa Özcürümez

Department of Internal Medicine, University Hospital Knappschaftskrankenhaus Bochum, Ruhr-University Bochum, Bochum, Germany

Dominik Heider

Department of Mathematics and Computer Science, University of Marburg, Marburg, Germany.

Ali Canbay (✉ ali.canbay@rub.de)

Department of Internal Medicine, University Hospital Knappschaftskrankenhaus Bochum, Ruhr-University Bochum, Bochum, Germany

Jan-Peter Sowa

Department of Internal Medicine, University Hospital Knappschaftskrankenhaus Bochum, Ruhr-University Bochum, Bochum, Germany

Research Article

Keywords: liver, obesity, non-invasive prediction, liver enzymes, NAFLD, alanine-aminotransferase, aspartate-aminotransferase, gamma-glutamyltransferase, NAFLD activity score, SAF score

Posted Date: December 21st, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-129749/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Prevalence of non-alcoholic fatty liver disease (NAFLD) in the German population is 20–30%. Liver biopsy to detect non-alcoholic steatohepatitis (NASH) let alone for NAFLD monitoring is not feasible. Current practice regards elevated serum concentrations of liver enzymes as indicator for NAFLD or NASH. In this study we analyzed if an adjustment of the upper limit of normal (ULN) for serum liver enzymes can improve their diagnostic accuracy.

Methods: Data from 363 morbidly obese patients (42.5 ± 10.3 years old; mean BMI: 52 ± 8.5 kg/m²), who underwent bariatric surgery were retrospectively analyzed. All patients had histologically confirmed NAFL or NASH (NAS and SAF).

Results: In 121 women (45%) and 45 men (46%) elevated values for at least one serum parameter (ALT, AST, γ GT) were present. The serum concentrations of ALT ($p < 0.0001$), AST ($p < 0.0001$) and γ GT ($p = 0.0023$) differed significantly between NAFL and NASH, independent of classification method (NAS, SAF). Concentrations of all three serum parameters correlated significantly positively with the NAS and the SAF score, with correlation coefficients between 0.33 (ALT/NAS) and 0.40 (GGT/SAF). The AUROCs to separate NAFL and NASH by liver enzymes achieved a maximum of 0.70 (ALT applied to NAS-based classification). For 95% specificity the ULN for ALT would be 47.5 U/l; for 95% sensitivity, the ULN for ALT would be 17.5 U/l, resulting in 62% uncategorized patients.

Conclusion: ALT, AST and GGT are unsuitable for non-invasive screening or diagnosis of NAFL or NASH. Utilizing liver enzymes as an indicator for NAFLD/NASH should generally be questioned.

Introduction

Nonalcoholic fatty liver disease (NAFLD) is a highly prevalent disease occurring in 20–25% of the global population and is currently the main chronic liver disease in the Western world ¹. Since obesity and insulin resistance (IR) are crucial elements in the pathophysiology of NAFLD, recently a new term has been proposed to label this disease as MAFLD: Metabolism-associated fatty liver disease ^{2,3}. This is important because the textbook definition of NAFLD strictly excludes relevant alcohol consumption and other liver related diseases. In reality humans are diverse in habits and diet and cases with co-existing NAFLD/MAFLD with alcohol consumption or, i.e., hepatitis C infection are not uncommon.

Due to the increasing prevalence of obesity in the last five decades, detection and characterization of NAFLD has become more and more important for two main reasons. Firstly, NAFLD is the hepatic manifestation of the metabolic syndrome and can therefore be a part of the risk stratification for cardiovascular diseases and other conditions associated to the metabolic syndrome ^{4,5}. Secondly, Nonalcoholic steatohepatitis (NASH), the progressive form of NAFLD with inflammation and ballooning of hepatocytes in the liver is associated with substantial risk of morbidity and mortality ^{6–8}. It is debatable if NAFL (liver steatosis without relevant inflammation) and NASH have similar risk for

cardiovascular and other metabolic comorbidities and both develop significant fibrosis⁹⁻¹². But it is clear that NASH is associated with a greater risk of cirrhosis and developing a hepatocellular carcinoma (HCC) with or without cirrhosis^{13,14}. The risk of liver-related mortality is also greater in NASH.

A common method in the detection of NAFLD is ultrasound. However, while NAFLD is histologically defined by the presence of steatosis in at least 5% of hepatocytes, usually 20% liver lipid content (depending on operator experience and the examination conditions) is necessary for the detection of liver steatosis by ultrasound¹⁵. Elevated serum liver enzymes (ALT, AST, γ GT) are the most common symptom prompting further examinations. However, up to 75% of NAFLD patients do not show elevated liver enzymes, when applying current reference values¹⁶. This situation leads to severe underdiagnosis of NAFLD and NASH in real life practice.

Liver biopsy and histological assessment is still considered the gold standard for the diagnosis of NAFLD and the separation of its subgroups non-alcoholic fatty liver (NAFL) and NASH¹⁵. The diagnosis is based on semi-quantitative combination of the histological features steatosis, ballooning, and lobular inflammation. By now two scoring systems have been established: The NAFLD activity score (NAS) and the Steatosis activity fibrosis score (SAF score). Whereas NAS is the unweighted summation of semi-quantitative evaluation of steatosis, ballooning, and lobular inflammation (with an increasing likelihood of NASH depending on the scoring points), the SAF score requires each one of those criteria to be met for the diagnosis NASH^{15,17}. This gold standard is limited in applicability, as there is a small risk for complications (about 1%), sampling variability and intra-observer differences can be substantial, since one biopsy only represents about 1/50000 of liver tissue^{18,19} and an effective screening of 20% cannot be performed via liver biopsy.

Due to the increasing global prevalence of NAFLD and the limitations of above described diagnostic methods described above, non-invasive screening methods are urgently warranted. Aim of the present study was to test whether serum liver enzyme concentrations can be applied to detect NAFLD and/or separate NAFL from NASH, independent from current reference values.

Results

Differences of demographic data between NAFL and NASH depend on assessment method

In total, data of 363 morbidly obese patients undergoing bariatric surgery was analyzed. The whole cohort was comprised of 270 women (74%) and 93 men (26%) with an mean age of 42.5 ± 10.3 years and mean BMI of 52 ± 8.5 kg/m². The mean ALT serum concentration before surgery was 37.2 ± 26.4 U/l, and mean serum AST was 31.3 ± 19.0 U/l, respectively. The median NAS was 4 (1–8) and the median SAF score 6 (1–12).

When NAS was applied for categorization 160 patients (44%) had NAFL (NAS 1–4; 80% female, 20% male) and 203 patients (56%) had NASH (NAS \geq 5; 70% female, 30% male; Fisher's exact test $p = 0.03$).

There was no significant difference in age (41.9 ± 10.4 vs. 43.1 ± 10.3) and BMI (51.5 ± 8.7 kg/m² vs. 52.4 ± 8.7 kg/m²) between NAFL and NASH.

Applying the SAF score 106 patients (29%) were categorized as NAFL (78% female, 22% male) and 257 patients (71%) were categorized as NASH (73% female, 27% male). Based on the SAF classification the patients with NAFL were significantly younger than the patients with NASH (41.0 ± 11.1 vs. 43.2 ± 9.9 ; $p = 0.048$), but there was no significant difference in the BMI (51.2 ± 8.8 kg/m² vs. 52.3 ± 8.4 kg/m²) and sex. Disagreement between the two classification systems was 20% of all patients with 64 patients categorized as NASH (by SAF) instead of NAFL and 10 patients categorized as NAFL (by SAF) instead of NASH. Accordingly, distribution into NAFL and NASH between NAS and SAF differed significantly ($p < 0.0001$ Fisher's exact test).

Serum concentrations of liver enzymes differ significantly between NAFL and NASH

The majority of patients did not show an elevation of liver enzyme serum concentrations above the current upper limit of normal (AST/ALT: 35 U/l, γ GT: 40 U/l for women; AST/ALT: 50 U/l, γ GT: 60 U/l for men). In 121 women (45%) and 45 men (46%) elevated concentrations were found for at least one of the factors ALT, AST or γ GT. Serum concentrations of ALT ($p < 0.0001$), AST ($p < 0.0001$) and γ GT ($p < 0.0023$) differed significantly between NAFL and NASH, regardless of classification by NAS (Fig. 1) or SAF score (Fig. 2)

Liver enzyme serum concentrations are correlated with NAS and SAF score

As all liver enzyme concentrations differed significantly between NAFL and NASH, we investigate if there was a direct correlation between histological scores and serum concentrations. Calculation of spearman rank coefficients resulted in significantly positive correlations (all $p < 0.0001$) of all three serum parameters (ALT, AST γ GT) with the NAS as well as the SAF score. Correlation coefficients ranged between 0.33 (ALT / NAS) and 0.40 (γ GT / SAF), with a detailed overview given in Table 2. Overall, slightly, but not significantly weaker correlations were detected between NAS and the serum concentrations compared to those with the SAF score. The highest correlation coefficients were reached by γ GT with NAS or the SAF.

Table 2
Overview of spearman rank correlation coefficients of serum liver enzyme concentrations with histological scores for NAFLD assessment.

Liver enzyme (serum concentration)	Correlation with NAS [r; 95% confidence interval]	Correlation with SAF [r; 95% confidence interval]
ALT	0.3343; 0.2369–0.4251	0.3728; 0.2780–0.4605
AST	0.3709; 0.2759–0.4587	0.3867; 0.2929–0.4731
γ GT	0.3814; 0.2869–0.4686	0.4096; 0.3173–0.4943
All correlations reached a $p < 0.0001$.		

None of the serum liver enzymes is suitable for discrimination of NAFL from NASH

Following the above described analyses we assessed whether serum concentrations of liver enzymes can separate NAFL and NASH. To this end receiver operating characteristics (ROC) curves were generated for all three parameters to discern NAFL and NASH, either classified by NAS or SAF. The resulting area under the curve (AUC) represents the ability of each parameter to distinguish NAFL from NASH (with 0.5 equivalent to random guessing).

In the present study, the highest AUC of 0.7073 was achieved for ALT to discriminate NAS-based NAFL vs. NASH, which demonstrates a poor to moderate separation. ROCs with detailed analyses for each combination of serum liver enzyme and binary NAFL/NASH classification by NAS or SAF are depicted in Figs. 3 and 4. None of the potential classifiers exhibited sufficient performance for a clinically relevant separation of NAFL vs. NASH. Despite this result cut-off values for ALT, AST and γ GT for either 95% specificity or 95% sensitivity were calculated (Table 3). For example, to achieve 95% specificity, the threshold for ALT would have to be 17.5 U/l and a threshold of 47.5 U/l would result in 95% sensitivity. The diagnostic gap between these thresholds leaves 62% patients uncategorized. With the same two-threshold approach (each at 95 specificity/sensitivity) 91% of the patients would remain uncategorized for AST and 77% patients for γ GT.

Table 3
Required serum liver enzyme cut-offs for high sensitivity or specificity separation between NAFL and NASH

Liver enzyme (serum concentration)	Cut-off for 95% sensitivity	Cut-off for 95% specificity	Diagnostic Gap*
ALT	17.5 U/l	47.5 U/l	62% (52% of these NASH)
AST	15.5 U/l	65.5 U/l	91% (55% of these NASH)
GGT	19.5 U/l	95 U/l	77% (58% of these NASH)
*: % of patients with value between cut-offs			

Discussion

With a global prevalence of NAFLD at 20–30% and a similar estimated prevalence in the German population¹ it has become more and more important to enhance diagnostic efficiency for this disease. Liver biopsy is still considered gold standard for confirmation of NAFLD, for risk assessment of liver cirrhosis and HCC, and the only method to detect and diagnose NASH¹⁵. Given the limitations of this method and the very low proportion of patients, who actually undergo liver biopsy, it is no surprise that

NAFLD and in particular asymptomatic progressive liver disease (i.e. NASH with compensated cirrhosis) remain severely underdiagnosed⁷. As elevations of serum liver enzyme concentrations are still the most common symptom resulting in further clinical workup and referral to hepatologists^{20,21}, in the present study utility of serum liver enzymes for detection or diagnosis of NAFLD or NASH was tested. Despite significant differences of ALT, AST and γ GT serum concentrations between NAFL and NASH and significant correlations of these factors with NAS and SAF, performance to separate NAFL and NASH was insufficient for clinical application.

The central finding of this retrospective data analysis is that even in morbidly obese individual serum liver enzyme concentrations cannot efficiently separate NAFL and NASH. All AUROCs calculated to discern NAFL from NASH demonstrated only poor to mediocre performance of liver serum enzymes as classifiers with a maximum AUROC of 0.7073. For clinical use, this value should be at least above 0.85. Additionally 55% of the whole cohort did not exhibit elevations of any of the three tested liver enzymes and 63% of those without elevations had NASH. This confirms previous findings that normal serum concentrations of liver transaminases cannot reliably rule out NAFLD or the need for a liver biopsy. In a study analyzing histological and clinical features of NAFLD it has been shown that 59% of patients with normal serum liver transaminases had NASH and 35% had advanced fibrosis or cirrhosis²². Another study aimed at serum based identification of NASH found that in patients with normal serum transaminases 46% had NASH²³. In a direct comparison between 63 NAFLD patients with normal transaminase and 395 NAFLD patients with elevated transaminases, no difference in the distribution of fibrosis grades was found in 59% of patients with normal and in 74% of patients with elevated transaminases NASH was present. In a large population based study with more than 4900 participants (Heinz Nixdorf Recall Study) we identified the proportion of diabetics as 13.7%¹⁶. In that study elevated transaminases were found in only 2% of the participants, although 90% of patients with type 2 diabetes mellitus have NAFLD and approximately 38% NASH²⁴. The authors of some of these works suggest that including measures of glucose metabolism and IR (i.e. HbA1c) should be included into detection algorithms²³. This is in line with a serum based score to separate NAFL and NASH, recently introduced by us²⁵. The included parameters age, γ GT, HbA1c, M30, and adiponectin were selected from all available parameters by a nonbiased ensemble feature selection, including machine learning algorithms²⁶. Of note, with adiponectin and HbA1c two factors associated to glucose metabolism and IR (and adipose tissue status) were selected. To develop an efficient, clinically applicable serum based diagnostic algorithm sufficiently powered prospective studies assessing clinical outcome and development of various markers over the disease course are required. Measures of IR and glucose metabolism should be included in such an analysis.

It has been argued that adapting the current limits of normal for serum liver enzymes would improve their diagnostic accuracy. However, when calculating optimized cut-offs for either 95% sensitivity or specificity from our data between 62% and 91% of patients would remain unclassified with a lower boundary of 15 to 20 U/l. In the past, various studies aimed to lower the upper limits of normal for ALT in different patient populations. Studies from China (22.15 U/l for men, 22.40 U/l for women), USA (30 U/l and 19 U/l), Taiwan (22 U/l and 17 U/l) and India (27 U/l and 17 U/l) resulted in dramatically lower thresholds, in a

similar range as those identified by us²⁷⁻³⁰. Based on these and further studies, current AASLD guidelines recommend a hepatitis B-screening at ALT serum concentrations above 35 U/l for men and 25 U/l for women³¹. However, to achieve a 95% sensitivity for the detection of NASH in our cohort, the cut-off would have to be even lower (AST: 15.5 U/l; γ GT: 19.5 U/l). Lowering of the upper limit of normal values to the point where adequate sensitivity is reached, would lead to a very high rate of false positive results for NAFLD/NASH when used as screening in the general population. In our opinion these findings of other groups and our own rule out any applicability of the liver serum values for diagnostic or screening purposes in NAFLD.

Another result of this study is that histological evaluation by NAS and by SAF differed in a substantial proportion (20%) of the study population. NAS and SAF were designed for different purposes and the NAS classification is usually broken down into three groups with 1-2 points as definitely no NASH, 3-4 as borderline NASH (requiring further examination or follow up) and ≥ 5 as definite NASH. To reduce ambiguity and allow more robust statistical evaluation we applied a simplified classification, interpreting NAS up to 4 as NAFL. This might be one cause, why 64 patients more were classified as NASH by SAF than by NAS. However, 10 patients were classified as NAFL by SAF and NASH by NAS. A study with a more diverse patient population, different outcome scenarios, and rigorous histological work-up would be required to investigate if one scoring system is superior in identifying patients at higher health risk. With our current understanding³²⁻³⁴ the observed differences between NAS and SAF classification might be a matter of interpretation in borderline NASH cases.

A direct comparison of serum liver enzyme concentrations yielded a significant difference between NAFL and NASH, independent of classification by NAS or SAF. The effect size was small and the overlap of the range of values was quite broad. Furthermore, all three serum parameters correlated significantly positively with the NAS and the SAF score, with slightly stronger correlations with the SAF score. Overall the rank correlation coefficients indicate a modest correlation between the histological assessed liver injury and serum concentrations of enzymes. However, due to this positive correlation and the small but significant difference between NAFL and NASH serum liver enzymes might support surveillance. When the diagnosis NAFLD is established (i.e. histologically or by ultrasound) an increase of serum liver enzymes, even within the current normal range, would indicate disease progression. Though, this has to be tested in an appropriately designed study. Our finding that the serum parameters exhibited slightly stronger correlations with the SAF score than with NAS can only be attributed to the addition of the fibrosis grade to the SAF score. This clearly indicates that serum liver enzymes reflect not only injury of the hepatic parenchyma but also fibrogenesis. It has been shown multiple times that fibrogenesis and progression to advanced fibrosis occurs in a similar proportion of NAFL and NASH patients^{10,35-37}, albeit possibly with different pace. The hypothesis that serum liver parameters reflect liver injury and fibrogenesis in parallel would explain the limited clinical benefit for screening or diagnostic purposes in NAFLD, as progression of fibrosis is rather rare in this disease.

In summary, our data fortify findings of the past 15 years, that elevation of serum liver enzymes ALT, AST and γ GT is no reliable sign for NASH or progressive NAFLD. Conversely, serum concentrations of these

factors in normal range do not exclude NASH. An adaptation of normal ranges would probably dramatically increase false positive results without enhancing clinical risk assessment. Serum liver enzymes might still have use in disease surveillance, when NAFLD or NASH will be established by other diagnostic measures. It is time to move on from serum liver enzyme elevations in NAFLD detection or risk assessment and focus studies on other markers, in particular those related to adipose tissue dysfunction, glucose metabolism, and IR.

Methods

Patients

Morbidly obese patients undergoing bariatric surgery were recruited at the Alfried-Krupp-Krankenhaus Essen, Essen, Germany. Dietary and exercise counselling was offered to patients for 6 months prior surgery, no calorie restriction was imposed. A blood sample was collected for assessment of serum derived factors on the day of surgery (prior surgery) and liver tissue was sampled during bariatric surgery. In liver tissue steatosis, ballooning, lobular inflammation, and fibrosis were assessed by two independent pathologists. NAFLD and severity thereof was diagnosed histologically according to Kleiner et al.⁴⁰. The SAF was calculated according to Bedossa et al.¹⁷. All data shown were recorded on the day of surgery.

Overall 694 patients were recruited from 2004 until 2017. For this data analysis patients were eligible, when measurements of alanine- and aspartate aminotransferase serum concentrations (ALT, AST) and liver histology were available, with a NAS of at least 1. To limit groups for a robust analysis a NAS of 1–4 was categorized as NAFL and $\text{NAS} \geq 5$ was categorized as NASH for the present study. 363 patients were eligible for the retrospective study and detailed demographic and clinical information of this cohort is given in Table 1.

Table 1
Overview of demographic and clinical data of analyzed patients with NAFLD.

Parameter	All NAFLD patients n = 363	NAFL (n = 106 by SAF)	NASH (n = 257 by SAF)
Sex [w / m]	270 / 93	83 / 23	187 / 70
Age	45.5 ± 10.3	41.0 ± 11.1	43.2 ± 9.9
BMI [kg/m ²]	52 ± 8.5 kg/m ²	51.2 ± 8.8 kg/m ²	52.3 ± 8.4 kg/m ²
ALT [U/l]	37.2 ± 26.4	28.7 ± 18.6	40.6 ± 28.3
AST [U/l]	31.3 ± 19.0	24.7 ± 9.7	33.9 ± 21.2
γGT [U/l]	47.1 ± 72.4	34.0 ± 30.9	52.4 ± 83.2
bilirubin	0.54 ± 0.25	0.57 ± 0.29	0.53 ± 0.24
INR	0.984 ± 0.110	0.996 ± 0.145	0.980 ± 0.092
HbA1c	6.18 ± 1.43	6.00 ± 1.39	6.26 ± 1.45
Total cholesterol	201.7 ± 36.8	197.9 ± 35.8	203.2 ± 37.1
LDL	133.9 ± 34.0	133.4 ± 28.4	134.1 ± 36.0
HDL	47.0 ± 12.0	48.1 ± 13.1	46.5 ± 11.6

Statistical analysis

Statistical data analyses were performed with Prism Version 7 (Graphpad Inc., La Jolla, CA, USA). All data are presented as mean ± standard error of the mean (SEM) unless specified otherwise. Correlation analysis was performed using Spearman's rank correlation coefficient. To assess diagnostic capability of serum parameters ROC curves for separation of NAFL vs. NASH were constructed.

Declarations

Acknowledgements

Prof. Hideo A. Baba, Dorothe Möllmann und Martin Schlattjan (Institute for Pathology, University Hospital Essen), and Prof. Johannes Haybäck (Institute for Pathology, University Hospital Magdeburg) for histological preparation and assessments. We also thank Prof. Niedergethmann (Department for General- and Visceral Surgery, Alfried Krupp Hospital, Essen, Germany) and Prof. Hasenberg (Helios Hospital Niederberg) for sample collection.

Author Contributions

Study concept and design: AC, JPS

Analysis and interpretation of data: PL, JPS

Drafting of the manuscript: PL, NS

Critical revision of the manuscript for important intellectual content: TB, MPO, DH, AC, JPS

Statistical analysis: JPS, DH

Obtained funding: AC

Study supervision: AC

All authors approved the final draft submitted.

Additional Information

Funding: This work was supported by the German research foundation (DFG CA267/14-1) and the Wilhelm Laupitz Foundation (AC).

Competing Interests Statement: All authors state that there are no conflicts of interest to declare.

Ethics approval and patient consent: The study protocol conformed to the ethical guidelines of 1975 Declaration of Helsinki and was approved by the Institutional Review Board (IRB; Ethik-Kommission der Medizinischen Fakultät der Universität Duisburg-Essen; Germany; 15-6356-BO). Due to the retrospective nature of the data analysis the IRB waived requirement of written informed consent. All procedures adhered to the Declaration of Helsinki and the requirements of the IRB.

At the time of IRB approval, TB, AC and JPS were affiliated to the University Hospital Essen, Department of Gastroenterology and Hepatology. Patient recruitment / data collection for the present retrospective analysis was performed at this department and the affiliated Alfried-Krupp Krankenhaus in Essen, Germany.

Data availability: Raw data are available upon reasonable request to the corresponding author.

References

1. Younossi, Z. M. *et al.* Global epidemiology of nonalcoholic fatty liver disease—Meta-analytic assessment of prevalence, incidence, and outcomes. *Hepatology* **64**, 73–84 (2016).
2. Balmer, M. L. & Dufour, J. F. Nicht-alkoholische steatohepatitis - von NAFLD zu MAFLD. *Ther. Umschau* **68**, 183–188 (2011).
3. Younossi, Z. M. *et al.* From NAFLD to MAFLD: Implications of a premature change in terminology. *Hepatology* (2020). doi:10.1002/hep.31420

4. Rosato, V. *et al.* NAFLD and extra-hepatic comorbidities: Current evidence on a multi-organ metabolic syndrome. *International Journal of Environmental Research and Public Health* **16**, (2019).
5. Targher, G., Day, C. P. & Bonora, E. Risk of cardiovascular disease in patients with nonalcoholic fatty liver disease. *New England Journal of Medicine* **363**, 1341–1350 (2010).
6. Mantovani, A. *et al.* Complications, morbidity and mortality of nonalcoholic fatty liver disease. *Metabolism*. **111**, (2020).
7. Canbay, A. *et al.* Patterns and predictors of mortality and disease progression among patients with non-alcoholic fatty liver disease. *Aliment. Pharmacol. Ther.* **52**, 1185–1194 (2020).
8. Younossi, Z. M., Henry, L., Bush, H. & Mishra, A. Clinical and Economic Burden of Nonalcoholic Fatty Liver Disease and Nonalcoholic Steatohepatitis. *Clinics in Liver Disease* **22**, 1–10 (2018).
9. Ekstedt, M. *et al.* Long-term follow-up of patients with NAFLD and elevated liver enzymes. *Hepatology* **44**, 865–873 (2006).
10. Singh, S. *et al.* Fibrosis Progression in Nonalcoholic Fatty Liver vs Nonalcoholic Steatohepatitis: A Systematic Review and Meta-analysis of Paired-Biopsy Studies. *Clinical Gastroenterology and Hepatology* **13**, 643-654.e9 (2015).
11. Long, M. T. *et al.* Hepatic fibrosis associates with multiple cardiometabolic disease risk factors: The Framingham Heart Study. *Hepatology* hep.31608 (2020). doi:10.1002/hep.31608
12. Gottlieb, A. *et al.* Lipoprotein and Metabolic Profiles Indicate Similar Cardiovascular Risk of Liver Steatosis and NASH. *Digestion* (2020). doi:10.1159/000510600
13. Ertle, J. *et al.* Non-alcoholic fatty liver disease progresses to hepatocellular carcinoma in the absence of apparent cirrhosis. *Int. J. Cancer* **128**, 2436–2443 (2011).
14. Best, J. *et al.* GALAD Score Detects Early Hepatocellular Carcinoma in an International Cohort of Patients With Nonalcoholic Steatohepatitis. *Clin. Gastroenterol. Hepatol.* **18**, 728-735.e4 (2020).
15. Kleiner, D. E. *et al.* Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* **41**, 1313–1321 (2005).
16. Kälsch, J. *et al.* Normal liver enzymes are correlated with severity of metabolic syndrome in a large population based cohort. *Sci. Rep.* **5**, (2015).
17. Bedossa, P. *et al.* Histopathological algorithm and scoring system for evaluation of liver lesions in morbidly obese patients. *Hepatology* **56**, 1751–1759 (2012).
18. Lindor, K. D. *et al.* The role of ultrasonography and automatic-needle biopsy in outpatient percutaneous liver biopsy. *Hepatology* **23**, 1079–1083 (1996).
19. Ratziu, V. *et al.* Sampling variability of liver biopsy in nonalcoholic fatty liver disease. *Gastroenterology* **128**, 1898–1906 (2005).
20. Armstrong, M. J. & Marchesini, G. Referral pathways for NAFLD fibrosis in primary care – No longer a ‘needle in a haystack’. *Journal of Hepatology* **71**, 246–248 (2019).
21. Blais, P. *et al.* Nonalcoholic fatty liver disease is underrecognized in the primary care setting. *Am. J. Gastroenterol.* **110**, 10–14 (2015).

22. Mofrad, P. *et al.* Clinical and histologic spectrum of nonalcoholic fatty liver disease associated with normal ALT values. *Hepatology* **37**, 1286–1292 (2003).
23. Gholam, P. M., Flancbaum, L., MacHan, J. T., Charney, D. A. & Kotler, D. P. Nonalcoholic fatty liver disease in severely obese subjects. *Am. J. Gastroenterol.* **102**, 399–408 (2007).
24. Younossi, Z. M. *et al.* The global epidemiology of NAFLD and NASH in patients with type 2 diabetes: A systematic review and meta-analysis. *J. Hepatol.* **71**, 793–801 (2019).
25. Canbay, A. *et al.* Non-invasive assessment of NAFLD as systemic disease—A machine learning perspective. *PLoS ONE* **14**, (2019).
26. Neumann, U. *et al.* Compensation of feature selection biases accompanied with improved predictive performance for binary classification by using a novel ensemble feature selection approach. *BioData Min.* **9**, 1–14 (2016).
27. Zhang, P. Determination of the upper cut-off values of serum alanine aminotransferase and aspartate aminotransferase in Chinese. *World J. Gastroenterol.* **21**, 2419 (2015).
28. Ruhl, C. E. & Everhart, J. E. Upper limits of normal for alanine aminotransferase activity in the United States population. *Hepatology* **55**, 447–454 (2012).
29. Wu, W.-C. *et al.* Updated thresholds for serum alanine aminotransferase level in a large-scale population study composed of 34 346 subjects. *Aliment. Pharmacol. Ther.* **36**, 560–568 (2012).
30. Kumar, S., Amarapurkar, A. & Amarapurkar, D. Serum aminotransferase levels in healthy population from western India. *Indian J. Med. Res.* **138**, 894–899 (2013).
31. Terrault, N. A. *et al.* Update on prevention, diagnosis, and treatment of chronic hepatitis B: AASLD 2018 hepatitis B guidance. *Hepatology* **67**, 1560–1599 (2018).
32. Hannah, W. N., Torres, D. M. & Harrison, S. A. Nonalcoholic Steatohepatitis and Endpoints in Clinical Trials. *Gastroenterol. Hepatol. (N. Y.)* **12**, 756–763 (2016).
33. Rastogi, A. *et al.* Non-alcoholic fatty liver disease – histological scoring systems: a large cohort single-center, evaluation study. *APMIS* **125**, 962–973 (2017).
34. Brunt, E. M. Nonalcoholic fatty liver disease and the ongoing role of liver biopsy evaluation. *Hepatol. Commun.* **1**, 370–378 (2017).
35. Ekstedt, M. *et al.* Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up. *Hepatology* **61**, 1547–1554 (2015).
36. Wong, V. W. S. *et al.* Disease progression of non-alcoholic fatty liver disease: A prospective study with paired liver biopsies at 3 years. *Gut* **59**, 969–974 (2010).
37. McPherson, S. *et al.* Evidence of NAFLD progression from steatosis to fibrosing-steatohepatitis using paired biopsies: Implications for prognosis and clinical management. *J. Hepatol.* **62**, 1148–1155 (2015).
38. Dam-Larsen, S. *et al.* Final results of a long-term, clinical follow-up in fatty liver patients. *Scand. J. Gastroenterol.* **44**, 1236–1243 (2009).

39. Adams, L. A., Sanderson, S., Lindor, K. D. & Angulo, P. The histological course of nonalcoholic fatty liver disease: A longitudinal study of 103 patients with sequential liver biopsies. in *Journal of Hepatology* **42**, 132–138 (J Hepatol, 2005).
40. Kleiner, D. E. *et al.* Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* **41**, 1313–1321 (2005).

Figures

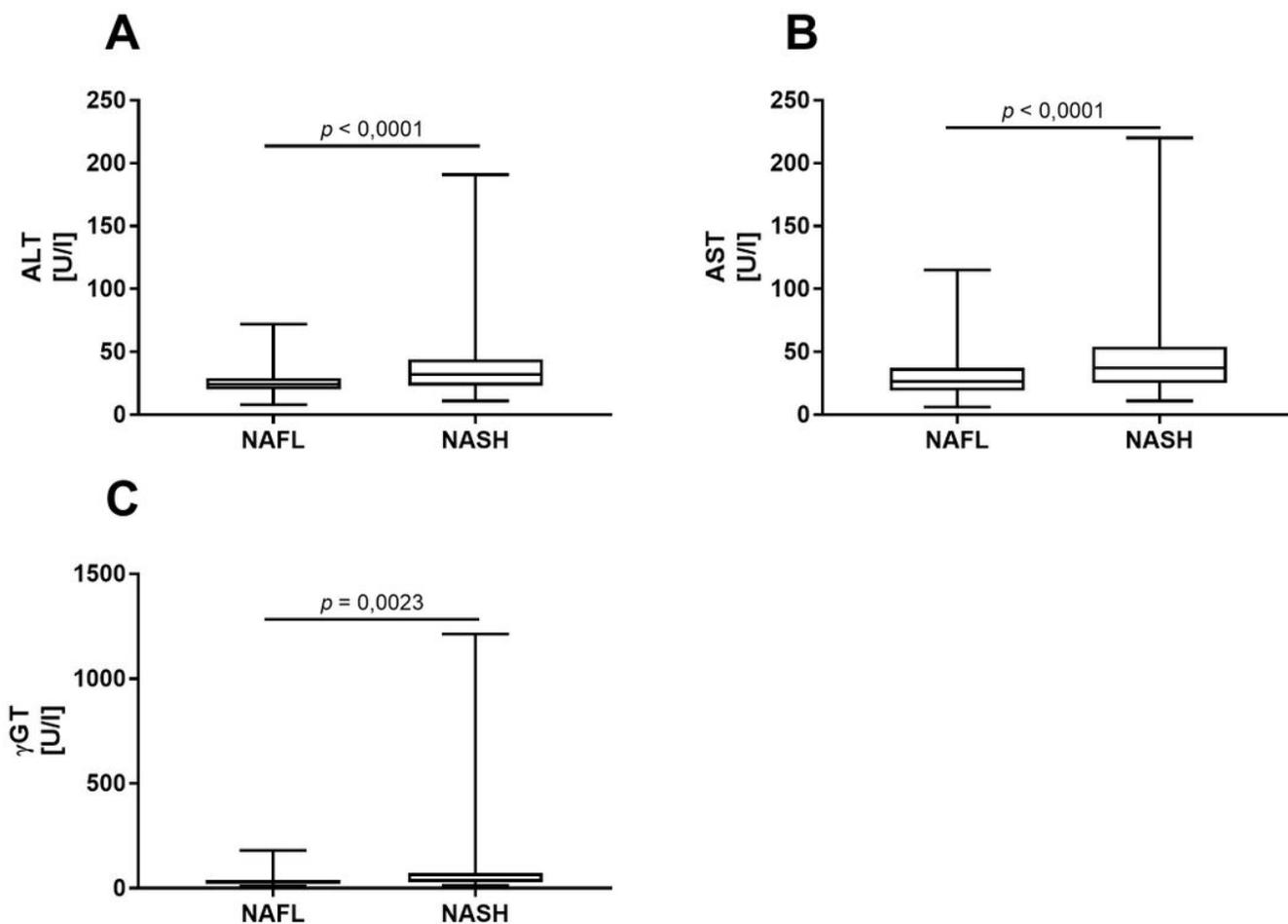


Figure 1

Serum liver enzyme concentrations differ between NAFL and NASH according to classification by NAS. When morbidly obese NAFLD patients were grouped by NAS (1-4 = NAFL; ≥ 5 = NASH), concentrations of all three liver enzymes, alanine-aminotransferase (ALT; A), aspartate-aminotransferase (AST; B) and gamma-glutamyltransferase (γ GT; C) differed significantly between NAFL and NASH. Statistical significance was tested by Mann-Whitney-U-Test (not normally distributed data) and significance was assumed at $p < 0.05$.

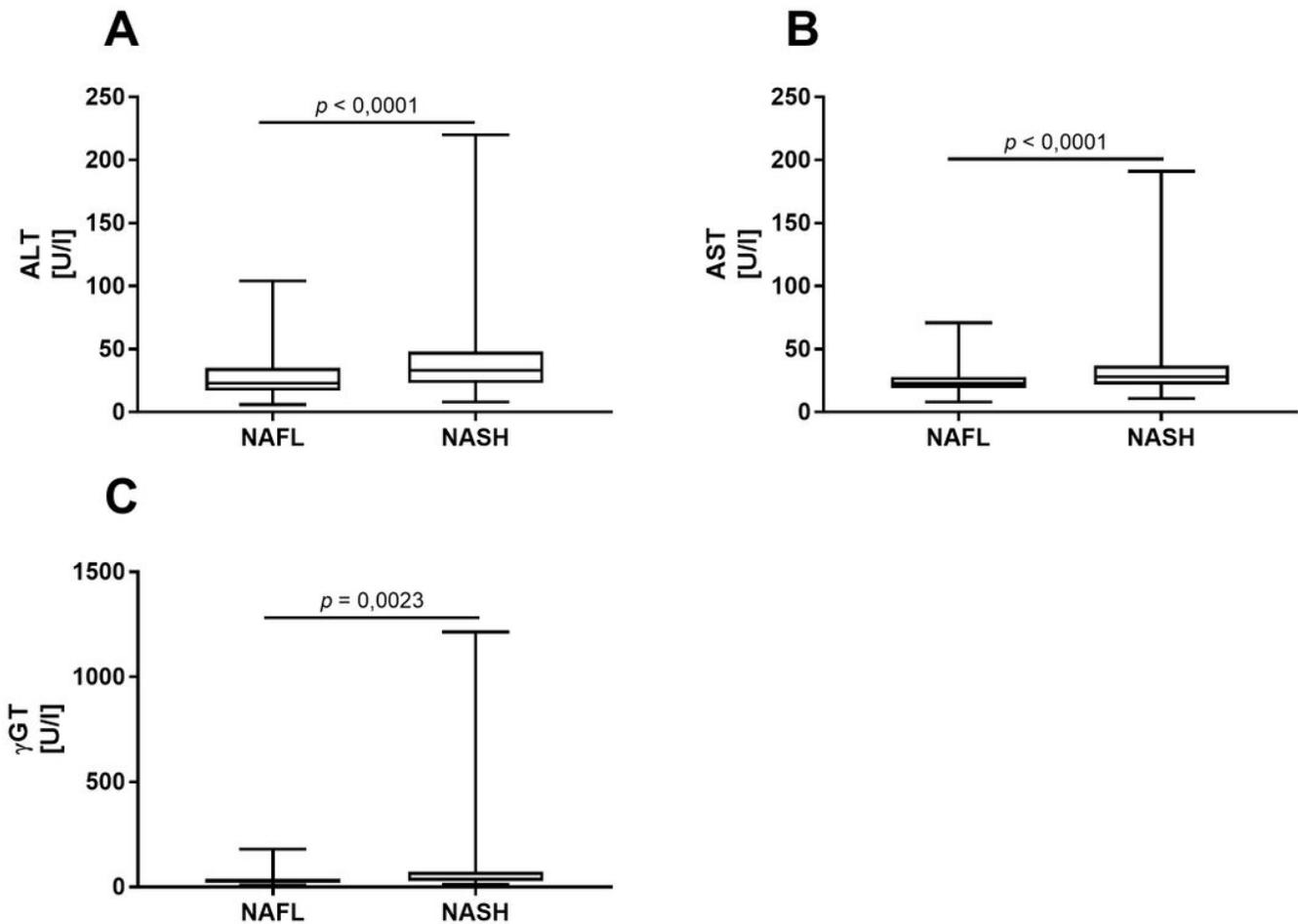


Figure 2

Serum liver enzyme concentrations differ between NAFL and NASH according to classification by SAF. Morbidly obese NAFLD patients were grouped as NAFL or NASH according to the SAF score (REF Bedossa). The serum concentrations of all three liver enzymes, alanine-aminotransferase (ALT; A), aspartate-aminotransferase (AST; B) and gamma-glutamyltransferase (γ GT; C) differed significantly between NAFL and NASH. Statistical significance was tested by Mann-Whitney-U-Test (not normally distributed data) and significance was assumed at $p < 0.05$.

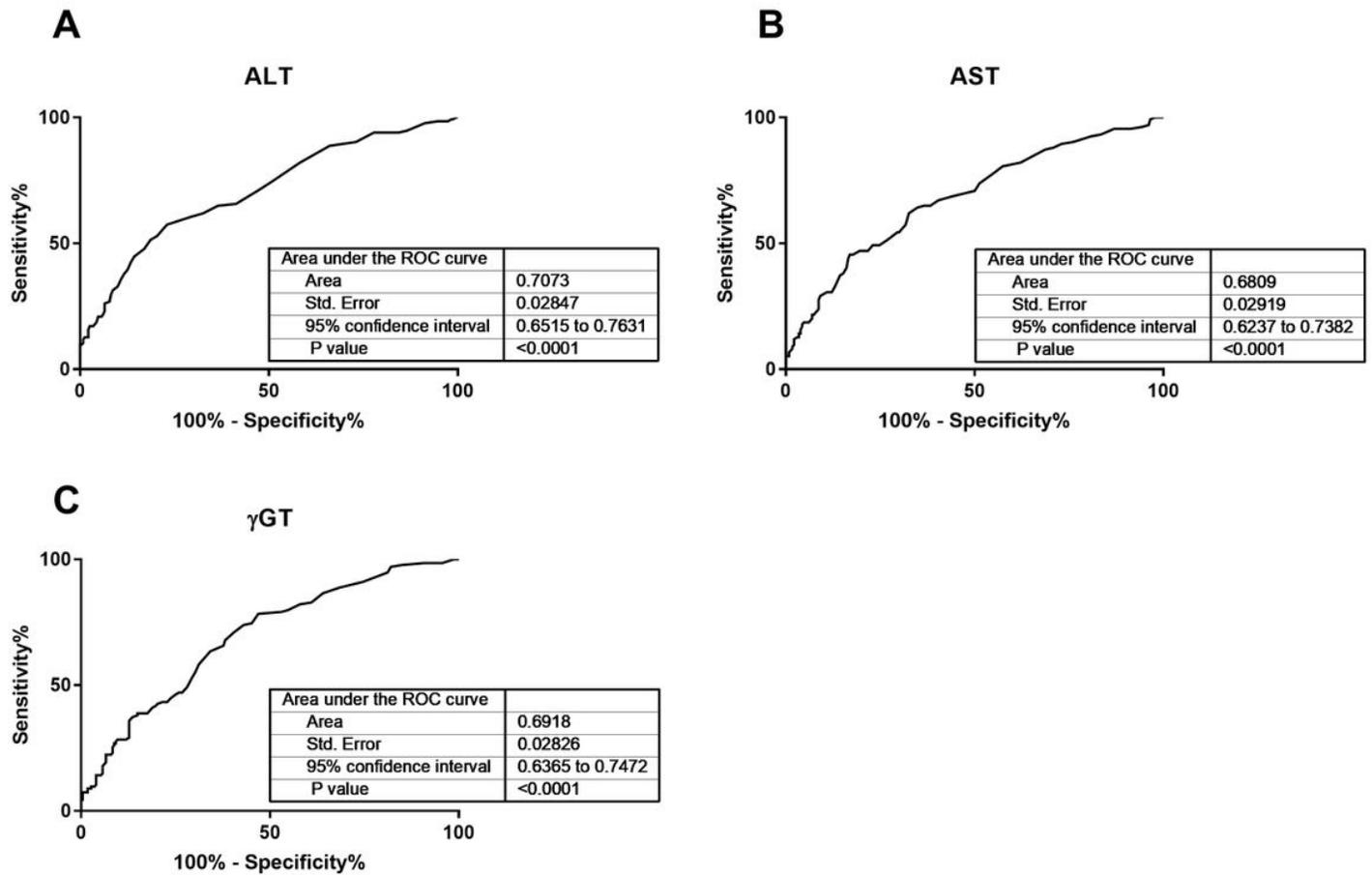


Figure 3

ROC curves for performance of serum liver enzyme concentrations to separate NAFL and NASH after NAS-based classification. For each serum parameter alanine-aminotransferase (ALT; A), aspartate-aminotransferase (AST; B) and gamma-glutamyltransferase (γ GT; C) ROC curves were calculated for separation of NAFL from NASH according to NAS (1-4 = NAFL; ≥ 5 = NASH). The area under the curve (AUC) was calculated for all curves resulting in poor to moderate discriminatory performance of 0.68 (AST) to 0.71 (ALT).

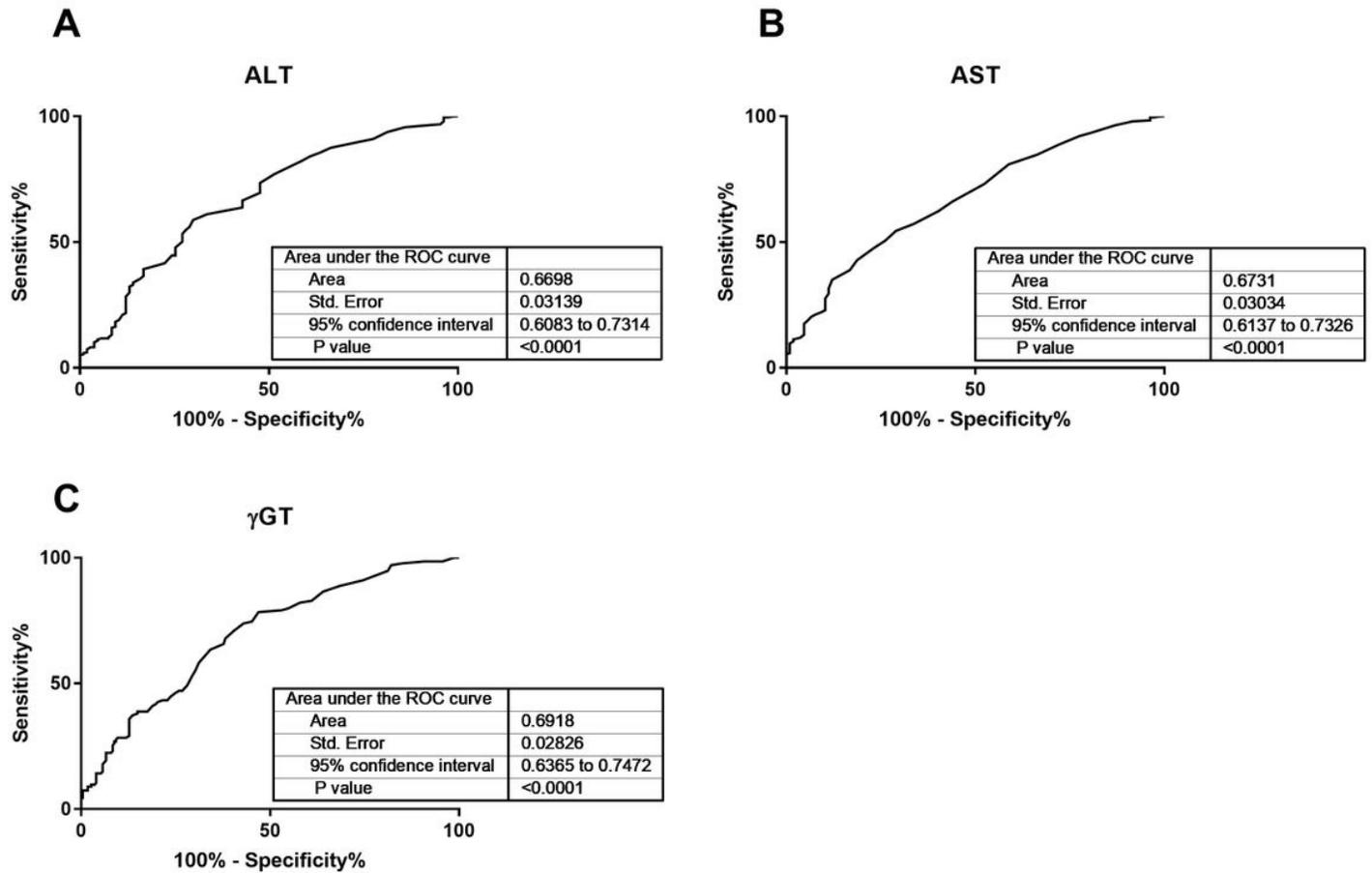


Figure 4

ROC curves for performance of serum liver enzyme concentrations to separate NAFL and NASH defined by SAF. For each serum parameter alanine-aminotransferase (ALT; A), aspartate-aminotransferase (AST; B) and gamma-glutamyltransferase (γ GT; C) ROC curves were calculated to separate NAFL from NASH based on SAF classification (REF Bedossa). The area under the curve (AUC) were calculated for all curves resulting in poor to moderate discriminatory performance of 0.67 (ALT) to 0.69 (ALT).