Risk Factor Prevalence and their Relative Influence on Fatty Liver and Gallstone Disease: A Cross-Ethnic Study Comparing Two High-Risk Populations from Chile and Northeast Germany

Peter J. Meffert, PhD 1*; Maximilian A. Schwittay, MS 1; Markus M. Lerch, MD, FRCP 2; Jochen Hampe, MD 3; Stephan Buch, PhD 3; Flavio Nervi, MD 4; Luis Villarroel, PhD 5; Henry Völzke, MD, PhD 1†; Juan F. Miquel, MD 4†

†These authors co-supervised this work and contributed equally

1Institute for Community Medicine, Ernst Moritz Arndt University of Greifswald, Greifswald, Germany
2Department of Medicine A, Ernst Moritz Arndt University of Greifswald, Greifswald, Germany
31st Medical Department, University Hospital Dresden, Technical University Dresden, Dresden, Germany
4Departamento de Gastroenterologia, Facultad de Medicina, Pontificia Universidad Catolica de Chile, Santiago, Chile
5División Salud Pública y Medicina Familiar, Facultad de Medicina, Pontificia Universidad Catolica de Chile, Santiago, Chile

ABSTRACT

Objective: Fatty Liver Disease (FLD) and Gallstone disease (GSD) are among the most common gastrointestinal disorders worldwide. Based on data of two ethnically distant populations from Santiago de Chile and Northeast Germany sharing remarkably high prevalence for both diseases, we aimed to estimate (1) the standardised prevalence of FLD and GSD, (2) the population-specific prevalence of known risk factors and (3) the relative influence of each risk factor for the occurrence of the respective outcome.

Design: We used data from two population-based studies from Chile (n=742) and Germany (n=3632). Regression methods were applied to assess the prevalence of risk factors at a specific age. To quantify the relative contribution of the risk factors to the occurrence of FLD and GSD, we used hierarchical variation partitioning.

Results: The standardised prevalence of FLD ranged from 20.7% in Chilean men to 37.9% in German men and the standardised prevalence of GSD from 14.5% in German men to 40.5% in Chilean women. Risk profiles differed considerably between populations and sexes. Overall, the contributions of the risk factors to the occurrence of FLD and GSD coincided widely. For FLD, age and BMI were the predominant factors, followed by serum triglyceride levels and the Single-Nucleotide Polymorphism (SNP) rs738409 C>G of the PNPLA3 gene. In GSD age was the predominant determinant, in women additionally BMI. Alcohol consumption played a marked role in Chilean men, being positively associated with GSD, similarly the use of contraceptives in Chilean women. The relative contribution of the lithogenic SNP rs11887543 was low in both populations, with marginal effect in Chileans.

Conclusion: Despite marked differences in risk profiles, relative influences of risk factors coincided widely between study populations. Knowledge about importance of risk factors for fatty liver and gallstone disease may guide therapy and advises for lifestyle changes.

KEYWORDS: PNPLA3; ABCG8; Epidemiology; Hierarchical variation partitioning.
INTRODUCTION

Gallstone disease (GSD) is one of the most common gastrointestinal disorders, affecting 10-20% of the population in Western countries and causing high costs in the health-care systems, e.g. in the USA about 6.5 billion Dollars annually.1,2 Fatty Liver Disease (FLD) is even more common reaching a prevalence of up to 45% in the general population3,4 and causing substantial health-care costs.5 GSD and FLD are showing a rising prevalence since many decades.6,7

Both FLD and GSD are multifactorial disorders. In FLD, immutable risk factors are higher age, male sex and genetic risk factors such as the SNP rs738409 (variant p.I148M) of the enzyme adiponutrin (PNPLA3).8,9 Mutable risk factors include somatometric measurements such as BMI, nutrition, physical activity, alcohol consumption and drugs (e.g. anovulants, steroids, glucocorticoids) as well as biomarkers such as serum triglyceride levels and diabetes type 2. Non-modifiable risk factors for GSD include advancing age, being female, Amerindian ancestry and carrying common genetic variants such as the SNP rs11887543 (ABCG8 p.D19H variant).10 Among the modifiable risk factors are diabetes type 2, obesity, rapid weight loss, diet, hemolysis, drugs and physical activity.11,12,13 Women are in general of higher risk than men,14 a phenomenon possibly connected with pregnancy and higher oestrogen levels.15

Our study populations were a Hispanic population from the commune La Florida in the Santiago Metropolitan Region in Chile and a Caucasian from Western Pomerania in Northeast Germany. Chileans and Western Pomeranians are among the populations with the highest prevalence of GSD and FLD worldwide.3,16 The two populations did not differ significantly in the frequency of carrying risk allele for GSD identified so far.17,18 In this study we aimed to assess standardised prevalence of GSD and FLD, prevalence of risk factors and ascertain the specific contributions of these risk factors in the two populations to the development of FLD and GSD. Consequently, our research questions are: (1) How do prevalence of FLD and GSD differ between the two populations? (2) Are differences in prevalence explained by differences in risk profiles or rather by different susceptibilities for certain risk factors?

METHODS

Populations and Study Designs

The data from the Chilean study population originate from the first follow-up examination of an epidemiological study in an urban area of south Santiago, La Florida, which is representative of the predominant socioeconomic strata and European-Amerindian admixture of the Hispanic Chilean population with 328,881 inhabitants in 1992. Blocks were randomly selected, and then all individuals residing there were enumerated. In the baseline phase, of 2558 selected individuals, 1678 agreed to participate, a response of 65.6%. The current data of 964 subjects aged from 28-89 originate from the first follow-up examination that has been conducted between 2000 and 2001 which corresponds to 57.4% of the initial sample.19,20 Of these, 63 datasets were incomplete, primarily in genotyping (n=159). Thus, we used 742 complete datasets for the present analyses. For prevalence estimates we used 911 complete datasets.

The German data come from the baseline examination of the Study of Health in Pomerania (SHIP)30 that includes 4308 subjects. From the total population of Western Pomerania comprising 213,057 inhabitants in 1996, a two-stage stratified random cluster sample of adults aged 20-79 years was drawn. The net sample comprised 6265 eligible subjects. Of these, 4308 subjects (n=2192 women; response of 68.8%) participated in the SHIP baseline examination between 1997 and 2001. In 449 datasets data in the variables of interest were missing, primarily in genotyping (n=252) and alcohol consumption (n=207), thus, the analytic sample comprised of 3632 individuals. For prevalence estimates we used 4124 complete datasets. All patients and controls gave written informed consent prior to the study, and all study protocols have been approved by the Institutional Review and Ethics Committees at the respective sites.

Data Collection and Definitions

The examination protocols included a pre-coded questionnaire on socioeconomic data and medical history, physical and abdominal ultrasonographic examinations (Chile: 3.5-MHz linear transducer, Toosbee, Toshiba, Japan; Germany: 5 MHz transducer and a high-resolution instrument, Vingmed VST Gateway) and biomarker analyses. In both studies, body mass index (BMI) was calculated as kg/m². Alcohol consumption was measured as the mean intake of alcohol over the last 30 days, calculated from interview answers. Serum triglycerides and serum glucose were measured with a fasting time of more than eight hours in Chile; in Germany, the median fasting time was 3.5 hours. Diabetes was defined as a pathological serum glucose value of 11.1 mmol/L or higher, antidiabetic medication and/or a self-reported physician’s diagnosis. Parity was defined as ever being pregnant or not, also postmenopausal hormone intake and contraceptive intake was defined as ‘yes’ or ‘no’.

FLD was diagnosed after visualisation of a hyperechogenic liver pattern compared to renal parenchyma. Indistinctness of hepatic blood vessels or the diaphragm, decreased acoustic penetration and hepatomegaly were taken into the diagnosis as additional criteria. We did not separate non-alcoholic fatty liver disease (NAFLD) from alcoholic fatty liver disease (AFLD) because we were explicitly interested in the contribution of alcohol consumption to the occurrence of FLD. ABLD and NAFLD share similar histopathological patterns and are only caused by different risk factors.21 Moreover, the definition of ABLD does depend on the arbitrarily chosen cut-off value for daily alcohol consumption. Based on ultrasound in both studies, GSD was defined as the presence of gallstones or a previous cholecystectomy.19,20 The SNP rs738409 (common risk variant p.I148M) of the enzyme adiponutrin encoded in the PNPLA3 gene is a major risk factor for FLD,22 similarly the SNP rs11887543 (common
risk variant p.D19H) of the ABCG8 gene for GSD.\textsuperscript{11} Genotyping of these variants was carried out using the TaqMan SNP genotyping allelic discrimination method (Applied Biosystems, Foster City, CA, USA).

Statistical Analyses

Prevalence data were converted to the world population by direct standardisation by age and sex\textsuperscript{23} using the ‘dstdize’ command in Stata (Stata Corp., College Station, TX, USA). We decided in favour of sex-specific modelling because risk profiles differed substantially between men and women for both phenotypes. Additionally, pathomechanisms are not clear enough to rule out or assume specific sex-specific interactions. Models for FLD were adjusted for age, BMI, alcohol consumption, triglycerides and glucose serum levels, diabetes and SNP rs738409. Models for GSD were adjusted in the same manner but with the SNP rs11887543 and in women additionally for parity, contraceptives and postmenopausal hormones. To quantify the contribution of each risk factor to the phenotypes, we used hierarchical variation partitioning\textsuperscript{24} implemented in the package ‘hier.part’\textsuperscript{25} in the software package R (R Foundation, Vienna, Austria).\textsuperscript{26} This method calculates the model fit with all possible combinations of predictor variables and gives the proportion of explained variance for each variable that is explained solely by this variable. We applied the log-likelihood as goodness-of-fit measurement. To account for possible non-linear relations between the predictors and the outcome, we used fractional polynomials.\textsuperscript{27} If a two-degree polynomial was fitted for the relationship, we summed up their relative influences resulting from the variation partitioning. To quantify the prevalence of risk factors, we applied median regression for each risk factor and adjusted for age, because median regression does not rely on certain distributional assumptions as e.g. linear (mean) regression. Afterwards, we calculated the respective marginal mean at an age of 50 years to account for changes with age. All models for quantitative outcomes were calculated with robust standard errors.

RESULTS

The analyses included 3632 individuals from Germany (51.1% women) and 742 from Chile (64.6% women). Mean age was 49 years for both samples (standard deviation for Chile: 13 years, SHIP: 16 years).

Prevalence of Fatty Liver Disease, Gallstone Diseases and their Risk Factors

Crude and standardised prevalence data for FLD and GSD are given in Table 1. FLD was much more frequent in German men (37.9%) compared to Chilean men (20.7%). In women, Chileans had a similar prevalence (23.0%) to Germans (21.8%). In contrast, the frequency of GSD was similar between men (Chile: 15.1%, Germany: 14.5%) but differed clearly among women with a much higher proportion in Chileans (40.5%) compared to Germans (26.8%).

Further, we estimated the prevalence of risk factors for FLD and GSD. To enable comparability between the two study groups, we performed median regressions adjusted by age and calculated marginal means at an age of 50 years. Median values of risk factors are given in Table 2. Median alcohol consumption was in general much lower in women than in men and considerably higher in the German compared to the Chilean population. BMI differed significantly only in women with Chilean having proportions comparably more frequent in Chile. Similarly but less pronounced the risk G allele frequency of the genetic variant rs738409 at the PNPLA3 gene was strikingly more frequent in Chile. The risk G allele frequency of the genetic variant rs11887534 at the ABCG8 gene was more frequent in the Chilean population. Parity was marginally higher in Chilean women whereas intake of postmenopausal hormones was more frequent in Germany; there was no significant difference in the use of contraceptives.

Relative Influence of Risk Factors on Gallstone Disease and Fatty Liver Disease

We performed regression analyses and then calculated the contribution of each risk factor on the occurrence of FLD and GSD by hierarchical variation partitioning. As a result, this

| Sex/cohort        | Fatty liver disease |                                  |                                | Gallstone disease |                                  |
|-------------------|---------------------|----------------------------------|--------------------------------|-------------------|----------------------------------|
|                   | N                   | Crude Percentage | Standardised Percentage | 95% CI            | Crude Percentage | Standardised Percentage | 95% CI          |
| Men               |                     |                     |                             |                   |                     |                             |                  |
| La Florida (Chile)| 311                 | 25.7%               | 20.7%                        | 16.5%             | 24.9%               | 14.8%                        | 15.1% | 11.1% | 19.0% |
| Northeast Germany | 2011                | 37.3%               | 37.9%                        | 35.8%             | 39.9%               | 14.8%                        | 14.5% | 13.0% | 15.9% |
| Women             |                     |                     |                             |                   |                     |                             |                  |
| La Florida (Chile)| 600                 | 24.5%               | 23.0%                        | 19.6%             | 26.3%               | 41.3%                        | 40.5% | 36.8% | 44.2% |
| Northeast Germany | 2113                | 21.9%               | 21.8%                        | 20.2%             | 23.5%               | 27.1%                        | 26.8% | 25.1% | 28.6% |

Table 1: Prevalence of fatty liver and gallstone disease for the two cohorts, given is crude and standardised prevalence for the world population.
method gives the independent influence of each variable.

Logistic regression models for the occurrence of FLD and GSD comprised age, BMI, alcohol consumption, serum triglycerides, serum glucose, and diabetes type 2, in FLD additionally the SNP rs738409 and in GSD additionally the SNP rs11887534 as well as parity, contraceptives and postmenopausal hormones in women. For non-linear associations, dependent variables were transformed using fractional polynomials. The regression tables are given in the Appendix. The regression models for FLD had AUC (area under the Receiver-Operation-Characteristic curve) values ranging from 0.79 to 0.79 (Chilean men: AUC=0.742, women: AUC=0.741; German men: AUC=0.772, women: AUC=0.789), pseudo-R² values were lower compared to FLD models (Chilean men: pseudo R²=0.102, women: pseudo R²=0.133; German men: pseudo R²=0.148, women: pseudo R²=0.190). In German men only age and BMI had a relative influence above 10%, whereas in Chilean men age and alcohol consumption contributed considerably to the occurrence of GSD (Figure 2A). Serum triglycerides, serum glucose and diabetes played only a marginal role, in Chile BMI was also not an important factor. While in German men alcohol was non-significant inversely related to GSD (OR=0.993, 95% CI: 0.983-1.002, p=0.019, Table A 5 of the Appendix) being borderline statistically significant. In women both age and BMI contributed substantially to the occurrence of GSD, in Chiles additionally use of contraceptives (Figure 2 B). In German women, the C allele of the SNP rs11887534 contributed with 11.9% in Chilean men, 39.3% in German men, 31.1% in Chilean women and 28.8% in German women.

In GSD models, AUC ranged from 0.74 to 0.79 (Chilean men: AUC=0.742, women: AUC=0.741; German men: AUC=0.772, women: AUC=0.789), pseudo-R² values were lower compared to FLD models (Chilean men: pseudo R²=0.102, women: pseudo R²=0.133; German men: pseudo R²=0.148, women: pseudo R²=0.190). In German men only age and BMI had a relative influence above 10%, whereas in Chilean men age and alcohol consumption contributed considerably to the occurrence of GSD (Figure 2A). Serum triglycerides, serum glucose and diabetes played only a marginal role, in Chile BMI was also not an important factor. While in German men alcohol was non-significant inversely related to GSD (OR=0.993, 95% CI: 0.983-1.002, p=0.019, Table A 5 of the Appendix) being borderline statistically significant. In women both age and BMI contributed substantially to the occurrence of GSD, in Chiles additionally use of contraceptives (Figure 2 B). In German women, the C allele of the SNP rs11887534 contributed with 11.9% in Chilean men, 39.3% in German men, 31.1% in Chilean women and 28.8% in German women.

DISCUSSION
We compared the prevalence of FLD and GSD between two ethnically different populations from Chile and Germany. Additionally, we assessed the occurrence of known risk factors for both disorders. Finally, we calculated the relative contribution of each risk factor to the occurrence of FLD and GSD.

| Risk factor                          | La Florida | Northeast Germany | p value |
|-------------------------------------|-----------|-------------------|---------|
| Alcohol (g/d)                       | 1.0       | 0.1               | <0.001  |
| BMI (kg/m²)                         | 27.5      | 28.8              | <0.001  |
| Diabetes (yes)                      | 9.4%      | 8.7%              | 0.442   |
| Serum triglycerides (mmol/l)        | 1.50      | 1.35              | <0.001  |
| Serum glucose (mmol/l)              | 5.06      | 4.95              | <0.001  |
| SNP rs738409 (number of G allele copies) | 1.03     | 1.10              | <0.001  |
| SNP rs11887534 (number of C allele copies) | 0.16     | 0.19              | <0.001  |
| Parity                              | –         | 96.8%             | <0.001  |
| Postmenopausal hormones             | –         | 11.5%             | <0.001  |
| Contraceptives                      | –         | 50.6%             | <0.001  |

Table 2: Frequencies of risk factors for gallstone disease and fatty liver disease in La Florida (Chile) and Northeast Germany; values are margins at an age of 50 years (except SNP rs738409 and rs11887534; contraceptives are given for an age of 25 years) of median regressions for continuous outcomes (given are medians), logistic regression for binary outcomes (given are probabilities) and Poisson regression for number of G allele copies of rs738409 and rs11887534, respectively.
Fatty Liver Disease

In our populations, FLD had a standardised prevalence of 20.8% to 23.1% with the exception of men from the SHIP that showed even a 37.5% prevalence (Table 1). Standardisation to the world population decreased the prevalence in Chile by five percent points. Worldwide prevalence of NAFLD, which make up the vast majority of FLD, has an assumed median of 20%. Population-based data from South America are scarce; in Mexico the prevalence of NAFLD was 17.1%, and 23% in the same Chilean population we analysed for the current analysis. Regarding the exclusion of individuals with high levels of alcohol consumption for NAFLD in these reports, this might be of about the same prevalence as our findings. Two southwest German cohorts showed a lower prevalence of 24.0% and 18.7% (men and women combined) compared to our data. In the United States, the prevalence were of similar magnitude as in Western Pomerania: the Dallas Heart Study found 33.6% FLD measured by 1H magnetic resonance spectrography, in another population-based study the prevalence of NAFLD was 21.4%. The estimated prevalence highly depends on diagnostic tools (ultrasound, tomography, enzymes, biopsy, autopsy), on age and sex distributions and study design. The high prevalence in German men is plausible in the light of the high alcohol consumption, the high prevalence of diabetes and the high serum triglyceride levels (Table 2).

The relative contribution of each risk factor to the occurrence of FLD varied between the Chilean and the German population as well as between men and women. The most obvious difference between the study populations was the influence of age and BMI. This could possibly be attributed to the fact that elder Germans were on average more obese than younger ones, whereas elder Chileans have on average a lower BMI compared to middle-aged subjects. Here, the cross-sectional nature of the data has to be kept in mind: possibly, a change in lifestyle had occurred only in younger Chileans but not in elder ones. In consequence, age and BMI should have to be considered together rather than as single factors. In fact, the sum of the relative influences of age and BMI is comparable for Chile and the SHIP. Alternatively, Chileans may develop FLD and obesity earlier in life compared to Germans or independent of age because of life styles or genetic factors such as rs738409 or others to be discovered.

The SNP rs738409 had an influence of about 10% in SHIP and Chilean men despite the prevalence of the risk G allele was more than 130% higher in the Chilean compared to the German population in men and women. However, a major finding of this cross-ethnic study is that in Chilean women the relative contribution of the PNPLA3 risk variant was above 30%. Inversely, serum triglycerides accounted for about 20% of the variation of FLD in men and women from the SHIP but only for 5% in Chilean women. Possibly, PNPLA3 gene is very closely connected with triglyceride storage and might be an explanation for the low impact of serum triglycerides in Chilean women. For the Chilean data but not for the German data, we also confirm previous findings that the influence of the SNP rs738409 on liver fat content is stronger in women than in men. An association of the rs738409 G allele with lower fasting serum triglycerides in
overweight and obese individuals was found in two genetic studies and a cohort in northern Europe. Conceivably, the G allele of the rs738409 variant could cause an impaired hepatic triglyceride efflux resulting in lower serum triglycerides via a reduction of the very low density lipoprotein (VLDL) lipidation. Finally, the risk G allele of the PNPLA3 gene has been reported not only to be strongly associated with altered lipometabolism, but with chronic liver disease severity and progression. Since heritability of NAFLD has been estimated to be as high as 50% in some populations, this single PNPLA3 risk variant may explain about 60% of the heritability of fatty liver disease in Chilean women. Thus, at least in Chilean women, the evaluation of this risk variant may be a cornerstone in implementation of early detection and preventive actions for chronic liver diseases.

The influences of alcohol consumption, serum glucose and diabetes on FLD were low in both populations and sexes.

**Gallstone Disease**

The standardised prevalence of GSD in Chile was 15.1% for men and 40.5% for women, very similar to the original data reported for the entire cohort and one of the highest prevalence reported worldwide. The mean prevalence of GSD in European populations is 10-12%, a large study from Italy shows a prevalence of 8.5% for men and 18.9% for women. As reported earlier, Northeast Germany is a high-prevalence area for GSD. In contrast, two studies from Denmark report a much lower prevalence (men: 5.7%, women: 11.9% and men: 6.0%, women: 13.1%, respectively). Possibly, these dissimilarities are caused by considerable differences in BMI between Denmark and Germany. However, in the last-mentioned study, differences between known risk factors were only partly able to explain the difference in GSD prevalence between Denmark and Northeast Germany.

In GSD, age was the superior explanation in all groups, consistent with the chronic nature of this metabolic disease. In men, only alcohol consumption in Chileans had additionally a considerable relative influence with higher alcohol consumption being associated with a higher risk of GSD. Since German men drank much more alcohol, our data suggest that the amount of alcohol consumption is in Chileans more closely related to the occurrence of GSD. This could be attributed to genetic differences that cause differences in metabolism. A case-control study from Australia reported a protective effect of alcohol, but confounding cannot be ruled out for this finding. In women, the differences in the relative influence of BMI was higher for Chileans compared to Germans (22% versus 12%, respectively) and this trend is also observed in men, but the relative influence was lower. The differences in the relative contribution of serum triglycerides, serum glucose and diabetes were in general small.

Although the prevalence of risk factors for GSD differed substantially between German and Chilean women, their relative influences were comparable. Thus, what caused the extraordinary high prevalence in Chilean Women? Indeed, Chilean women had higher BMI, serum glucose levels and were more often pregnant than German women, but they also used less contraceptives and postmenopausal hormones. We assume that beyond BMI, an unmeasured factor contributed to the high prevalence in Chilean women. Model performances for GSD were worse for the Chilean data than for the German data, with smaller AUC and pseudo-\(R^2\) values. This could be indicative of one or several risk factors that occur only in the Chilean population. One possibility are specific not yet discover lithogenic genes, most probably associated to their Amerindian admixture. Indeed, we and others have proposed that heritability of GSD is significantly higher in Native Americans and admixture populations with Amerindian ancestry compared to other populations. The SNP rs11887534 indeed did explain some variation in GSD and had an odds ratio of about two in the German population, but it was able to explain only a comparatively small proportion of variance in the data. It is noteworthy that, although the frequency of this risk allele is significantly higher in Chileans, its effect on GSD risk was very small in Chileans compared to Germans. Possibly, other yet undiscovered genetic variants as well as gene-gene and gene-environment interactions may act beyond the impact of the single SNP, especially in Hispanics. BMI and serum triglycerides were slightly more important in the SHIP, whereas contraceptives had a stronger influence in Chile. Contraceptives may differ in their impact on GSD, and possibly, compounds used in Chile differ from those used in Germany; unfortunately data on the specific compounds were not available in our study. Moreover, the impact of contraceptives might be modified by different levels of BMI. The relative influence of BMI on GSD was in both populations much higher in women than in men. This corresponds to previous findings where obesity was associated with an increased risk of GSD in young women only.

**STRENGTHS AND LIMITATIONS**

The strength of our data is the population-based design for both cohorts and a comparable data acquisition in a relatively short period in between 1997 to 2001. Limiting factors are that we do not have comparable data on socioeconomic status, diet, family history and physical activity and are thus not able to address the influences of these factors.

**CONCLUSIONS**

Overall, the contributions of the risk factors to the occurrence of FLD and GSD coincided widely, but not for all variables. A particularly strong influence on the occurrence of FLD had age in Germans and BMI in Chileans, further the SNP rs738409 in Chilean women, whereas the influence of serum triglycerides was exceptionally weak in this group. For GSD, particularly strong influence had alcohol consumption in Chilean men and use of contraceptives in Chilean women.

**ACKNOWLEDGEMENTS**

The SHIP is part of the Community Medicine Research net of
the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ20103, 01ZZ0403 and 03IS2061A), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania. The Chilean study has been supported by Fondo Nacional Cientifico y Tecnológico (FONDECYT, grant no. 1130303 to JFM and no. 1100020 to FN) and the FONDAP Center for Genome Regulation (grant no. 15090007) to JFM. The authors thank the community of La Florida, without whose participation and encouragement these studies could not have been performed.

CONFLICTS OF INTEREST: None.

REFERENCES

1. Shaffer EA. Gallstone disease: Epidemiology of gallbladder stone disease. Best Pract Res Clin Gastroenterol. 2006; 20(6): 981-996. doi: 10.1016/j.bpg.2006.05.004

2. Lammert F, Sauerbruch T. Mechanisms of disease: the genetic epidemiology of gallbladder stones. Nat Clin Pract Gastroenterol Hepatol. 2005; 2(9): 423-433. doi: 10.1038/ncpgasthep0257

3. Völzke H, Baumeister SE, Alte D, et al. Independent risk factors for gallstone formation in a region with high cholelithiasis prevalence. Digestion. 2005; 71(2): 97-105. doi: 10.1159/000084525

4. Browning JD, Szczepaniak LS, Dobbins R, et al. Prevalence of hepatic steatosis in an urban population in the United States: a 5-year observation. Am J Gastroenterol. 2014; 24(10): 1404-1414. doi: 10.1038/ajg.2014.155

5. Meffert PJ, Baumeister SE, Lerch MM, Mayerle J, Kratzer W, Völzke H. Development, external validation, and comparative assessment of a new diagnostic score for hepatic steatosis. Hepatology. 2006; 40(6): 1387-1395. doi: 10.1002/hep.20466

6. Baumeister SE, Völzke H, Marschall P, et al. Impact of fatty liver disease on health care utilization and costs in a general population: a 5-year observation. Gastroenterology. 2008; 134(1): 85-94. doi: 10.1053/j.gastro.2007.10.024

7. Plant JC, Percy I, Bates T, Gastard J, de Nercy YH. Incidence of gallbladder disease in Canada, England, and France. Lancet. 1973; 2(7823): 249-251. doi: 10.1016/S0140-6736(73)93148-6

8. Bateson MC. Gallstones and cholecystectomy in modern Britain. Postgrad Med J. 2000; 76(901): 700-703. doi: 10.1136/pgmj.76.901.700

9. Welsh JA, Karpen S, Vos MB. Increasing prevalence of non-alcoholic fatty liver disease among United States adolescents, 1988-1994 to 2007-2010. J Pediatr. 2013; 162(3): 496-500. doi: 10.1016/j.jpeds.2012.08.043

10. Sookoian S, Pirola CJ. Meta-analysis of the influence of I148M variant of patatin-like phospholipase domain containing 3 gene (PNPLA3) on the susceptibility and histological severity of nonalcoholic fatty liver disease. Hepatology. 2011; 53(6): 1883-1894. doi: 10.1002/hep.24283

11. von Kampen O, Buch S, Nothnagel M, et al. Genetic and functional identification of the likely causative variant for cholesterol gallstone disease at the ABCG5/8 lithogenic locus. Hepatology. 2013; 57(6): 2407-2417. doi: 10.1002/hep.26009

12. Shaffer EA. Epidemiology and risk factors for gallstone disease: has the paradigm changed in the 21st century? Curr Gastroenterol Rep. 2005; 7(2): 132-140. doi: 10.1007/s11894-005-0051-8

13. Shaffer EA. Epidemiology of gallbladder stone disease. Best Pract Res Clin Gastroenterol. 2006; 20(6): 981-996. doi: 10.1016/j.bpg.2006.05.004

14. Aerts R, Penninckx F. The burden of gallstone disease in Europe. Aliment Pharmacol Ther. 2003; 3: 49-53. doi: 10.1046/j.0953-0673.2003.01721.x

15. Jorgensen T. Gall stones in a Danish population: fertility period, pregnancies, and exogenous female sex hormones. Gut. 1988; 29(4): 433-439. doi: 10.1136/gut.29.4.433

16. Miquel JF, Covarrubias C, Villaroel L, et al. Genetic epidemiology of cholesterol cholelithiasis among Chilean Hispanics, Amerindians, and Maoris. Gastroenterology. 1998; 115(4): 937-946

17. Buch S, Schafmayer C, Völzke H, et al. Loci from a genome-wide analysis of bilirubin levels are associated with gallstone risk and composition. Gastroenterology. 2010; 139(6): 1942-1951. doi: 10.1053/j.gastro.2010.09.003

18. Buch S, Schafmayer C, Völzke H, et al. A genome-wide association scan identifies the hepatic cholesterol transporter ABCG8 as a susceptibility factor for human gallstone disease. Nat Genet. 2007; 39(8): 995-999. doi: 10.1038/ng2101

19. Nervi F, Miquel JF, Alvarez M, et al. Gallbladder disease is associated with insulin resistance in a high risk Hispanic population. J Hepatol. 2006; 45(2): 299-305. doi: 10.1016/j.jhep.2006.01.026

20. Völzke H, Alte D, Schmidt CO, et al. Cohort profile: The study of health in pomerania. Int J Epidemiol. 2005; 34(2): 393-398. doi: 10.1093/ije/dyp394

21. Völzke H. Multicausality in fatty liver disease: Is there a rationale to distinguish between alcoholic and non-alcoholic origin? World J Gastroenterol. 2012; 18(27): 3492-3501. doi: 10.3748/wjg.v18.i27.3492
22. Romeo S, Kozlitina J, Xing C, et al. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. Nat Genet. 2008; 40(12): 1461-1465. doi: 10.1038/ng.257

23. Ahmad OB, Boschi-Pinto C, Lopez AD, Murray CJ, Lozano R, Inoue M. Age standardization of rates: a new WHO standard. GPE Discussion Paper Series: World Health Organization, 2001. Web site. http://www.who.int/healthinfo/paper31.pdf. Accessed December 30, 2015.

24. Chevan A, Sutherland M. Hierarchical partitioning. Am Stat. 1991; 45(2): 90-96. doi: 10.2307/2684366

25. Walsh C, MacNally R. Hierarchical partitioning. The R Project for Statistical Computing. 2004.

26. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing Vienna Austria 2010.

27. Sauerbrei W, Royston P. Building multivariable prognostic and diagnostic models: transformation of the predictors by using fractional polynomials. J R Stat Soc: Series A. 1999; 162(1): 71-94. doi: 10.1111/1467-985X.00122

28. Vernon G, Baranova A, Younossi ZM. Systematic review: the epidemiology and natural history of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in adults. Aliment Pharmacol Ther. 2011; 34(3): 274-285. doi: 10.1111/j.1365-2036.2011.04724.x

29. Lopez-Velazquez JA, Silva-Vidal KV, Ponciano-Rodriguez G, et al. The prevalence of nonalcoholic fatty liver disease in the Americas. Ann Hepatol. 2014; 13(2): 166-178. Web site. http://www.annalsofhepatology.com/revista/numeross/2014/03_142_v13n2_2014_PrevalenceNonalcoholic.pdf. Accessed December 30, 2015.

30. Lizardi-Cervera J, Laparra D, Chavez-Tapia N, Ostos M, Esquivel M. Prevalence of NAFLD and metabolic syndrome in asymptomatic subjects. Rev Gastroenterol Mex. 2006; 71: 453-459.

31. Riquelme A, Arrese M, Soza A, et al. Non-alcoholic fatty liver disease and its association with obesity, insulin resistance and increased serum levels of C-reactive protein in Hispanics. Liver Int. 2009; 29(1): 82-88. doi: 10.1111/j.1478-3231.2008.01823.x

32. Boehm BO, Claudi-Boehm S, Yildirim S, et al. Prevalence of the metabolic syndrome in southwest Germany. Scand J Clin Lab Invest Suppl. 2005; 240: 122-128.

33. Szczepaniak LS, Nurenberg P, Leonard D, et al. Magnetic resonance spectroscopy to measure hepatic triglyceride content: prevalence of hepatic steatosis in the general population. Am J Physiol-endoc M. 2005; 288(2): E462-E468. doi: 10.1152/ajpendo.00064.2004

34. Lazo M, Hernaez R, Eberhardt MS, et al. Prevalence of Non-alcoholic Fatty Liver Disease in the United States: The Third National Health and Nutrition Examination Survey, 1988-1994. Am Journal Epidemiol. 2013; 178(1): 38-45. doi: 10.1093/aje/kws448

35. Kollerits B, Coassin S, Beckmann ND, et al. Genetic evidence for a role of adiponutrin in the metabolism of apolipoprotein B-containing lipoproteins. Hum Mol Genet. 2009; 18(23): 4669-4676. doi: 10.1093/hmg/ddp424

36. Krawczyk M, Gruenhage F, Mahler M, Tirzhu S, Acalovschi M, Lammert F. The common adiponutrin variant p.1148M does not confer gallstone risk but affects fasting glucose and triglyceride levels. J Physiol Pharmacol. 2011; 62(3): 369-375. Web site. http://www.jpp.krakow.pl/journal/archive/06_11/pdf/369_06_11_article.pdf. Accessed December 30, 2015.

37. Palmer CN, Maglio C, Pirazzi C, et al. Paradoxical lower serum triglyceride levels and higher type 2 diabetes mellitus susceptibility in obese individuals with the PNPLA3 148M variant. PLoS One. 2012; 7(6): e39362. doi: 10.1371/journal.pone.0039362

38. Krawczyk M, Portincasa P, Lammert F. PNPLA3-associated steatohepatitis: toward a gene-based classification of fatty liver disease. Semin Liver Dis. 2013; 33(4): 369-379. doi: 10.1055/s-0033-135825

39. Schwimmer JB, Celedon MA, Lavine JE, et al. Heritability of nonalcoholic fatty liver disease. Gastroenterology. 2009; 136(5): 1585-1592. doi: 10.1053/j.gastro.2009.01.050

40. Kratzer W, Mason RA, Kachele V. Prevalence of gallstones in sonographic surveys worldwide. J Clin Ultrasound. 1999; 27(1): 1-7. doi: 10.1002/(SICI)1097-0096(199901)27:1<1::AID-JCU1>3.0.CO;2-H

41. Attili AF, Carulli N, Roda E, et al. Epidemiology of gallstone disease in Italy: prevalence data of the Multicenter Italian Study on Cholelithiasis (M.I.COL.). Am J Epidemiol. 1995; 141(2): 158-165.

42. Jorgensen T. Prevalence of gallstones in a Danish population. Am J Epidemiol. 1987; 126(5): 912-921.

43. Friedrich N, Völzke H, Hampe J, Lerch MM, Jorgensen T. Known risk factors do not explain disparities in gallstone prevalence between denmark and northeast Germany. Am J Gastroenterol. 2009; 104(1): 89-95. doi: 10.1038/ajg.2008.13

44. Scragg RK, McMichael AJ, Baghurst PA. Diet, alcohol, and relative weight in gall stone disease: a case-control study. Br Med J. 1984; 288(6424): 1113-1119.

45. Nervi F, Miquel JF, Marshall G. The Amerindian epidemics of cholesterol gallstones: the North and South connection. Hep-
46. Krawczyk M, Miquel JF, Stokes CS, et al. Genetics of biliary lithiasis from an ethnic perspective. *Clin Res Hepatol Gastroenterol.* 2013; 37(2): 119-125. doi: 10.1016/j.clinre.2012.09.002

47. Thijs C, Knipschild P. Oral contraceptives and the risk of gallbladder disease: a meta-analysis. *Am J Public Health.* 1993; 83(8): 1113-1120.

48. Etminan M, Delaney JA, Bressler B, Brophy JM. Oral contraceptives and the risk of gallbladder disease: a comparative safety study. *CMAJ.* 2011; 183(8): 899-904. doi: 10.1503/cmaj.110161

49. Wang FZ, Shen XF, Guo XR, Peng YZ. Oral contraceptives and risk of gallbladder disease. *CMAJ.* 2011; 183(13): 1517. doi: 10.1503/cmaj.111-2065
### APPENDIX

#### Regression table for fatty liver disease in Chilean women.

|                      | Odds Ratio | \( p \) value | 95% CI     |
|----------------------|------------|---------------|------------|
| Age (years)          | 0.986      | 0.352         | 0.958      | 1.015      |
| BMI polynomial 1\(^2\) (kg/m\(^2\)) | 2.981      | <0.001\(^a\) | 1.506      | 5.901      |
| BMI polynomial 2\(^2\) (kg/m\(^2\)) | 0.507      | 0.324         | 0.794      |
| Alcohol (g/d)        | 1.014      | 0.127         | 0.996      | 1.033      |
| Serum triglycerides polynomial 1\(^2\) (mmol/l) | 1.2E-04    | 0.014\(^a\)  | 7.5E-10    | 20.00      |
| Serum triglycerides polynomial 2\(^2\) (mmol/l) | 1.1E-16    | 5.7E-29       | 2.1E-04    |
| Serum glucose (mmol/l) | 0.832      | 0.607         | 1.140      |
| Diabetes (yes)       | 1.037      | 0.963         | 0.224      | 4.795      |
| SNP rs738409         | 1.799      | 0.014         | 1.125      | 2.876      |

\(^a\)transformation: (BMI/10)^\(0.5\)

#### Regression table for fatty liver disease in Chilean men.

|                      | Odds Ratio | \( p \) value | 95% CI     |
|----------------------|------------|---------------|------------|
| Age polynomial 1\(^2\) (years) | 204.0      | <0.001\(^a\) | 42.8       | 971.6      |
| Age polynomial 2\(^2\) (years) | 0.910      | 0.880         | 0.940      |
| BMI (kg/m\(^2\))     | 1.178      | <0.001        | 1.139      | 1.218      |
| Alcohol (g/d)        | 1.014      | 1.007         | 1.020      |
| Serum triglycerides (mmol/l) | 1.431      | <0.001        | 1.304      | 1.571      |
| Serum glucose (mmol/l) | 1.154      | 1.055         | 1.262      |
| Diabetes (yes)       | 1.072      | 0.727         | 1.581      |
| SNP rs738409         | 1.521      | <0.001        | 1.251      | 1.849      |

\(^a\)transformation: (age/10)^\(0.5\)

#### Regression table for fatty liver disease in men from the SHIP.

|                      | Odds Ratio | \( p \) value | 95% CI     |
|----------------------|------------|---------------|------------|
| Age (years)          | 0.992      | 0.419         | 0.973      | 1.012      |
| BMI polynomial\(^2\) (kg/m\(^2\)) | 2.9E-13    | <0.001        | 7.0E-17    | 1.2E-09    |
| Alcohol (g/d)        | 0.973      | 0.563         | 0.883      | 1.073      |
| Serum triglycerides (mmol/l) | 1.438      | 1.076         | 1.921      |
| Serum glucose (mmol/l) | 1.095      | 0.959         | 1.250      |
| Diabetes (yes)       | 2.891      | 0.936         | 8.926      |
| SNP rs738409         | 1.813      | 1.276         | 2.577      |

\(^a\)transformation: (BMI/10)^\(0.5\)

#### Regression table for fatty liver disease in Chilean women.

|                      | Odds Ratio | \( p \) value | 95% CI     |
|----------------------|------------|---------------|------------|
| Age (years)          | 0.992      | 0.419         | 0.973      | 1.012      |
| BMI polynomial\(^2\) (kg/m\(^2\)) | 2.9E-13    | <0.001        | 7.0E-17    | 1.2E-09    |
| Alcohol (g/d)        | 0.973      | 0.563         | 0.883      | 1.073      |
| Serum triglycerides (mmol/l) | 1.438      | 1.076         | 1.921      |
| Serum glucose (mmol/l) | 1.095      | 0.959         | 1.250      |
| Diabetes (yes)       | 2.891      | 0.936         | 8.926      |
| SNP rs738409         | 1.813      | 1.276         | 2.577      |

\(^a\)transformation: (BMI/10)^\(0.5\)
### Odds Ratio, p value, 95% CI

| Variable | Odds Ratio | p value | 95% CI     |
|----------|------------|---------|------------|
| Age (years) | 6.3E-09 | <0.001 | 2.4E-11 1.7E-06 |
| BMI (kg/m²) | 1.157 | <0.001 | 1.123 1.192 |
| Alcohol polynomial 1st (g/day) | 1.191 | <0.001 | 1.066 1.306 |
| Alcohol polynomial 2nd (g/day) | 1.015 | <0.001 | 1.007 1.022 |
| Serum triglycerides (mmol/l) | 2.128 | <0.001 | 1.829 2.477 |
| Serum glucose (mmol/l) | 1.175 | 0.013 | 1.035 1.333 |
| Diabetes (yes) | 2.559 | <0.001 | 1.517 4.316 |
| SNP rs738409 | 1.866 | <0.001 | 1.490 2.339 |

1 transformation: (age/10)^2
2 transformation: ((alcohol)/100)^0.5
3 transformation: ((alcohol)/100)^0.5 * ln((alcohol)/100)
4 joint test for both variables

A 4: Regression table for fatty liver disease in women from the SHIP.

### Odds Ratio, p value, 95% CI

| Variable | Odds Ratio | p value | 95% CI     |
|----------|------------|---------|------------|
| Age (years) | 1.052 | 0.004 | 1.017 1.090 |
| BMI (kg/m²) | 1.041 | 0.318 | 0.962 1.126 |
| Alcohol (g/d) | 1.017 | 0.066 | 0.983 1.036 |
| Serum triglycerides (mmol/l) | 1.211 | 0.198 | 0.904 1.623 |
| Glucose (mmol/l) | 0.718 | 0.061 | 0.508 1.016 |
| Diabetes (yes) | 2.659 | 0.265 | 0.476 14.862 |
| SNP rs7187645 | 1.174 | 0.751 | 0.434 3.174 |

5 transformation: (glucose/10)^2

A 5: Regression table for gallstone disease in Chilean men.

### Odds Ratio, p value, 95% CI

| Variable | Odds Ratio | p value | 95% CI     |
|----------|------------|---------|------------|
| Age (years) | 1.062 | <0.001 | 1.051 1.073 |
| BMI (kg/m²) | 1.078 | <0.001 | 1.036 1.122 |
| Alcohol (g/d) | 0.993 | 0.119 | 0.983 1.002 |
| Serum triglycerides (mmol/l) | 1.080 | 0.070 | 0.994 1.173 |
| Glucose polynomial 1st (mmol/l) | 1.148 | 0.053 | 0.998 1.319 |
| Diabetes (yes) | 1.532 | 0.036 | 1.029 2.281 |
| SNP rs7187645 | 1.926 | 0.001 | 1.319 2.813 |

6 transformation: (glucose/10)^2

A 6: Regression table for gallstone disease in men from the SHIP.

### Odds Ratio, p value, 95% CI

| Variable | Odds Ratio | p value | 95% CI     |
|----------|------------|---------|------------|
| Age (years) | 1.046 | <0.001 | 1.027 1.065 |
| BMI (kg/m²) | 1.059 | 0.007 | 1.015 1.104 |
| Alcohol (g/d) | 1.032 | 0.206 | 0.983 1.083 |
| Serum triglycerides (mmol/l) | 0.964 | 0.776 | 0.752 1.237 |
| Glucose (mmol/l) | 1.049 | 0.468 | 0.922 1.193 |
| Diabetes (yes) | 1.485 | 0.444 | 0.539 4.093 |
| Parity (yes) | 2.572 | 0.121 | 0.778 8.497 |
| Contraceptives (yes) | 0.476 | 0.039 | 0.235 0.962 |
| Postmenopausal hormones (yes) | 1.562 | 0.177 | 0.818 2.985 |
| SNP rs7187645 | 1.456 | 0.125 | 0.901 2.352 |

7 transformation: (age/10)^2

A 7: Regression table for gallstone disease in Chilean women.
### Odds Ratio, p value, 95% CI

| Variable                        | Odds Ratio | p value | 95% CI    |
|---------------------------------|------------|---------|-----------|
| Age (years)                     | 1.058      | <0.001  | 1.048 1.068 |
| BMI (kg/m²)                     | 1.085      | <0.001  | 1.060 1.110 |
| Alcohol (g/d)                   | 1.001      | 0.905   | 0.985 1.017 |
| Triglycerides (mmol/l)          | 1.140      | 0.010   | 1.032 1.260 |
| Glucose (mmol/l)                | 0.931      | 0.130   | 0.849 1.021 |
| Diabetes (yes)                  | 1.350      | 0.172   | 0.878 2.075 |
| Parity (yes)                    | 1.665      | 0.022   | 1.075 2.577 |
| Contraceptives (yes)            | 1.157      | 0.497   | 0.760 1.760 |
| Postmenopausal hormones (yes)   | 0.971      | 0.826   | 0.745 1.285 |
| SNP rs11887534                  | 2.150      | <0.001  | 1.590 2.906 |

**A 8:** Regression table for gallstone disease in women from the SHIP.

**Fatty liver disease in Chilean men**

Area under ROC curve = 0.8410

**A 9:** ROC curve for fatty liver disease in Chilean men.

**Fatty liver disease in German men**

Area under ROC curve = 0.7936

**A 10:** ROC curve for fatty liver disease in German men.

**Fatty liver disease in Chilean women**

Area under ROC curve = 0.8296

**A 11:** ROC curve for fatty liver disease in Chilean women.

**Fatty liver disease in German women**

Area under ROC curve = 0.8832

**A 12:** ROC curve for fatty liver disease in German women.
Gallstone disease in Chilean men

Area under ROC curve = 0.7418

A 13: ROC curve for gallstone disease in Chilean men.

Gallstone disease in German men

Area under ROC curve = 0.7720

A 14: ROC curve for gallstone disease in German men.

Gallstone disease in Chilean women

Area under ROC curve = 0.7498

A 15: ROC curve for gallstone disease in Chilean women.

Gallstone disease in German women

Area under ROC curve = 0.7887

A 16: ROC curve for gallstone disease in German women.