



Published in final edited form as:

Clin Gastroenterol Hepatol. 2008 November ; 6(11): 1243–1248. doi:10.1016/j.cgh.2008.07.013.

Placebo in Nonalcoholic Steatohepatitis: Insight into Natural History and Implications for Future Clinical Trials

Rohit Loomba, M.D. MHSc^{1,2}, Robert Wesley, Ph.D.³, Frank Pucino, Pharm D.⁴, T. Jake Liang, M.D.¹, David E. Kleiner, M.D., Ph.D.⁵, and Joel E. Lavine, M.D., PhD.⁶

¹ Liver Diseases Branch, National Institute of Diabetes and Digestive Diseases (NIDDK), National Cancer Institute, National Institutes of Health, Bethesda, MD

² Division of Gastroenterology, Department of Medicine, University of California at San Diego, La Jolla, CA

³ Hospital Epidemiology, National Cancer Institute, National Institutes of Health, Bethesda, MD

⁴ Pharmacy Department, Clinical Center, National Cancer Institute, National Institutes of Health, Bethesda, MD

⁵ Laboratory of Pathology, National Cancer Institute, National Institutes of Health, Bethesda, MD

⁶ Division of Pediatric Gastroenterology, Department of Pediatrics, University of California at San Diego, La Jolla, CA

Abstract

Background and aims—Changes in biochemical and histologic parameters related to nonalcoholic steatohepatitis (NASH) in placebo-treated patients may provide an insight into natural history and help in defining treatment endpoints in NASH. The aim of our study was to assess the biochemical and histologic changes seen in the placebo-arm of the randomized-placebo-controlled trials in adult patients with NASH.

Methods: Data Sources—Medline was searched (through May 2008) for studies published in the English language.

Study Selection—Randomized, placebo-controlled trials of at least 6-month duration in patients with NASH that provided biochemical and/or histologic data of placebo-arm.

Data Extraction—One investigator performed literature search and data extraction. Two investigators independently confirmed that the studies met pre-specified criteria. Pooled estimates of biochemical and histologic parameters associated with NASH were calculated.

Results—Five randomized-controlled trials met the predefined criteria and included 162 placebo-treated and 189 active-treatment patients. Mean serum alanine (ALT) and aspartate (AST) aminotransferase decrease on placebo. 1-point improvement in steatosis, ballooning degeneration, lobular inflammation, NASH fibrosis and combined inflammation scores is seen in 31%, 15%,

© 2008 The American Gastroenterological Association. Published by Elsevier Inc. All rights reserved.

Address Correspondence to: Rohit Loomba, M.D. MHSc, Division of Gastroenterology, University of California at San Diego, UC 303, MC0063, 9500 Gilman Drive, La Jolla, CA 92093, Telephone: 301-379-8587, Fax: 301-480-7926, roloomba@ucsd.edu.

Conflict of Interest: No conflict of interest exist

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

33%, 22% and 32% of patients, respectively. 2-point improvement in NASH histologic scores is rarely seen.

Conclusions—Serum ALT may decrease on placebo and is not a reliable measure of treatment response. Although a 1-point improvement is seen in a third of patients, a 2-point improvement in histologic parameters is rarely seen in placebo-arm and may be more reliable in assessing treatment response. These data may have important implications in designing future clinical trials in NASH.

Keywords

Nonalcoholic steatohepatitis; randomized placebo-controlled double-blind study; clinical trial; placebo-effects; natural history of disease; pooled analysis; meta-analysis; treatment end-points; fatty liver disease; fibrosis; serum aminotransferase levels; liver histology and histologic scoring systems

Introduction

Randomized, double-blind, placebo-controlled studies are considered the gold standard in evaluating the efficacy of medical interventions in clinical trials(1). These individual placebo-controlled studies are mainly utilized to examine and report treatment efficacy as compared to placebo. Additionally, placebo-controlled studies provide important information to both the clinicians and researchers regarding variability in the disease related parameters in the placebo-arm that may be explained by either Hawthorne-effect (a short-term improvement in outcome variable simply due to the increased attention to the participants who know that they are being studied and closely followed), or sampling variability, or repeat measurements by chance, or rarely, true placebo effect (2, 3). Furthermore, it provides valuable information regarding the natural history of disease. A meta-analysis of placebo versus no-placebo control arm studies show that placebo effects were rarely significant except in subjective measures, especially when measured as a continuous variable(2).

A placebo-arm can also guide researchers in identifying suitable treatment endpoints for pilot studies to screen for potential interventions or medications. Recognition of appropriate treatment endpoints is especially critical in chronic diseases that lack a single surrogate marker of disease activity or progression, and may utilize a composite endpoint composed of two or more parameters, or disease activity index (4). Composite end points have been commonly utilized in assessing efficacy of medications in liver diseases. Classically, improvement in liver histology has been considered an endpoint in establishing efficacy of medications in the treatment of chronic liver diseases such as viral hepatitis and nonalcoholic steatohepatitis (NASH)(5, 6). NASH is a clinico-pathologic entity that is seen in individuals who consume little alcohol and have evidence of necroinflammation, ballooning degeneration and steatosis with or without peri-sinusoidal fibrosis (7). The diagnosis is dependent on both clinical context and liver histology. Currently, there is no serologic or non-invasive surrogate marker of disease activity or progression in NASH. Furthermore, there is no uniformly recognized treatment for NASH and several randomized, placebo-controlled studies have been conducted to examine the efficacy of various treatment approaches. Screening potential agents for the treatment of NASH requires valid treatment endpoints (8). Whether changes in serum aminotransferase alone can be considered an adequate treatment endpoint, in lieu of improvement in liver histology, is an important clinical issue (9). It is also necessary to define whether any improvement in a specific parameter or 2-point decrease in a disease-specific parameter such as steatosis, or ballooning degeneration, or lobular inflammation may provide better assessment of efficacy.

Knowledge of the placebo-arm of randomized-controlled studies in NASH may provide valuable information to better understand changes in biochemical or histologic parameters over time and help define valid treatment endpoints for subsequent trials. We conducted a pooled analysis of biochemical and histologic changes in seen in placebo-treated patients from the eligible randomized, placebo-controlled studies in patients with NASH.

Methods

Retrieval of clinical trials

The Medline database was searched for manuscripts written in English through May, 2008. Indexing terms included *nonalcoholic steatohepatitis or NASH* in combination with *randomized-controlled trials*, and *nonalcoholic fatty liver disease or NAFLD* in combination with *randomized-controlled trials*. A manual review of the bibliographies of seminal primary and review articles also was performed to identify additional studies.

Criteria for study inclusion in the meta-analysis included: 1) Randomized, placebo-controlled clinical trials in patients with NASH (not NAFLD alone); 2) Minimum 24 weeks of treatment and 3) Well-defined treatment outcomes (defined by reporting at least one of the following: changes in serum ALT or AST, or liver histologic parameters related to NASH).

Trials were excluded if relevant data were not extractable, if the case-mix included both NASH and NAFLD patients, lacked inter-independence with other trials or lacked peer review (e.g., meeting abstract). Additionally, case reports or series and uncontrolled studies were excluded.

Data Extraction and Quality Assessment

One investigator (RL) performed the initial literature search and data extraction. Two investigators independently (RW and FP) reviewed whether the studies met pre-specified criteria and verified the extracted data. Quality was assessed based upon the design of the studies. All randomized, placebo-controlled clinical trials were considered as level 1 in the quality of evidence. As the total number of studies was small we elected to keep the descriptive labels such as allocation-concealment, and intention to treat analysis, rather than numeric scores to represent quality in the Results Section.

Due to heterogeneity in biochemical, anthropometric and histologic outcome measurements and use of different histologic systems the following assumptions were made- the average change in pre-placebo and post-placebo mean values should have a good correlation irrespective of the unit of measurement. As the outcome measurement is dependent upon average change in biochemical, anthropometric, and histologic parameters therefore, measures can be pooled to achieve reliable estimates.

Statistical analysis

Pooled estimates of each variable of interest were calculated and presented as a mean. The cumulative rate of each outcome of interest was calculated in the placebo arm of eligible studies. Pooled summary estimates were compared before and after placebo-treatment to test for statistically significant differences. For categorical variables, a stratified Fisher Exact Test was used to determine whether a statistically significant difference exists in the rates of outcome before and after placebo treatment using the StatXact software version 6 (Cytel Software Corp., Cambridge, MA).

To derive the overall meta-analysis P-values for most parameters, we formed a test statistic based on summing the individual mean changes from each usable study; under the null hypothesis testing no change between pre- and post-values, this sum will be (asymptotically) normally distributed with a variance dependent on the standard errors for the mean changes in variables in each study. The resulting overall p-values are only approximate; due to partially missing results in various studies (e.g. for the standard errors for average changes in the variables or sometimes even for post-baseline standard deviations), we estimated certain parameters based upon deduction from the complete data available in other studies. For the outcome of combined necroinflammation, because even the minimal information necessary to employ the above approach was absent, we computed the overall P-value (based on the results of the studies that had any usable information) using Fisher's method for combining P-values (10). To determine the effect of study duration on biochemical and anthropometric variables, a Least squares regression was used to assess trend across studies with a 6, 12 and 24 month duration. To evaluate histologic changes over time, an exact test for trend based on Jonckheere-Terpstra test as applied to contingency table data was used for studies with available data. Statistical significance for the two-tailed p-values was set *a priori* as <0.05 unless otherwise specified.

Role of the Funding Sources

Results

Seven studies met the specified criteria on initial screen. We excluded one study because of short follow-up (11) and a second study because patients had NAFLD and not exclusively NASH (12). Of 5 clinical trials in the pooled analysis, 2 were conducted in North America (13, 14), and one each in Switzerland (15), Iran (16) and France (17). Of the 5 studies deemed eligible, 4 were multi-center (13–15, 17), and one single center (16). The quality indicators and study characteristics are shown in Table 1. A total of 162 patients served as placebo-controls and a total of 189 patients were in the treatment arm. The Lindor, Belfort and Ratziu studies provided the best estimates of placebo changes based upon the histologic data presented in the manuscript and therefore were classified as level 1A in quality. The Dufour study was level 1B, and the Merat study was level IC in quality.

Pooled-analysis

The results of the pooled analysis are shown in Table 2 and 3. The mean serum alanine aminotransferase (ALT) and aspartate aminotransferase levels decreased from 85 U/L to 65 U/L (p-value <0.001), and 56 U/L to 48 U/L (p-value= 0.07), respectively in patients who were receiving placebo. Serum ALT decreased on placebo in 4 studies and did not change in one study. Serum AST levels decreased in 4 but increased marginally in one study. However, there was no change in weight or BMI on placebo.

The mean scores of all the NASH associated histologic parameters such as steatosis, ballooning degeneration, lobular inflammation, and NASH fibrosis decreased on placebo and changes in steatosis scores were statistically significant (p-value<0.001). However, the median scores did not change. Minor changes in individual patient scores are common. A one point improvement in steatosis, ballooning degeneration, lobular inflammation, NASH fibrosis, and combined or overall inflammation scores is seen in 31%, 15%, 33%, 22%, and 32% of patients, respectively. A two-point improvement in steatosis, ballooning degeneration, lobular inflammation, NASH fibrosis, and combined (or overall) inflammation scores is seen in 1%, 2%, 0%, 4%, and 8%, respectively.

Effect of Study Duration

We separately analyzed the effect of treatment duration on the results of the pooled analyses. The trend results across studies from 6 to 12 to 24 months were borderline significant ($p=0.047$ for trend) for ALT only, with greater reduction observed with longer treatment duration. Serum AST, BMI, and histologic outcome parameters were not significantly affected.

Discussion

The main finding of this pooled analysis is that serum ALT and AST declined in placebo-treated patients with NASH. Similar reductions are seen in mean changes in NASH associated histologic parameters. The greatest decline in mean histologic scores is seen in the steatosis score, which is statistically significant. Overall decrease in these parameters is small but depending upon how an outcome measure is defined it may become statistically significant. One point reduction in any of the NASH associated histologic parameters can be seen in up to one-third of patients but two-point improvement in these parameters is rarely seen. This study provides important information regarding the natural history of NASH and supports the notion that serum ALT alone is unreliable as an outcome measure of severity of NASH or efficacy of treatment.

A possible explanation for why placebo-group ALT and AST decreases during follow-up is that these values fluctuate in NASH, and that decisions to biopsy may be based on high(er) ALT values. Thus, the patients who have higher serum ALT values are more likely to receive a liver biopsy or a referral to a tertiary care center and enter into clinical trials (later fluctuating to a lower serum ALT value), while the patients with lower serum ALT values are less likely to receive a liver biopsy or a referral to a tertiary care center and not enter into clinical trials. Therefore, there may be a selection bias in favor of recruiting patients who have higher serum ALT levels into clinical studies. Another explanation for decline in serum ALT and AST during follow-up in placebo treated patients could be due to the phenomenon of regression to the mean. Additionally, one-point improvement in the NASH histologic scores could occur in up to a third of patients possibly because of repeat measurement or sampling variability. Furthermore, it also signifies that up to two-thirds of patients either do not improve or worsen. Therefore, a one-point improvement in an individual NASH score in an individual patient with NASH should not signify a positive treatment effect.

However, if the majority of patients in a treatment-arm achieve a one-point improvement in a particular NASH histologic parameter it may signify a positive drug effect and may be suggestive of mechanism of action of a drug in NASH. In this pooled analysis, we report that a two-point, rather than mean decline or one-point, improvement in the NASH histology parameters is rarely seen in placebo-treated patients and therefore, it may be a more reliable indicator of treatment effect. A two-point decline in one or more parameters of NASH activity index may be better in certain settings. Placebo-treated individuals who achieved a two-point improvement in any histologic parameter may have significantly improved their eating habits or changed their life-style by incorporating regular exercise. However, it was not possible to extract these data from the available studies to evaluate the reasons for these changes. As the ballooning degeneration, lobular inflammation and fibrosis scores did not change in the placebo treated patients, it would be desirable to construct outcome criteria that would not allow improvement in steatosis alone (even a 2 point improvement) to be considered "clinically significant histologic improvement". In order to suggest "clinically significant histologic improvement" in a clinical trial the cohort as a whole should show improvement in more than one histologic parameter in a consistent manner.

These data have important implications for designing future clinical trials and reporting study results. We propose that future studies utilize composite histologic response criteria as utilized by the NASH-Clinical Research Network or similar previously proposed criteria. This would greatly enhance the generalizability and set a stage for standard approach in nomenclature and reporting in NASH studies that is critically required. In future NASH clinical trials, authors should consider providing detailed changes in histologic scores either as shown by Belfort et al or Lindor et al along with summary estimates. This would facilitate future meta-analyses to estimate true medication (and placebo) effects in diverse international settings.

Although the average BMI in these studies was 31 kg/m² (in the obese range), there was no weight loss in the placebo-arms on an average. These data suggest that weight loss in placebo-arms is seldom seen in NASH patients.

Strengths of the analysis

The strengths of this study include the good quality level of the studies included in the meta-analysis. The results were uniform across studies with minor variability. The patients included in the meta-analysis were derived from three continents and there was adequate follow-up. Importantly, all patients had biopsy-proven NASH. Additional analyses were conducted to address the effect of treatment duration on the results of the research synthesis.

Limitations of this analysis

Racial/ethnic or gender based differences in these four studies could not be analyzed due to unavailability of these data. Only three studies reported detailed histologic data before and after placebo. The internal validity of histologic findings is satisfactory as most patients were enrolled in these three studies (Lindor et al., Belfort et al. and Ratziu et al.) and these three studies also had a better quality score as described in the Results Section. Studies included in this research synthesis used different histologic criteria for entry into the study and used different end point and histologic scoring systems to assess treatment responses. In addition, the score range for ballooning degeneration and steatosis differs between various NASH histologic scoring system, which makes it difficult to compare the magnitude of change in these parameters across treatment trials. Despite this limitation, unidirectional change consistently across two or three histologic scoring system may be a reliable finding that may be suggestive of true treatment effect. However, this limitation may increase the generalizability of our findings and at the same time underscores the need for standardization of outcomes, histologic scoring system, and reporting in future NASH clinical trials. Due to unavailability of the data, we could not assess the effect of placebo on subjective measures such as fatigue, right upper quadrant pain or quality of life. It is possible that the placebo-treated patients may have received varied nutritional counseling and lifestyle intervention in each study. We believe that it is unlikely to affect the results of our study as none of the trials showed any significant weight loss in placebo patients. This research synthesis is not able to assess the placebo-mind effects and their impact in NASH.

Implications and Conclusions

Serum ALT and AST change during follow-up in NASH patients and may not reflect histologic improvement. Minor improvements in NASH associated histologic scores are seen in up to a third of patients although a two-point improvement in steatosis, ballooning degeneration or lobular inflammation is rarely seen in placebo-treated individuals. These data have important implications for designing and reporting future clinical trials in NASH and suggest a need for standardization of terminologies, histologic scoring system and treatment end-points.

Acknowledgments

Financial and NIH Support: This study was supported by the intramural research programs of the National Institute of Diabetes and Digestive and Kidney Diseases and the National Cancer Institute, National Institutes of Health.

Role of sponsor: No conflict of interest exists.

Authors thank Dr. Jay H. Hoofnagle for providing helpful comments at various stages of the study.

Abbreviations

NASH	nonalcoholic steatohepatitis
NAFLD	nonalcoholic fatty liver disease
ALT	alanine aminotransferase
AST	aspartate aminotransferase
RCT	randomized-controlled trial
NIDDK	National Institute of Diabetes and Digestive and Kidney Diseases
NIH	National Institutes of Health

References

- Greenfield S, Kravitz R, Duan N, Kaplan SH. Heterogeneity of treatment effects: implications for guidelines, payment, and quality assessment. *Am J Med.* 2007; 120:S3–9. [PubMed: 17403380]
- Hrobjartsson A, Gotzsche PC. Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment. *N Engl J Med.* 2001; 344:1594–1602. [PubMed: 11372012]
- Su C, Lewis JD, Goldberg B, Brensinger C, Lichtenstein GR. A meta-analysis of the placebo rates of remission and response in clinical trials of active ulcerative colitis. *Gastroenterology.* 2007; 132:516–526. [PubMed: 17258720]
- Su C, Lichtenstein GR, Krok K, Brensinger CM, Lewis JD. A meta-analysis of the placebo rates of remission and response in clinical trials of active Crohn's disease. *Gastroenterology.* 2004; 126:1257–1269. [PubMed: 15131785]
- Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, Ferrell LD, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology.* 2005; 41:1313–1321. [PubMed: 15915461]
- Promrat K, Lutchman G, Uwaifo GI, Freedman RJ, Soza A, Heller T, Doo E, et al. A pilot study of pioglitazone treatment for nonalcoholic steatohepatitis. *Hepatology.* 2004; 39:188–196. [PubMed: 14752837]
- Brunt EM, Janney CG, Di Bisceglie AM, Neuschwander-Tetri BA, Bacon BR. Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions. *Am J Gastroenterol.* 1999; 94:2467–2474. [PubMed: 10484010]
- Neuschwander-Tetri BA, Caldwell SH. Nonalcoholic steatohepatitis: summary of an AASLD Single Topic Conference. *Hepatology.* 2003; 37:1202–1219. [PubMed: 12717402]
- Suzuki A, Lymp J, Sauver JS, Angulo P, Lindor K. Values and limitations of serum aminotransferases in clinical trials of nonalcoholic steatohepatitis. *Liver Int.* 2006; 26:1209–1216. [PubMed: 17105586]
- Littell RaF J. Asymptotic Optimality of Fisher's Method of Combining Independent Tests. *J Am Statist Assoc.* 1971; 66:802–806.
- Chande N, Laidlaw M, Adams P, Marotta P. Yo Jyo Hen Shi Ko (YHK) improves transaminases in nonalcoholic steatohepatitis (NASH): a randomized pilot study. *Dig Dis Sci.* 2006; 51:1183–1189. [PubMed: 16944007]

12. Zelber-Sagi S, Kessler A, Brazowsky E, Webb M, Lurie Y, Santo M, Leshno M, et al. A double-blind randomized placebo-controlled trial of orlistat for the treatment of nonalcoholic fatty liver disease. *Clin Gastroenterol Hepatol*. 2006; 4:639–644. [PubMed: 16630771]
13. Lindor KD, Kowdley KV, Heathcote EJ, Harrison ME, Jorgensen R, Angulo P, Lymp JF, et al. Ursodeoxycholic acid for treatment of nonalcoholic steatohepatitis: results of a randomized trial. *Hepatology*. 2004; 39:770–778. [PubMed: 14999696]
14. Belfort R, Harrison SA, Brown K, Darland C, Finch J, Hardies J, Balas B, et al. A placebo-controlled trial of pioglitazone in subjects with nonalcoholic steatohepatitis. *N Engl J Med*. 2006; 355:2297–2307. [PubMed: 17135584]
15. Dufour JF, Oneta CM, Gonvers JJ, Bihl F, Cerny A, Cereda JM, Zala JF, et al. Randomized placebo-controlled trial of ursodeoxycholic acid with vitamin e in nonalcoholic steatohepatitis. *Clin Gastroenterol Hepatol*. 2006; 4:1537–1543. [PubMed: 17162245]
16. Merat S, Malekzadeh R, Sohrabi MR, Sotoudeh M, Rakhshani N, Sohrabpour AA, Naserimoghdam S. Probuocol in the treatment of non-alcoholic steatohepatitis: a double-blind randomized controlled study. *J Hepatol*. 2003; 38:414–418. [PubMed: 12663231]
17. Ratziu V, Giral P, Jacqueminet S, Charlotte F, Hartemann-Heurtier A, Serfaty L, Podevin P, Lacorte JM, Bernhardt C, Bruckert E, Grimaldi A, Poynard T. Rosiglitazone for Nonalcoholic Steatohepatitis: One-Year Results of the Randomized Placebo-Controlled Fatty Liver Improvement With Rosiglitazone Therapy (FLIRT) Trial. *Gastroenterology*. 2008 (in press).

Table 1

Characteristics of studies included in the analysis

Author	Year	Type of study	Country	Duration	Treatment group medication	Number of patients in treatment group	Number of patients in the placebo group	Design	Blinding	Allocation concealment (Yes/no)	Sample size calculations in manuscript (Yes/no)	ITT analysis (Yes/no)	Histologic scoring system used
Belfort	2006	Multi-center	USA	6 months	Pioglitazone	26	21	RCT	None	No	Yes	Yes	Kleiner et al.
Dufour	2006	Multi-center	Switzerland	24 months	Urso + Vit E; Ursol alone	15; 18	15	RCT	Double-blind	No	No	No	Promrat et al.
Lindor	2004	Multi-center	North-America	24 months	Ursol	80	86	RCT	Double-blind	No	No	No	(Mod) Brunt et al.
Merat	2003	Single-center	Iran	6 months	Probucol	18	9	RCT	Double-blind	No	No	No	Brunt et al.
Ratziu	2008	Multi-center	France	12 months	Rosiglitazone	32	31	RCT	Double-blind	No	Yes	No	Brunt et al. Kleiner et al (NAS*only).

Abbreviations: Ursol, Ursodiol; Vit, Vitamin; RCT, randomized-controlled trial; NAS, NASH activity score

Table 2

Pre- and post-treatment changes in biochemical and anthropometric variables in placebo-treated patients with NASH: Pooled results.

Variable	Study (year)	Duration (months)	Pre-Treatment Mean	Delta Mean	P-value
ALT (U/L) (n=137)	Belfort (2006)	6	61	-21	
	Merat (2003)	6	96.8	-0.6	
	Ratziu (2008)	12	84	-7.6	
	Dufour (2006)	24	76	-39	
	Lindor (2004)	24	108	-31.6	
	Combined		85	-20.0	<0.001
AST (U/L) (n=109)	Belfort (2006)	6	42	-9	
	Merat (2003)	6	57.6	-8	
	Ratziu (2008)	12	61	--	
	Dufour (2006)	24	50	6	
	Lindor (2004)	24	70.6	-20.7	
	Combined		56	-7.9	0.07
Weight (kg) (n=126)	Belfort (2006)	6	90.2	-0.5	
	Merat (2003)	6	83.4	0.89	
	Ratziu (2008)	12	--	-1	
	Dufour (2006)	24	83	--	
	Lindor (2004)	24	92	0.09	
	Combined		87.2	-0.13	0.76
BMI (kg/m²) (n=96)	Belfort (2006)	6	32.9	-0.2	
	Merat (2003)	6	30.2	--	
	Ratziu (2008)	12	30.5	--	
	Dufour (2006)	24	30	1	
	Lindor (2004)	24	31.7	-0.1	
	Combined		31.1	0.23	0.36

n, Number of patients that could be used to calculate Deltas and P-values out of a total N=162.

Table 3

Histologic changes in placebo-treated patients with NASH: Pooled results.

Variable	Study (year)	Duration (months)	Delta Mean	% Patients with Any Improvement	% Patients with 2-Point Improvement	P-value
Steatosis (n=124)	Belfort (2006)	6	-0.29	38.1% (8/21)	0% (0/21)	
	Merat (2003)	6	--	--	--	
	Ratziu (2008)	12	-0.16	16.1% (5/31)	--	
	Dufour (2006)	24	-0.50	--	--	
	Lindor (2004)	24	-0.30	36.8% (21/57)	1.8% (1/57)	
	Combined		-0.31	31.2% (34/109)	1.3% (1/78)	<0.001
Ballooning degeneration (n=124)	Belfort (2006)	6	-0.19	23.8% (5/21)	4.8% (1/21)	
	Merat (2003)	6	--	--	--	
	Ratziu (2008)	12	0.23	9.7% (3/31)	0% (0/31)	
	Dufour (2006)	24	0.17	--	--	
	Lindor (2004)	24	-0.10	14.0% (8/57)	1.8% (1/57)	
	Combined		0.03	14.7% (16/109)	1.8% (2/109)	0.68
Lobular inflammation (n=67)	Belfort (2006)	6	-0.29	28.6% (6/21)	0% (0/21)	
	Merat (2003)	6	--	--	--	
	Ratziu (2008)	12	-0.13	35.5% (11/31)	--	
	Dufour (2006)	24	0.17	--	--	
	Lindor (2004)	24	--	--	--	
	Combined		-0.08	32.7% (17/52)	0% (0/21)	0.19
NASH fibrosis (n=122)	Belfort (2006)	6	-0.14	33.3% (7/21)	4.8% (1/21)	
	Merat (2003)	6	--	--	--	
	Ratziu (2008)	12	-0.18	16.1% (5/31)	--	
	Dufour (2006)	24	0.05	--	--	
	Lindor (2004)	24	0.00	21.8% (12/55)	3.6% (2/55)	
	Combined		-0.07	22.4% (24/107)	3.9% (3/76)	0.44
Combined Necro-inflammation or Nash Activity Score (n=107)	Belfort (2006)	6	--	38.1% (8/21)	14.3% (3/21)	
	Merat (2003)	6	--	--	--	
	Ratziu (2008)	12	--	41.9% (13/31)	--	
	Dufour (2006)	24	--	--	--	

Variable	Study (year)	Duration (months)	Delta Mean	% Patients with Any Improvement	% Patients with 2-Point Improvement	P-value
	Lindor (2004)	24	--	23.6% (13/55)	5.5% (3/55)	
	<i>Combined</i>		--	31.8% (34/107)	7.9% (6/76)	0.12

n, Number of patients that could be used to calculate Deltas and P-values out of a total N=162.