Supplementary Appendix

Supplement to: Harrison SA, Bedossa P, Guy CD, et al. A phase 3, randomized, controlled trial of resmetirom in NASH with liver fibrosis. N Engl J Med 2024;390:497-509. DOI: 10.1056/NEJMoa2309000

This appendix has been provided by the authors to give readers additional information about the work.

Supplementary Appendix

Contents

List of Investigators (NCT03900429)

Australia: Leon Adams; John Bate; Mark Danta; Lay Gan; Jacob George; Mohammad Haque; Alexander Hodge; Nicholas Kontorinis; Christopher Leung; Stuart Roberts; Richard Skoien; Siddharth Sood; Simone Strasser; Martin Weltman

Austria: Stephanie Hametner-Schreil; Harald Hofer; Bertram Hölzl; Bernhard Ludvik; Andreas Maieron; Markus Peck-Radosavljevic; Michael Trauner

Belgium: Jochen Decaestecker; Sven Francque; Anja Geerts; Sebastien Kindt; Wim Laleman; Nicolas Lanthier; Luc Lasser; Christophe Moreno; Geert Robaeys; Christophe Van Steenkiste

Canada: Harpreet Bajaj; Meredith Borman; Natasha Chandok; Magdy El Khashab; Hughie Fraser; Hin Ko; Aldo Montano-Loza; Keyur Patel; Marco Puglia; Eric Yoshida

France: Armando Abergel; Edouard Bardou-Jacquet; Jerome Boursier; Jean-Pierre Bronowicki; Christophe Bureau; Laurent Castera; Vincent Di Martino; François Habersetzer; Dominique Larrey; Véronique Loustaud Ratti; Isabelle Mabile-Archambeaud; Philippe Mathurin; Marianne Maynard-Muet; Eric Nguyen-Khac; Stanislas Pol; Dominique Roulot; Lawrence Serfaty; Albert Tran; Anne Varaut

Germany: Keikawus Arastéh; Katrin Arelin; Susanne Beckebaum; Christel Contzen; Münevver Demir; Yvonne Doerffel; Andreas Geier; Patrick Ingiliz; Frank Lammert; Maria Lehretz; Ingolf Schiefke; Natalia Schütte; Stefan Zeuzem

Hungary: Tivadar Banyai; Miklós Lesch; Mihaly Makara; József Nemes; Tamás Nyéky; Márta Varga; István Zsolt Várkonyi

Israel: Ziv Ben-Ari; Marius Braun; Adi Francis; Helena Katchman; Yoav Lurie; Assy Nimer; Rifaat Safadi; Ella Veitsman; David Yardeni; Eli Zuckerman

Italy: Mario Angelico; Maria Del Ben; Stefano Fagiuoli; Pietro Lampertico; Alessandra Mangia; Fabio Marra; Alessandro Mattina; Luca Miele; Salvatore Petta; Mario Pirisi; Massimo Puoti; Giovanni Raimondo; Giuliano Rizzardini; Maurizia Rossana Brunetto; Massimo Zuin

Mexico: Alma Laura Ladron de Guevara; Laura Esthela Cisneros Garza; José Rodríguez Chávez; Jose Flores Figueroa; Mauricio Castillo Barradas

Spain: Juan M Pericàs, Jesús M Rivera-Esteban; Marta Casado Martín; Samaniego Ibañez; Martin Prieto; Jose Luis Calleja Panero; Xavier Torras Colell; Francisco Moises Diago Madrid; Javier Crespo Garcia; Juan Turnes Vazquez; Manuel Romero Gomez; Montserrat Garcia Retortillo; Agustin Albillos Martinez

Switzerland: Jean Francois Dufour; Lorenzo Magenta; Beat Müllhaupt

United Kingdom: William Alazawi; Michael Allison; Quentin Anstee; Sandra Connell; Pratiksha Dokhe; Susannah Eyre; Daniel Forton; Andrew Fowell; Matthew Foxton; Madhu Gowda; Shalini Iyengar; Christina Kyriakidou; Chin Lye Ch'ng; Pinelopi Manousou; Roger McCorry; Philip Newsome; Thomas Pembroke; Mahadev Ramjee; Stephen Ryder; Arun Shankar; David Sheridan; Cyril Sieberhagen; Hawys Thomas; Emmanouil Tsochatzis

United States: Manal Abdelmalek; Michael Adams; Victor Ankoma-Sey; Samir Arora; Vivek Awasty; Jennifer Bashour; Seth Baum; Anuj Bhargava; Brian Borg; Stephen Caldwell; Cornell Calinescu; Shekhar Challa; Michael Charlton; Christopher Christensen; Jeffery DeGrauw; Donna Desantis; Clarence Dunn; Winston Dunn; Michael Fine; Roberto Firpi-Morell; Bradley Freilich; Michael Fuchs; Marisela Gonzalez; Nadege Gunn; Paul Hellstern, Jr.; Stanley Hsia; Jason Huffman; Antoine Jakiche; George James; Robert Jenders; William Jennings; Thomas Kepczyk; Arun Khazanchi; Whitfield Knapple; Anita Kohli; Joseph Kosinski; Kris Kowdley; Marcelo Kugelmas; Eric Lawitz; Kurt Lesh; Kathryn Lucas; Richard Manch; Ravi Mani; Kevin Merkes; Robert Meyer; Apurva Modi; Ann Moore; Amilcar Morales; Abdullah Mubarak; Joseph Newberg; Grisell Ortiz Lasanta; Mangesh Pagadala; Rashmee Patil; John Poulos; Ravi Ravinuthala; Gary Reiss; Michele Reynolds; Mary Rinella; Peter Ruane; John Scott; Aasim Sheikh; Muhammad Sheikh; Mitchell Shiffman; Gerald Shockey; Dan Streja; Parvin Syal; Brent Tetri; George Therapondos; Dawn Torres; James Trotter; Gregg Wendorf; Alexander White; Johnny White; Mark Wilson; Ziad Younes

Supplementary Methods

Patient Prescreening Criteria

Potential patients to be screened for this study should:

- Not have any history of significant alcohol consumption (Exclusion Criterion #1)
- Have at least 3 metabolic risk factors using a slightly modified version of the International Diabetes Foundation (IDF) criteria (Synopsis Table 2)
- Have either
- Both AST and fibroscan requirements
- AST >17 IU (women) and AST >20 IU (men)

AND

- A fibroscan within 3 months of planned screening showing KpA ≥8.5 CAP ≥280. Fibroscan is a potential alternative to a historic eligible liver biopsy to meet inclusion 5 and should be obtained during the prescreening period if possible.

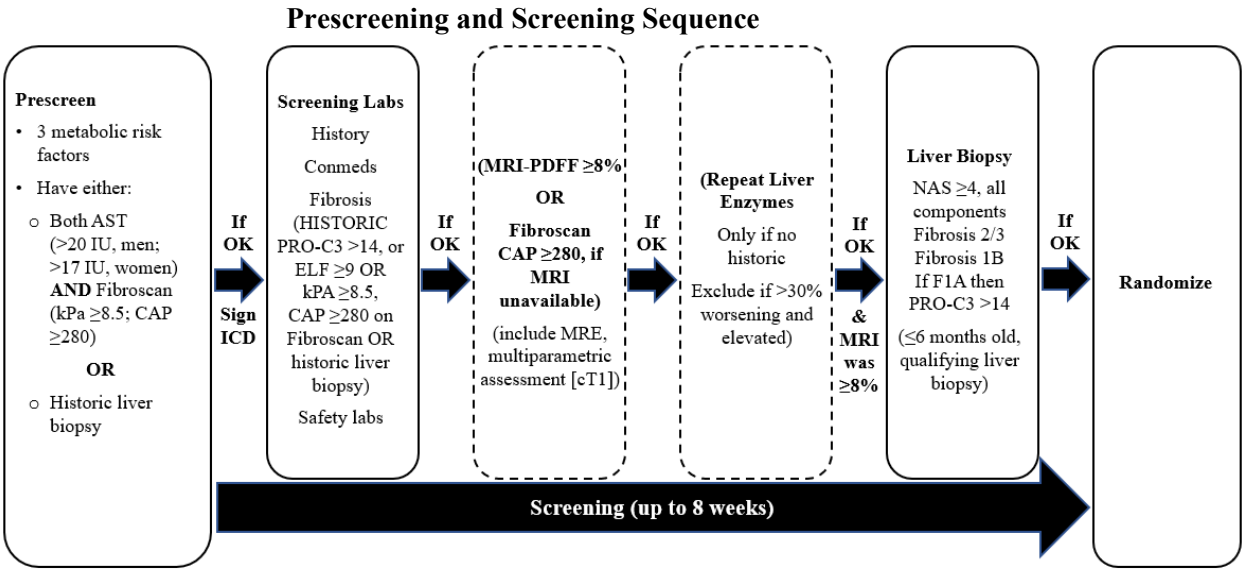
OR

- Have a historic liver biopsy <2 years old demonstrating fibrosis stage 1B, 2 or 3 with NASH (NAS ≥4, all components)

  NOTE: Meets inclusion #5. Prior biopsy reflects documentation of NASH fibrosis and is not necessarily eligible as a baseline biopsy which must have been obtained 6 months prior to randomization with NAS and fibrosis score confirmed as eligible by the central pathology reader.

The IDF risk factors may be assessed by the investigator based on historic values and/or use of concomitant medications for dyslipidemia, hypertension, and diabetes. Metabolic risk factors and the addition of an elevated serum AST value and fibroscan help ensure a higher degree of certainty that a patient will have biopsy- confirmed NASH, NAS≥4 with Stage 1A/1C (high-risk), 1B, 2, or 3 fibrosis with primary emphasis on identifying patients who will have Stage 2 or 3 fibrosis on liver biopsy.

A patient with a historical liver biopsy <2 years with confirmed NASH fibrosis as described above does not require an elevated AST or fibroscan with KpA ≥8.5. In patients with historic liver biopsies, a baseline fibroscan KpA obtained within 3 months of screening or during screening is still needed prior to randomization, but the fibroscan value does not determine eligibility. A baseline fibroscan in all randomized patients obtained during prescreening or screening is used to compare with serial fibroscans during the study.



Dotted box or parentheses, test not mandatory; prescreening Fibroscan values may be used for liver biopsy eligibility.

Inclusion and Exclusion Criteria

Inclusion Criteria

Only evaluate patients for study participation if they meet the Prescreening Criteria. Patients who do not initially meet eligibility criteria may be retested, based on Investigator judgment, to determine if they qualify to participate. Patients who meet all of the following criteria will be eligible to participate in the study:

Must be willing to participate in the study and provide written informed consent.

Male and female adults ≥18 years of age.

Female patients are eligible if they are of reproductive potential and have a negative serum pregnancy test (beta human chorionic gonadotropin), are not breastfeeding, and do not plan to become pregnant during the study and agree to use 2 highly effective birth control methods during the study OR if they are not of child bearing potential (ie, surgically [bilateral oophorectomy, hysterectomy, or tubal ligation] or naturally sterile [>12 consecutive months without menses]). Highly effective birth control methods include condoms with spermicide, diaphragm with spermicide, hormonal and non-hormonal intrauterine

device, hormonal contraception (estrogens stable ≥3 months), a vasectomized or sterile male partner, or sexual abstinence (defined as refraining from heterosexual intercourse), from Screening throughout the study and for at least 30 days after study drug administration. Reliance on abstinence from heterosexual intercourse is acceptable only if it is the subject's habitual practice.

Male subjects who are sexually active with a partner of child-bearing potential must either be sterile (vasectomy with history of a negative sperm count at least 90 days following the procedure); practice total abstinence from sexual intercourse as the preferred lifestyle (periodic abstinence is not acceptable); use a male condom with any sexual activity; or agree to use a birth control method considered to be appropriate by the Investigator (such as one of the methods identified above for female subjects of childbearing potential) from the time of Screening until 30 days after the last dose of study drug administration. Male subjects must agree not to donate sperm for a period of 30 days after the last dose of study drug administration.

Suspected or confirmed diagnosis of NASH fibrosis suggested by the historical data. Meet one of the following criteria that is consistent with NASH liver fibrosis:

Historical biochemical test for fibrosis: PRO-C3 >14 ng/mL; or ELF ≥9 (ELF is based on a historic value and is not obtained at screening; PRO-C3 is based on the historic PRO-C3 and not the screening PRO-C3).

Fibroscan with transient elastography ≥8.5 kPa; and controlled attenuation parameter ≥280 dB.m-1 (Fibroscan does not need to be repeated at screening if done at prescreening and/or a historical Fibroscan was done in the prior 3 months).

Historical liver biopsy obtained <2 years before expected randomization showing Stage 1B, 2 or 3 fibrosis with NASH (all 3 components) based on existing pathology review, with no significant change in body weight >5% or medication that might affect NAS or fibrosis stage.

NOTE: A biopsy that was > 6 months before the time of anticipated randomization is not eligible for study entry; a biopsy done ≤ 6 months is potentially eligible for study entry as a baseline biopsy only after confirmation of eligibility based on Inclusion #7 by the central pathology reviewer. If a Fibroscan is not used to meet Inclusion #5 (as in the case of historic eligible PRO-C3/ELF or liver biopsy), a baseline Fibroscan should be obtained prior to randomization. If other criteria are used to meet inclusion #5 and the MRI-PDFF ≥8% OR an eligible historical liver biopsy ≤ 6 months) has been confirmed by the central reader, and with confirmed no significant change in metabolic status since the time of that biopsy, the baseline Fibroscan can be KpA<8.5, CAP<280.

NOTE: Eligibility based on meeting Inclusion #5 should be determined based on historic medical and laboratory (PRO-C3/ELF, Fibroscan, liver biopsy) data and should be determined prior to informed consent and screening visit.

MRI-PDFF fat fraction ≥8% obtained during the screening period (Baseline MRI-PDFF).

NOTE: To be eligible to perform the screening MRI-PDFF (Baseline MRI-PDFF) a patient must first meet Criterion #5. An eligible MRI-PDFF with fat fraction ≥8% must be obtained prior to performing the baseline liver biopsy (Criterion #7). Patients with contraindications to an MRI-PDFF (eg, metal prosthetics or uncontrolled claustrophobia) examination or screened at an investigative site where MRI-PDFF is not available are eligible for a liver biopsy if they have a Fibroscan with CAP ≥280.

NOTE: An MRE and/or cT1 assessment will occur at sites with MRE equipment and/or multiparametric software. A historical MRI-PDFF ≥8% is eligible as a baseline MRI-PDFF if obtained ≤8 weeks prior to randomization.

NOTE: In cases with an eligible historical biopsy (≤ 6 months) confirmed by central reader, and with confirmation that there were no significant change in metabolic status since the time of the biopsy, MRI-PDFF <8% may be eligible. These cases would require review by Sponsor for confirmation.

Biopsy-proven NASH baseline liver biopsy) based on a liver biopsy obtained ≤6 months before anticipated date of randomization (if the biopsy is deemed acceptable for interpretation by the central reader) with fibrosis stage 1A/1C, 1B, 2, or 3 on liver biopsy and NAS of ≥4 with a score of at least 1 in each of the following NAS components:

Steatosis (scored 0 to 3)

Ballooning degeneration (scored 0 to 2)

Lobular inflammation (scored 0 to 3)

Fibrosis stage 1A/1C patients must also have elevated PRO-C3 (>14 ng/mL) obtained at Screening to be eligible to participate. Numbers of eligible F1A/F1C, F1B and F2 patients are defined in Number of Patients and Target Population.

NOTE: A historical biopsy obtained ≤6 months prior to anticipated date of randomization may be eligible as a baseline biopsy if the patient has had: (1) no significant change in metabolic status (diabetes control, lipid metabolism, and/or >5% weight gain or loss); (2) no change in the use of any prohibited medication(s) listed as exclusionary within 12 weeks prior to the anticipated date of randomization (3) no change in the use of a NASH therapeutic (eg therapeutic with documented impact on liver biopsy or GLP-1 agonist) since the time of the biopsy.

The historical biopsy must be evaluated for eligibility by the central pathology reader and confirmed as eligible. The biopsy should be sent for review by the central reader after screening labs and medical history/conmeds confirm I/E requirements are met. In cases where the MRI-PDFF is <8% but based on the local interpretation of the historical biopsy (≤6 months) the investigator believes that the subject would qualify then sending the biopsy for review by the central reader may be an option if discussed with the Sponsor.

NOTE: In cases with an eligible historical biopsy (≤6 months) confirmed by central reader, and with confirmed no significant change in metabolic status since the time of the biopsy, MRI-PDFF <8% may be eligible. These cases would require review by Sponsor for confirmation.

Estimated glomerular filtration rate (GFR) ≥45 by the Modification of Diet in Renal Disease 6-variable formula (MDRD-6).

Exclusion Criteria

Patients who meet any of the following criteria will be excluded from participation in the study. Patients who do not initially meet eligibility criteria may be retested or rescreened, based on Investigator judgment, to determine if they qualify to participate.

History of significant alcohol consumption for a period of more than 3 consecutive months within 1 year prior to Screening.

NOTE: Significant alcohol consumption is defined as equal to or greater than approximately 2 alcoholic drinks per day for males, and approximately 1.5 alcoholic drinks per day for females. One alcoholic drink is equal to 12 ounces (355 mL) of 5% alcohol by volume (ABV) beer, 5 ounces (148 mL) of 12% ABV wine, or 1.5 ounces (44.4 mL) of 40% ABV distilled spirits.

Regular use of drugs historically associated with NAFLD, which include, but are not limited, to the following: amiodarone, methotrexate, systemic oral glucocorticoids, tamoxifen, estrogens at doses greater than those used for hormone replacement or contraception, anabolic steroids except testosterone replacement, valproic acid, and known hepatotoxins for more than 4 weeks within the last 8 weeks prior to the initial Screening.

Thyroid diseases:

Active hyperthyroidism

NOTE: Patients with a history of hyperthyroidism are eligible to participate.

Untreated clinical hypothyroidism defined by thyroid stimulating hormone (TSH) >7 mIU/L with symptoms of hypothyroidism or >10 mIU/L without symptoms.

NOTE: TSH may be repeated once, and, if >10 mIU/L, even with normal FT4, patients may be stabilized on ≤75 µg thyroxine replacement therapy per day and rescreened for

eligibility. Patients with TSH >7 and <10 with no symptoms of hypothyroidism are eligible, and TSH may be monitored normally. Subclinical hypothyroidism and patients on stable thyroxine (T4) therapy up to 75 µg per day are eligible to participate. NOTE: Patients enrolled in the 'Subsequent 900' population or patients who have completed the Week 52 visit in the study may be on 100 µg/day. Other thyroid replacement therapies equivalent to up to 100 µg thyroxine are allowed (e.g. Armour thyroid).

Patients who have had a thyroidectomy and are on replacement thyroxine doses >75 µg per day are allowed.

History of bariatric surgery or intestinal bypass surgery within the 5 years prior to randomization or planned during the conduct of the study.

Weight gain or loss ≥5% total body weight within 12 weeks prior to randomization. (NOTE: This includes the Screening period.)

HbA1c >9.0%.

NOTE: Patients with HbA1c >8.0% and ≤10.0% should have documented efforts to control HbA1C to ≤8. If no prior documentation of efforts to control HbA1c, patients with HbA1C >8% and ≤9.0% may be treated with new or higher doses of existing diabetic medication(s) and continue screening. If screening HbA1C was >9.0% and a new antidiabetic therapy was initiated, they may have a repeat HbA1C 4 weeks after initiating a new antidiabetic therapy. Patients must be on stable treatment for all diabetes medications, including any new doses or medications, for ≥30 days prior to randomization.

NOTE: Insulin doses may be altered by up to 10% during the screening period. For screening HbA1C >9% and previous attempts to control HbA1C (no new therapy), HbA1C may be repeated once.

Glucagon-like peptide 1 [GLP-1] agonist therapy (eg, exenatide, liraglutide, lixisenatide, albiglutide, dulaglutide, semaglutide and albiglutide) unless stable dose for 24 weeks prior to biopsy. (NOTE: GLP-1 therapeutics may not be initiated or doses increased during the first 52 weeks of the study. However, GLP-1 therapeutics may be initiated or increased after the Week 52 visit is completed.)

Use of high dose vitamin E (>400 IU/day) unless stable for ≥24 weeks prior to an eligible screening liver biopsy. Vitamin E can be discontinued but dose cannot be increased during the first 52 weeks of the study.

Presence of cirrhosis on liver biopsy defined as stage 4 fibrosis.

Diagnosis of hepatocellular carcinoma (HCC).

MELD score ≥12, as determined at Screening, due to liver disease.

NOTE: MELD of ≥12 on screening labs must be the result of liver disease to be exclusionary, NOT isolated lab abnormalities such as elevated creatinine due to chronic kidney disease, INR abnormality secondary to anticoagulants or lab error, or bilirubin elevation due to Gilbert's Syndrome**.**

Hepatic decompensation or impairment defined as presence of any of the following:

History of esophageal varices, ascites, or hepatic encephalopathy

Serum albumin <3.5 g/dL, except as explained by non-hepatic causes

INR >1.4 unless due to therapeutic anticoagulants or laboratory error; NOTE: INR may be repeated once to reassess eligibility.

Total bilirubin >1.5 × upper limit of normal. NOTE: Patients with Gilbert's Syndrome are eligible with a total bilirubin above 1.5 × ULN if reticulocyte count is within normal limits (typically 0.5% to 2.5%),

hemoglobin is within normal limits (typically 13.5 to 17.5 g/dL for men; 12.0 to 15.5 g/dL for women), and direct bilirubin is <20% of total bilirubin.

Chronic liver diseases:

Primary biliary cholangitis

Primary sclerosing cholangitis

Hepatitis B positive (as defined in Appendix 3)

Hepatitis C as defined by presence of hepatitis C virus (HCV) antibody (HCV Ab) and positive HCV RNA (tested for known cured HCV infection, or positive HCV Ab at Screening). NOTE: Patients who are HCV antibody positive and HCV RNA negative who have a history of clearly documented HCV infection (history of positive HCV RNA) are eligible to participate if prior treatment for HCV was given, and they have a documented sustained virologic response (SVR) of at least two years prior to the baseline liver biopsy.

History or evidence of current active autoimmune hepatitis

History or evidence of Wilson's disease

History or evidence of alpha-1-antitrypsin deficiency

Evidence of genetic hemochromatosis (hereditary, primary)

Evidence of drug-induced liver disease, as defined on the basis of typical exposure and history

Known bile duct obstruction

Suspected or proven liver cancer.

Has an active autoimmune disease, including actively treated lupus, rheumatoid arthritis, inflammatory bowel disease, or autoimmune hepatitis, requiring systemic treatment within the past 12 weeks or a

documented history of clinically severe autoimmune disease, including autoimmune liver disease, or a syndrome that requires systemic steroids or immunosuppressive agents. NOTE: Patients with vitiligo or resolved childhood asthma/atopy would be an exception to this rule. Patients that require intermittent use of bronchodilators, topical, inhaled, or intranasal corticosteroids or local steroid injections are not excluded from the study.

Serum ALT >250 U/L.

NOTE: Given the intrinsic variability in ALT and AST in NASH patients, investigators should use the following guide in an attempt to establish a relatively stable baseline for ALT and AST. Investigator discretion is allowed. Documented historical (3 weeks to ≤ 6 months prior to study entry) ALT and AST levels consistent with the Screening ALT and AST values may help establish a stable baseline. This consistency may be established based on the following:

If the historical and Screening ALT and AST values are both ≤1.5 × ULN, there is no limit to the difference between the values.

Patients who do not have historical ALT and AST evaluations available will have their ALT and AST repeated during the Screening Period to help establish no worsening of >30% (both assessments during Screening period) with >2 weeks between assessments. If the historic ALT/AST are >1.5 × elevated and Screening ALT and AST are markedly improved (>50% decreased or normalized) relative to historic, then a third ALT/AST determination will be made during Screening to help establish a stable Baseline.

If at least 1 of the values is >1.5 × ULN and the second value is greater than the first value, the difference in the mean of ALT and AST values must be ≤30%. If the second value is greater than the first value by >30%, a third value assessed >2 weeks after the second value should be determined to help establish a lack of worsening trend in ALT/AST. If a worsening trend is confirmed (3 consecutive worsening values

with difference from first value and second value >30% and difference between second and third value >30%), patient will be a screen failure, but may be rescreened if ALT and AST stabilize.

Statins and/or other lipid-lowering therapies unless dose(s) is stable for ≥30 days prior to anticipated randomization. Statins must be taken in the evening for at least 2 weeks prior to randomization, and permitted statins include rosuvastatin up to 20 mg/day, atorvastatin up to 40 mg/day, pravastatin up to 40 mg/day, simvastatin up to 20 mg/day, pitavastatin up to 2 mg/day and lovastatin up to 40 mg/day. Other stable dyslipidemia therapies not specifically listed as excluded or dose-restricted such as PCSK9 inhibitors are allowed. Higher doses and other statins are excluded. Stable doses of bile acid sequestrants (eg, cholestyramine (Questran, Prevalite), colestipol (Colestid, Flavored Colestid), and colesevelam (Welchol)) are permitted only if taken at least 4 h after or at least 4 h before the dose of study drug.

Fenofibrate unless dose is stable for at least 6 weeks prior to anticipated randomization and unless taking fenofibrate for a history of and/or ongoing very high triglycerides (triglycerides >500 mg/dL).

> NOTE: Patients already enrolled who are taking fenofibrate even if not for very high triglycerides may remain in the study, because there are no safety concerns in most patients taking fenofibrate (Section 8.6.1).

Pioglitazone >15 mg per day. Pioglitazone treatment must be stable for ≥24 weeks prior to the eligible liver biopsy.

Platelet count <140,000/mm^3. Patients with platelets <140,000 and ≥120,000/mm^3 are eligible if Fib-4 <3.5.

Inability to safely obtain a liver biopsy.

History of biliary diversion.

Uncontrolled hypertension (either treated or untreated) defined as systolic blood pressure >170 mmHg or a diastolic blood pressure >100 mmHg at Screening.

New York Heart Association Class III or IV heart failure or known left ventricular ejection fraction <30%.

Uncontrolled cardiac arrhythmia.

Confirmed QT interval corrected using Fridericia's formula (QTcF) >450 sec for males and >470 msec for females at the Screening ECG assessment; At least 2/3 ECGs must show a prolongation and the average of the 3 ECGs must be prolonged to meet criteria for exclusion. Prolonged QTcF may be repeated and confirmed following machine calibration if needed.

NOTE: Patients with bundle branch block or other conditions in which a QTcF cannot be calculated are allowed.

Myocardial infarction, unstable angina, percutaneous coronary intervention, coronary artery bypass graft, or stroke within 12 weeks prior to randomization.

Use of illicit intravenous drugs within 5 years prior to randomization or a urine drug screen result positive for amphetamines, barbiturates, benzodiazepines, cocaine, methadone, opiates, or phencyclidine at Screening, unless a prescribed drug accounts for the positive test.

Active, serious medical disease with a likely life expectancy <2 years.

Participation in an investigational new drug trial in the 60 days or 5 half-lives, whichever is longer, prior to randomization. Patients previously treated with NASH therapeutics in an investigational trial are allowed if follow up liver biopsy at the end of trial continued to show active NASH fibrosis meeting eligibility criteria, and they have been off the NASH therapeutic for at least 24 weeks prior to expected randomization. If a potential NASH therapeutic studied revealed no safety issues, and, in fact, was not a

NASH therapeutic (no effect on liver biopsy compared to placebo) participation may occur 60 days or 5 half-lives, whichever is longer, after discontinuation of therapy.

History of major surgery (ie, surgery involving a risk to the life of the patient; specifically, an operation upon an organ within the cranium, chest, abdomen, or pelvic cavity) within 6 weeks prior to randomization.

History of cancer within the last 5 years (other than treated and believed to be cured basal or squamous cell carcinoma of the skin or resected carcinoma of the cervix).

Any other condition which, in the opinion of the Investigator, would impede compliance, hinder completion of the study, compromise the well-being of the patient, or interfere with the study outcomes.

Known immunocompromised status, including but not limited to individuals who have undergone organ transplantation, who are known to be positive for HIV, or who have recurrent or chronic systemic bacterial, fungal, viral, or protozoal infections.

Hypersensitivity to MGL-3196 or to any of the excipients or to placebo.

Compliance

We calculated compliance as the number of doses (pills) taken as a percentage of the number of days patients were identified as receiving IP. Compliance was defined in standard way, at least 80%. Overall compliance was high and no per protocol analysis was conducted.

Methodology for Central Pathologist Evaluation of Liver Biopsies

A liver biopsy review manual detailed the methodology for review of the liver biopsies. Briefly, glass slides were read by a central pathologist who was blinded to patient characteristics for eligibility at the time of screening based on the protocol definition of eligible biopsy. Approximately 4% of screening biopsies were considered technically inadequate. Biopsies were obtained using a 16-guage needle when possible and average biopsy length was ~22 mm at baseline and Week 52 as determined by both pathologists. Biopsy adequacy was confirmed for each biopsy by both pathologists. Week 52 biopsies blinded to patient ID were read by the central pathologist at the time the biopsy was obtained to determine if the biopsy showed cirrhosis. If cirrhosis was detected, the result was reported to the Sponsor and clinical site. The glass slides were digitized and stored at the digitization facility. For the primary analysis, each pathologist read all baseline and Week 52 glass slides as up to 100 slide batches of baseline and 100 slide Week 52 slides from the same patient to assure that baseline and Week 52 slides from the same patient were evaluated in batches at roughly the same time. Slides were read in the 6 months prior to the last patient Week 52 biopsy from the primary analysis population (966 patients with F1B, F2, or F3 fibrosis at baseline).

A secondary read was conducted of digitized images (read as patient pairs, blinded to time of biopsy). As a supportive analysis, consensus reads were conducted of digitized biopsies on which there was disagreement between the pathologists on responder status for the primary endpoints or the 2-stage fibrosis reduction.

Intrareader agreement (weighted kappa) was assessed between screening eligibility read and primary read for Path A, and for both pathologists between glass and digital reads. Interreader agreement was assessed for baseline and Week 52, respectively, for each component; steatosis 0.50, 0.60; lobular inflammation 0.30, 0.37; ballooning 0.34, 0.53 and fibrosis, 0.49, 0.65 (7 fibrosis

components). The weighted kappa statistic appeared to underestimate the degree of correlation between the two pathologists' scores, because Pathologist A consistently scored steatosis higher and Pathologist B consistently scored ballooning higher.

Baseline fibrosis stage was assigned the higher stage, when the two pathologists disagreed and scored F1B only when there was agreement between the two pathologists. Baseline F1a/F1c patients (n=84) were considered only for exploratory efficacy and safety analyses.

Patient's outcomes were classified and scored based on the following approach:

(1,1): Both pathologist's scores indicated the patient was a Responder – score of 1

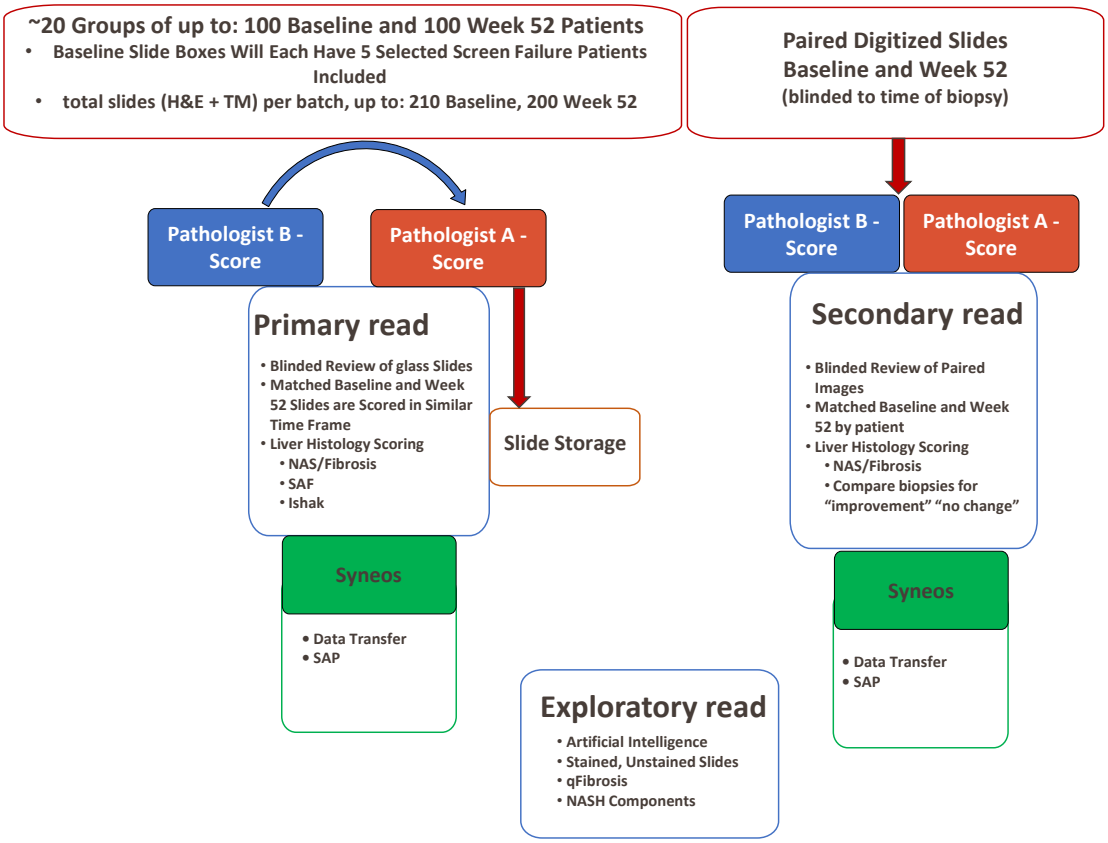
(1,0): Path A's scores indicated patient was a responder; Patht B's scores indicated a Non-Responder- score of 0.5

(0,1): Path B's scores indicated patient was a responder; Path A's scores indicated a Non-Responder- score of 0.5

(0,0): Both pathologist's scores indicated the patient was a Non-Responder – score of 0.

For each treatment, an average response rate was computed.

The primary analysis utilized a CMH test, stratified for type 2 diabetes status and baseline fibrosis stage. Estimates of risk difference and confidence interval were provided. Biopsies conducted at baseline and within 60 weeks of randomization and read as "adequate" on glass slides were considered valid. For patients without a Week 52 biopsy, response was imputed as "non-response".


~20 Groups of up to: 100 Baseline and 100 Week 52 Patients
• Baseline Slide Boxes Will Each Have 5 Selected Screen Failure Patients Included
• total slides (H&E + TM) per batch, up to: 210 Baseline, 200 Week 52
Paired Digitized Slides
Baseline and Week 52
(blinded to time of biopsy)
Pathologist B - Score
Pathologist A - Score
Pathologist B - Score
Pathologist A - Score
Primary read
• Blinded Review of glass Slides
• Matched Baseline and Week 52 Slides are Scored in Similar Time Frame
• Liver Histology Scoring
• NAS/Fibrosis
• SAF
• Ishak
Slide Storage
Secondary read
• Blinded Review of Paired Images
• Matched Baseline and Week 52 by patient
• Liver Histology Scoring
• NAS/Fibrosis
• Compare biopsies for "improvement" "no change"
Syneos
• Data Transfer
• SAP
Syneos
• Data Transfer
• SAP
Exploratory read
• Artificial Intelligence
• Stained, Unstained Slides
• qFibrosis
• NASH Components

Primary Liver Biopsy Analysis, Sensitivity Analyses Rationale and Methodology

The primary statistical analysis calculates within each pathologist the response rate using the Cochran-Mantel-Haenszel (CMH) model stratified for type 2 diabetes status and fibrosis stage. A single estimate of response difference from placebo is then obtained by averaging the difference obtained from each pathologist. The scoring of (1,1), (1,0), (0,1) and (0,0) provides higher weight to patients that are considered responders by both pathologists relative to when they disagree. The p-value was obtained from the CMH test using table scores to compare the active and placebo treatments. Patients with no valid biopsy within the Week 52 window that extended to Week 60 were considered non-responders for the Week 52 analysis, as were patients who experience a composite clinical endpoint (e.g., liver transplant, death) prior to their Week 52 biopsy.

Consensus Reads of Digitized Images of Glass Slides

Consensus reads of digitized images of glass slides by the two central pathologists were conducted in cases where the two pathologists scores disagreed as to whether there was a response for either NASH resolution (ballooning 0,1; 2-point NAS reduction and no worsening of fibrosis) OR ≥1 stage fibrosis reduction with no worsening of NAS (primary endpoints) OR a 2-stage reduction in fibrosis with no worsening of NAS. The study pathologists remained blinded to all clinical data, patient ID, and slide/image identity or time of biopsy. Study personnel remain blinded to patient level data and detailed group analyses. In total 387 total assessments took place requiring four Zoom meetings between the pathologists. The consensus meetings were conducted by an unblinded reviewer who showed the pathologists the initial scores on a spread sheet, and then the pathologists decided on the consensus score for the disagreed component.

Additional Sensitivity and Supportive Analyses of the Dual Primary Endpoints

Multiple additional sensitivity and supportive analyses were conducted on the dual primary endpoints.

Tipping Point Analyses: This analysis imputed missing placebo responses for NASH resolution (in increments from 0% to 100% placebo imputation) and fibrosis response (in increments from 0% up to 35% placebo imputation) at Week 52 as successful responses in order to determine how many missing responses (which were counted in the primary analysis labels as non-responders) could be imputed before losing statistical significance. Of note, the 100-mg treatment group continued to demonstrate significant improvement in NASH resolution (up to 100% missing placebo imputation) (nominal p = 0.0246) and fibrosis response (up to 35% missing placebo imputation) (nominal p = 0.0378) (Table S6).

Multiple Imputation Analysis: Multiple imputation analyses were conducted where data for patients with missing response at Week 52 were imputed under the missing at random assumption. Data for patients with missing response data at Week 52 were imputed under the missing at random assumption by running a simulation (100 times) which produces a correlated pair of binary 0/1 data for each patient that represents the patient's response status for each pathologist. The statistics were then calculated for each imputation using the same approach as for the primary endpoint. The normalized results from each dataset were combined using Rubin's rule. (Table S5).

General Estimating Equation (GEE) Model Analysis: A generalized estimation equation (GEE) analysis was performed as a sensitivity analysis. This approach treated the biopsy scores from the two pathologists as repeated measures (i.e. correlated binary outcomes) within a patient. Baseline diabetes status, fibrosis stage and interaction term between pathologist and treatment were also included in the model. Sensitivity analyses were performed using a GEE model in which the odds ratios favored both the 100- and 80-mg resmetirom treatment groups compared with placebo for NASH resolution (Table S5A and Table S5B).

Summary of Observed Cases Using In-window Week 52 Paired Liver Biopsies: Supportive analyses were performed on patients with in-window paired Week 52 biopsies. Missing Week 52 biopsies were not imputed.

Statistical Analysis of Change in LDL-C at Week 24

Within the manuscript, we provide analyses based on an ANCOVA model after having performed single imputation for LDL-C as described in the SAP. The ANCOVA model estimated LSMeans using treatment and baseline LDL-C as covariate.

To this end, the Week 24 LDL-C assessment and assessments at any visit in which LDL-C is measured (Weeks 4, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, and 52) that has been impacted by kit shortage at the prior visit (unavailability of IP during the 4 weeks preceding the visit due to the COVID-19 pandemic as indicated on the eCRF), were considered missing and were imputed from an unaffected assessment at surrounding visits or using data from subjects unaffected by this intercurrent event as described in Section 11.2 of the SAP. In summary, single imputation utilizing adjacent and valid LDL-D measurements were imputed.

For missing lipid data that are still missing after the single imputation approach described, those missing lipid data were imputed using the non-missing lipid values (including the singly imputed data) based on MAR-based MI. When applying the MAR-based MI, data were imputed separately by randomized treatment group and baseline stratification factors.

Supplementary Figures and Tables

Figure S1. Study Design

CAP, controlled attenuation parameter; LDL-C, low-density lipoprotein cholesterol; MRE, magnetic resonance elastography; MRI-PDFF, magnetic resonance imaging-proton density fat fraction; VCTE, vibration-controlled transient elastography.

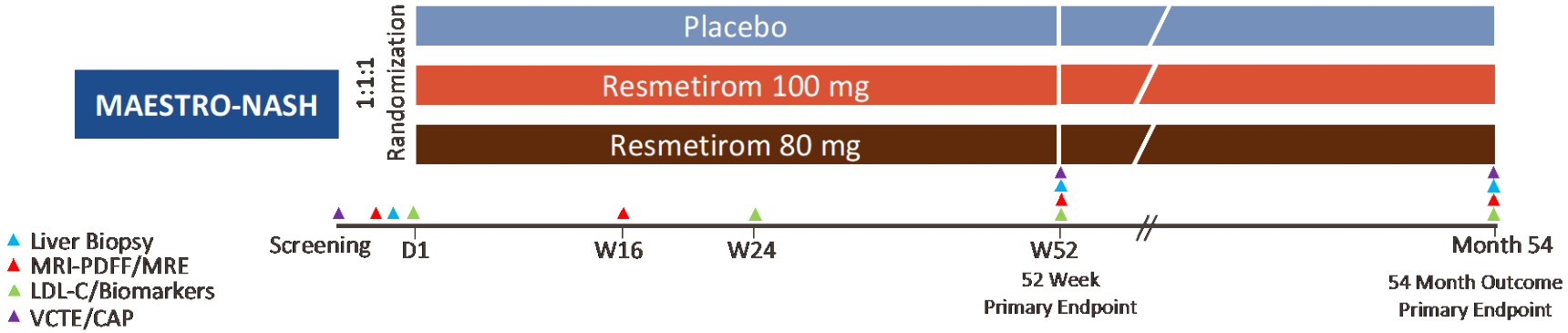


Figure S2. Statistical Testing Process for Primary and Key Secondary Endpoints at Week 52

LDL-C, low-density lipoprotein cholesterol.

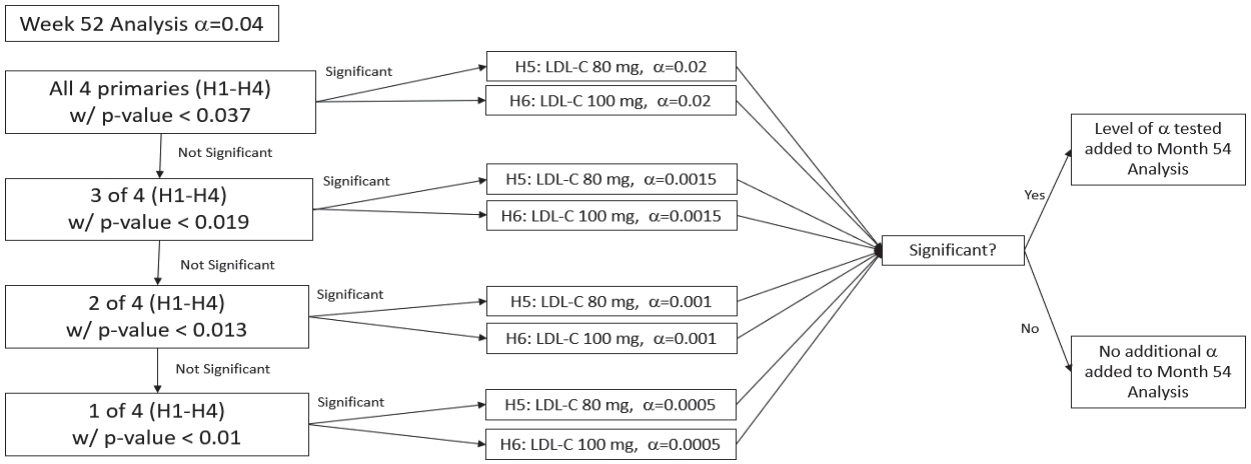


Figure S3. Patient Disposition

AE, adverse event; LTFU, lost to follow up. The primary reasons for screen failure included biopsy, withdraw of consent, MRI-PDFF <8%, HbA1c >9. The exploratory F1 group included baseline F1a/F1c patients (n = 84) that were considered only for exploratory efficacy and safety analyses. These patients received treatment but as prespecified in the statistical analysis plan, were not included in the primary analysis population.

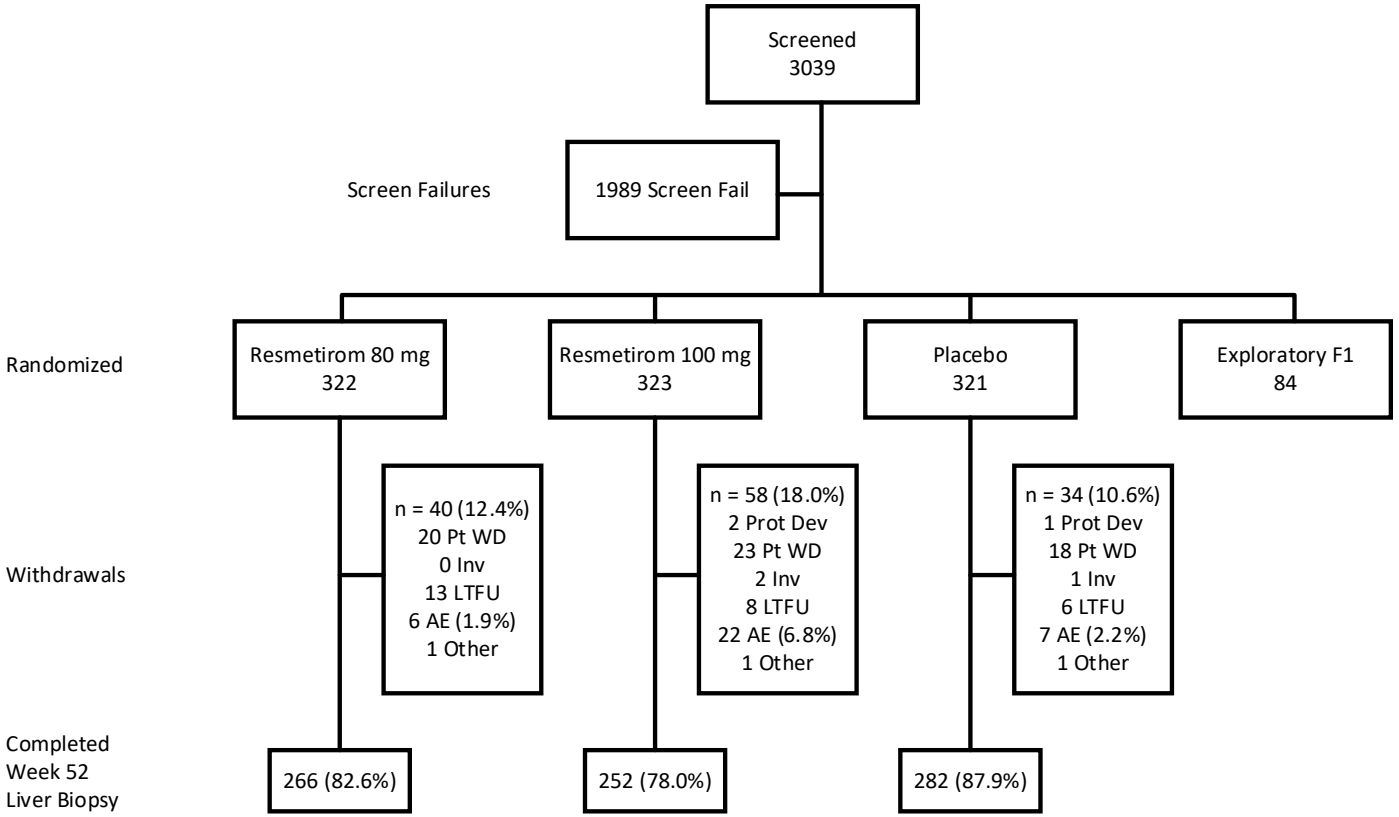
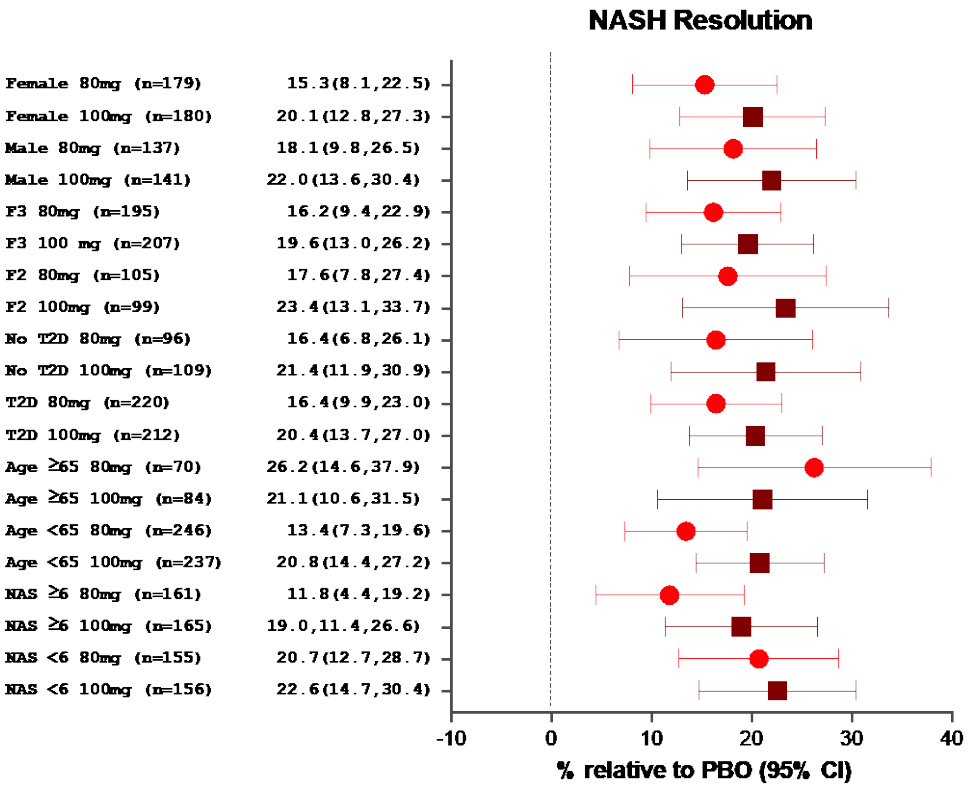


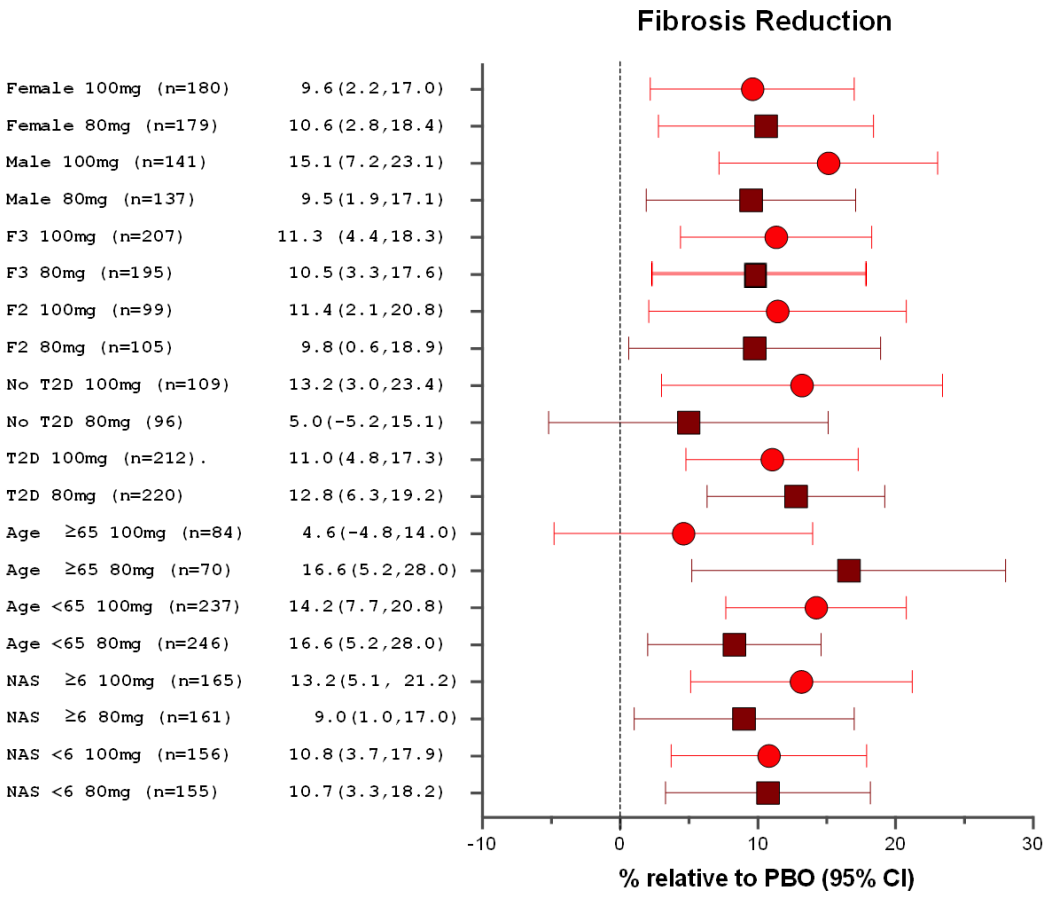
Figure S4. Subgroup analyses of the dual primary endpoints at Week 52

Resolution of nonalcoholic steatohepatitis (A&C) and improvement in fibrosis (B&D). Data are reported for the primary analysis population (n=955, after removal of the COVID impacted biopsies outside the Week 60 window). Eleven patients had a delay in their Week 52 biopsy due to COVID-19 biopsy site closure or related reasons and were removed from the analysis. Resolution of nonalcoholic steatohepatitis is defined as achievement of a ballooning score of 0, inflammation score of 0 or 1, and ≥2-point reduction in the nonalcoholic fatty liver disease activity score with no worsening of fibrosis. Fibrosis improvement is defined as achievement of ≥1-stage reduction in fibrosis with no worsening of the nonalcoholic fatty liver disease activity score. A 1-point improvement in fibrosis would be a change to F1A or F1C from F2 (a change of F2 to F1B is not considered a 1-point improvement). MGL-3196, resmetirom. Forest plots include prespecified subgroups with minor modifications. Body weight subgroups based on <=200, >200 pounds or BMI<=35, >35 were not informative. PDFF reduction in resmetirom groups are compared to all placebo patients with any Week 52 PDFF; % SHBG CFB in resmetirom groups is compared to all placebo patients with a Week 52 SHBG. A posthoc subgroup of <=100kg, >100 kg is shown in Table S9 that also includes subgroups for >=30% PDFF at Week 16 , region (US-ExUS), weight gain >=5% F2/F3, F1B. Please note that subgroup analyses by statin use and NAS were *post-hoc* analyses. Confidence interval widths have not been adjusted for multiplicity and may not be used for hypothesis testing.

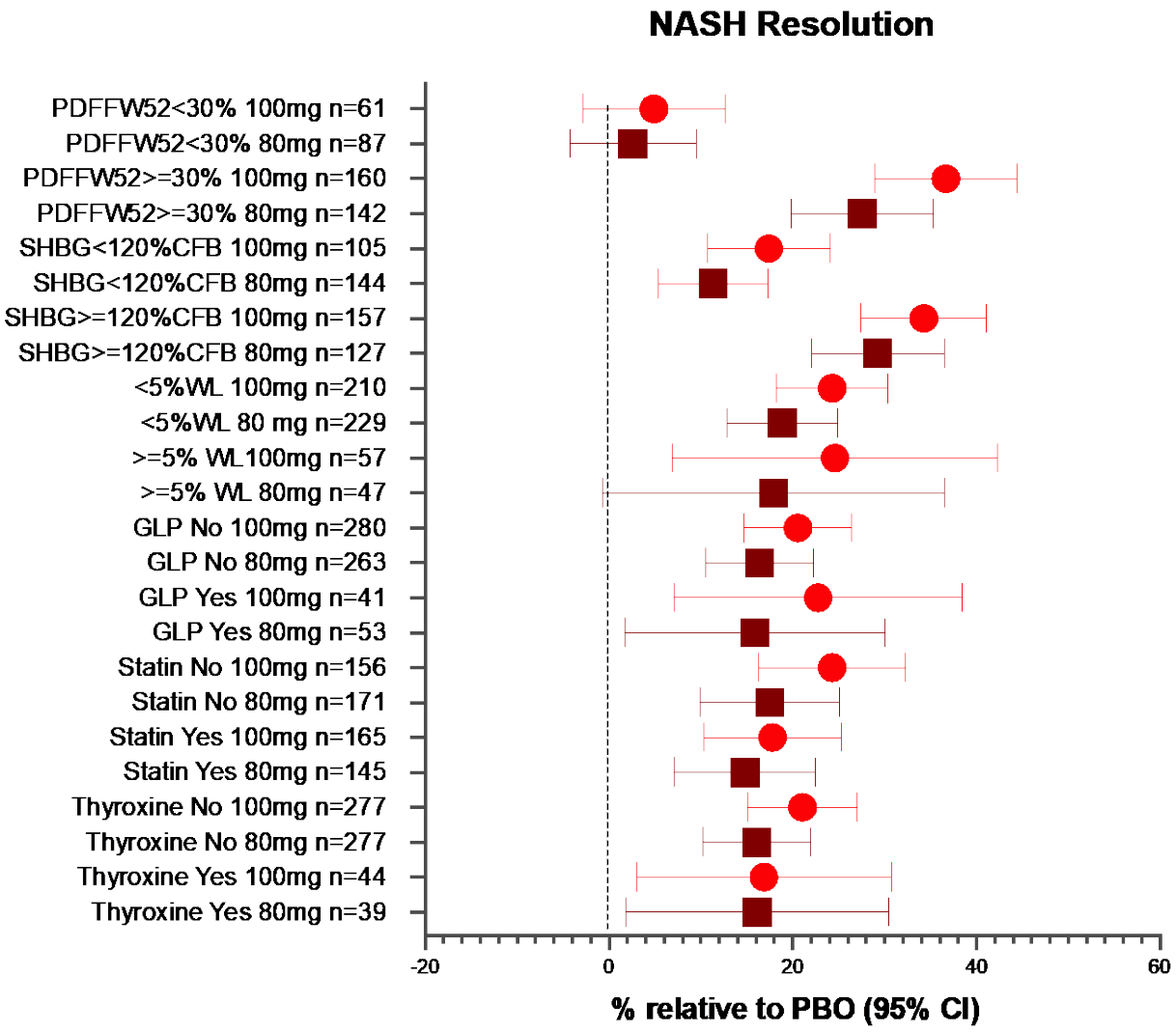
A


NASH Resolution
Female 80mg (n=179) 15.3(8.1,22.5)
Female 100mg (n=180) 20.1(12.8,27.3)
Male 80mg (n=137) 18.1(9.8,26.5)
Male 100mg (n=141) 22.0(13.6,30.4)
F3 80mg (n=195) 16.2(9.4,22.9)
F3 100 mg (n=207) 19.6(13.0,26.2)
F2 80mg (n=105) 17.6(7.8,27.4)
F2 100mg (n=99) 23.4(13.1,33.7)
No T2D 80mg (n=96) 16.4(6.8,26.1)
No T2D 100mg (n=109) 21.4(11.9,30.9)
T2D 80mg (n=220) 16.4(9.9,23.0)
T2D 100mg (n=212) 20.4(13.7,27.0)
Age ≥65 80mg (n=70) 26.2(14.6,37.9)
Age ≥65 100mg (n=84) 21.1(10.6,31.5)
Age <65 80mg (n=246) 13.4(7.3,19.6)
Age <65 100mg (n=237) 20.8(14.4,27.2)
NAS ≥6 80mg (n=161) 11.8(4.4,19.2)
NAS ≥6 100mg (n=165) 19.0,11.4,26.6)
NAS <6 80mg (n=155) 20.7(12.7,28.7)
NAS <6 100mg (n=156) 22.6(14.7,30.4)
-10
0
10
20
30
40
% relative to PBO (95% CI)

B


Fibrosis Reduction
Female 100mg (n=180) 9.6(2.2,17.0)
Female 80mg (n=179) 10.6(2.8,18.4)
Male 100mg (n=141) 15.1(7.2,23.1)
Male 80mg (n=137) 9.5(1.9,17.1)
F3 100mg (n=207) 11.3 (4.4,18.3)
F3 80mg (n=195) 10.5(3.3,17.6)
F2 100mg (n=99) 11.4(2.1,20.8)
F2 80mg (n=105) 9.8(0.6,18.9)
No T2D 100mg (n=109) 13.2(3.0,23.4)
No T2D 80mg (96) 5.0(-5.2,15.1)
T2D 100mg (n=212). 11.0(4.8,17.3)
T2D 80mg (n=220) 12.8(6.3,19.2)
Age ≥65 100mg (n=84) 4.6(-4.8,14.0)
Age ≥65 80mg (n=70) 16.6(5.2,28.0)
Age <65 100mg (n=237) 14.2(7.7,20.8)
Age <65 80mg (n=246) 16.6(5.2,28.0)
NAS ≥6 100mg (n=165) 13.2(5.1, 21.2)
NAS ≥6 80mg (n=161) 9.0(1.0,17.0)
NAS <6 100mg (n=156) 10.8(3.7,17.9)
NAS <6 80mg (n=155) 10.7(3.3,18.2)
-10
0
10
20
30
% relative to PBO (95% CI)

c


NASH Resolution
PDFFW52<30% 100mg n=61
PDFFW52<30% 80mg n=87
PDFFW52>=30% 100mg n=160
PDFFW52>=30% 80mg n=142
SHBG<120%CFB 100mg n=105
SHBG<120%CFB 80mg n=144
SHBG>=120%CFB 100mg n=157
SHBG>=120%CFB 80mg n=127
<5%WL 100mg n=210
<5%WL 80 mg n=229
>=5% WL100mg n=57
>=5% WL 80mg n=47
GLP No 100mg n=280
GLP No 80mg n=263
GLP Yes 100mg n=41
GLP Yes 80mg n=53
Statin No 100mg n=156
Statin No 80mg n=171
Statin Yes 100mg n=165
Statin Yes 80mg n=145
Thyroxine No 100mg n=277
Thyroxine No 80mg n=277
Thyroxine Yes 100mg n=44
Thyroxine Yes 80mg n=39
-20
0
20
40
60
% relative to PBO (95% CI)

D

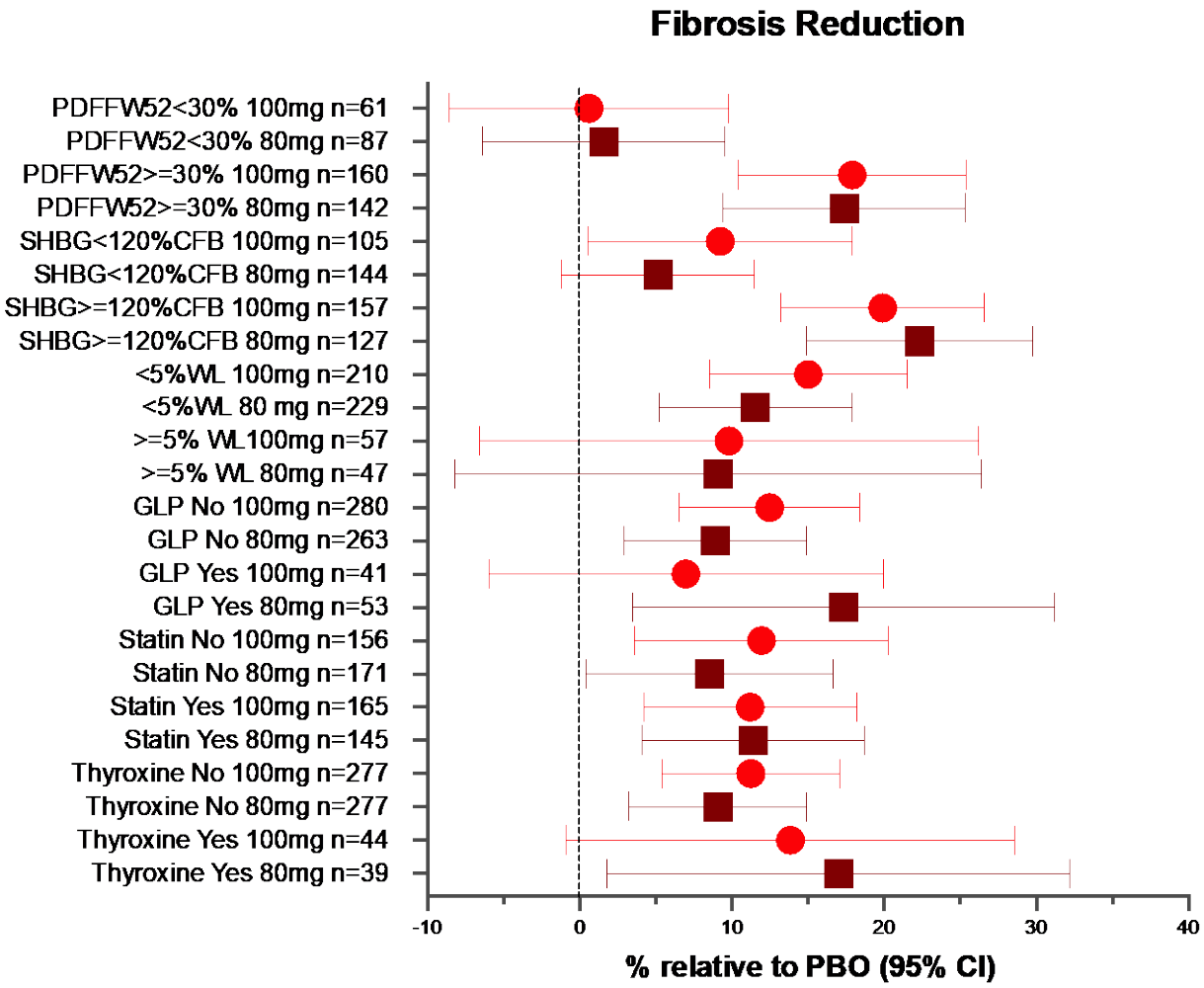

Fibrosis Reduction
PDFFW52<30% 100mg n=61
PDFFW52<30% 80mg n=87
PDFFW52>=30% 100mg n=160
PDFFW52>=30% 80mg n=142
SHBG<120%CFB 100mg n=105
SHBG<120%CFB 80mg n=144
SHBG>=120%CFB 100mg n=157
SHBG>=120%CFB 80mg n=127
<5%WL 100mg n=210
<5%WL 80 mg n=229
>=5% WL100mg n=57
>=5% WL 80mg n=47
GLP No 100mg n=280
GLP No 80mg n=263
GLP Yes 100mg n=41
GLP Yes 80mg n=53
Statin No 100mg n=156
Statin No 80mg n=171
Statin Yes 100mg n=165
Statin Yes 80mg n=145
Thyroxine No 100mg n=277
Thyroxine No 80mg n=277
Thyroxine Yes 100mg n=44
Thyroxine Yes 80mg n=39
-10
0
10
20
30
40
% relative to PBO (95% CI)

Figure S5. Percentage of Patients Who were F1B or F2 at Baseline with Worse (progressed to ≥F3), Stable (No Change), or Improved Fibrosis Stage at Week 52 Based on Liver Biopsy

In patients with a baseline and eligible Week 52 biopsy. 80 mg, 100 mg: resmetirom. The two pathologists' assessments were similar and were averaged to generate a single output.

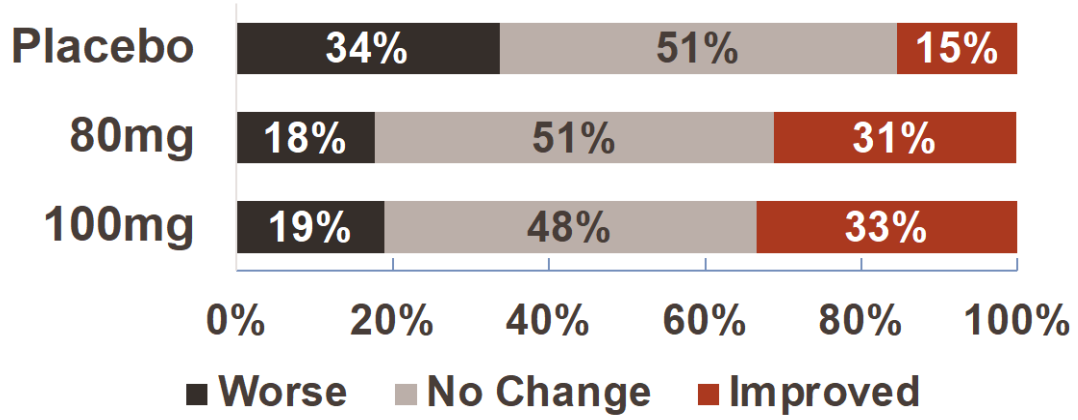


Figure S6. Percent of Patients with Worsened, Stable (No Change), or Improved Individual Components of the Nonalcoholic Fatty Liver Disease Activity Score (Ballooning, Inflammation, Steatosis)

In patients with a baseline and eligible Week 52 biopsy. The two pathologists' assessments were similar and were averaged to generate a single output.

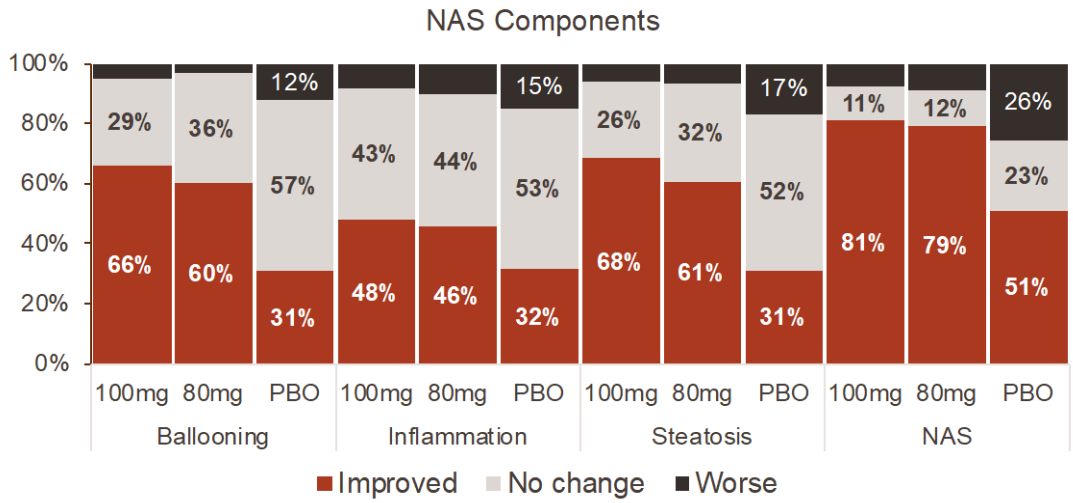


Figure S7. Percent Change from Baseline in Lipids and Lipoproteins at Weeks 24 and 52

ApoB, apolipoprotein B; ApoCIII, apolipoprotein CIII; LDL-C, low-density lipoprotein cholesterol; Lp(a), lipoprotein a; non-HDL-C, non–high-density lipoprotein cholesterol. 80mg, 100 mg, resmetirom.

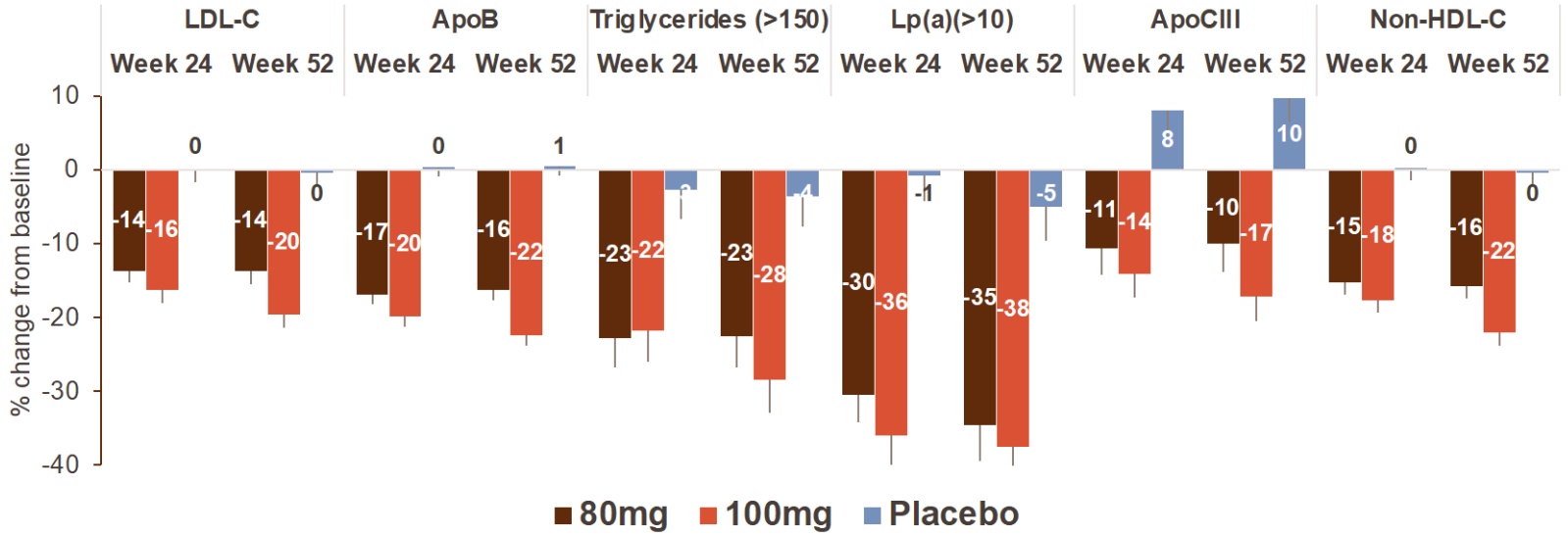


Figure S8. Percent Change from Baseline in Hepatic Fat as Measured by Magnetic Resonance Imaging-Proton Density Fat Fraction at Weeks 16 and 52, and Steatosis as Measured by FibroScan Controlled Attenuation Parameter at Week 52

80 mg, 100 mg: resmetirom, based on observed data, patients with a baseline and Week 52 assessment.

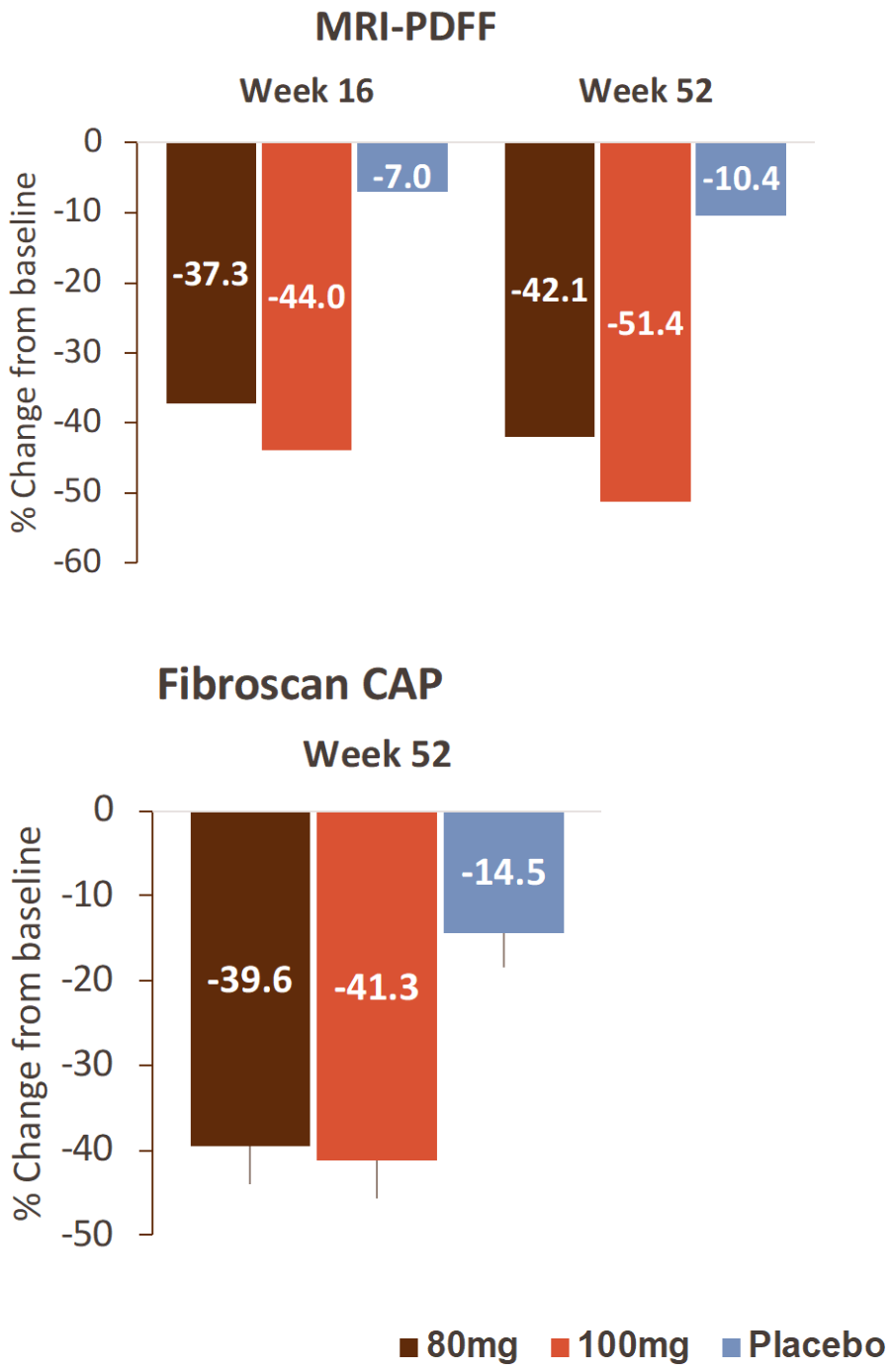


Figure S9. Percentage of Patients Achieving a ≥25% Reduction from Baseline in Liver Stiffness as Measured by FibroScan Vibration-controlled Transient Elastography at Week 52

Based on observed data, patients with a baseline and Week 52 assessment.

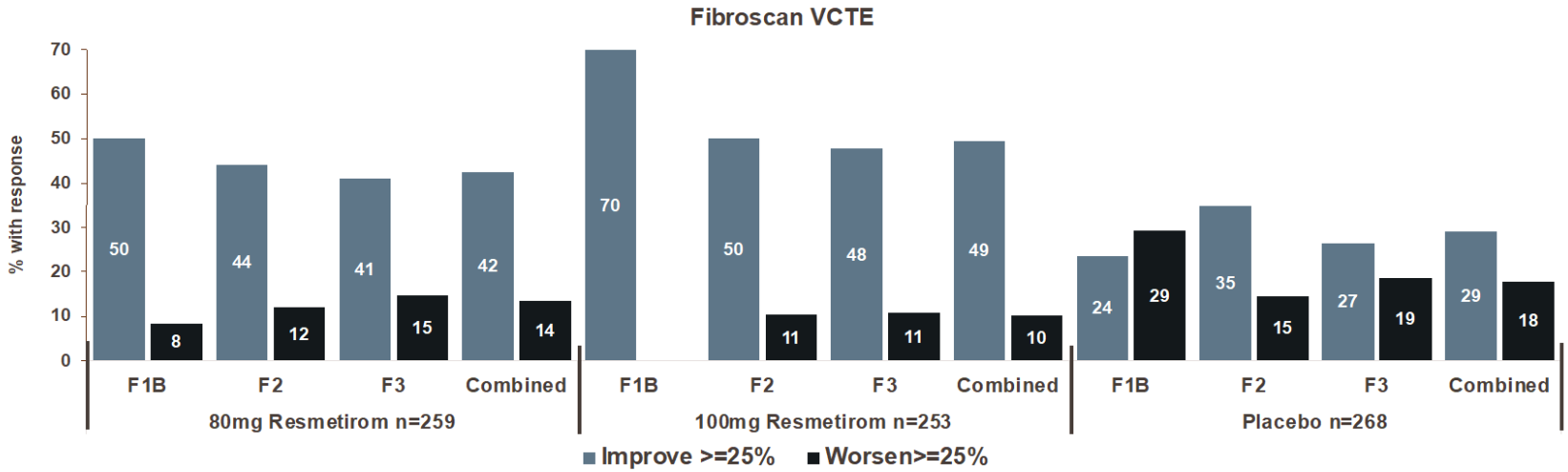


Figure S10. Improvement or Worsening from Baseline in Liver Stiffness As Measured by Magnetic Resonance Elastography at Week 52

80 mg, 100 mg resmetirom, based on observed data, patients with a baseline and Week 52 assessment.

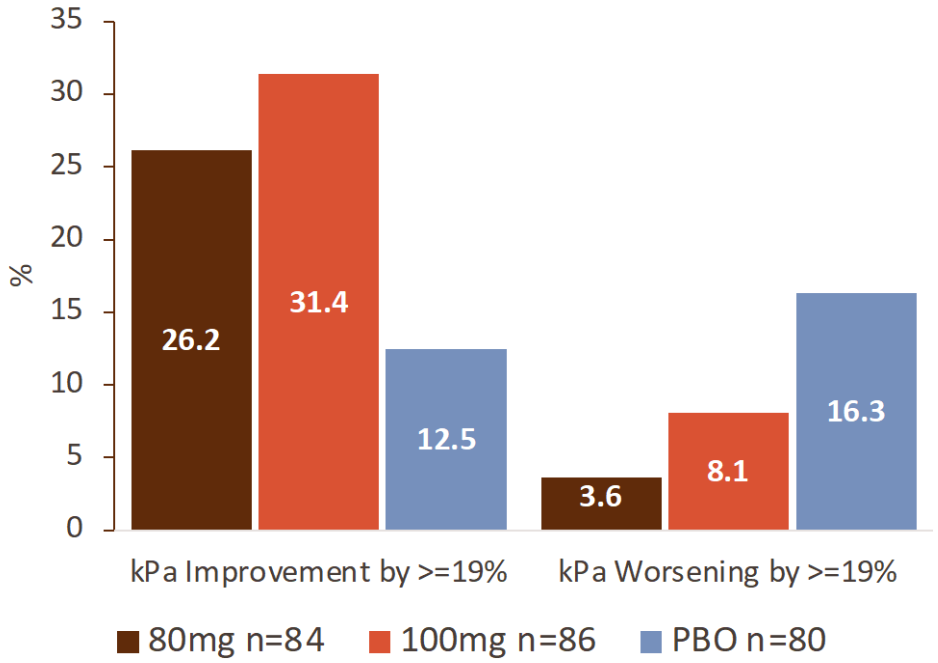


Figure S11. Percent Change from Baseline in Liver and Spleen Volume at Weeks 16 and 52

80 mg, 100 mg: resmetirom, based on observed data, patients with a baseline and Week 52 assessment.

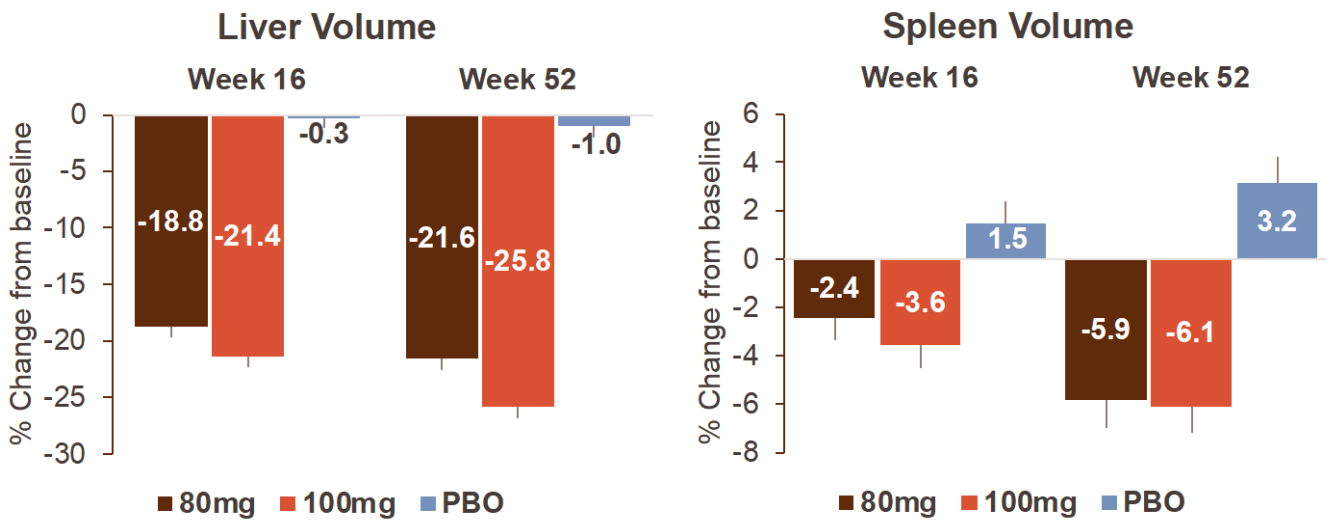


Figure S12. Change from Baseline in the Enhanced Liver Fibrosis Score, P3NP, and TIMP-1

80mg, 100mg: resmetirom. Based on observed data

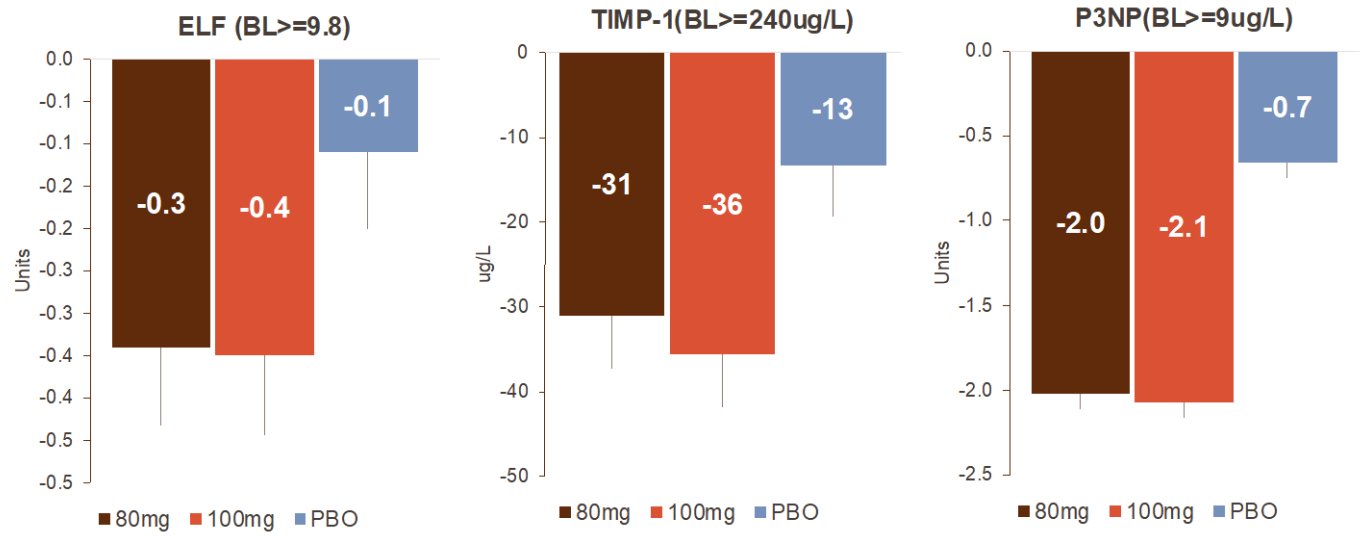


Figure S13. Time to Onset of First Gastrointestinal Adverse Event

MGL-3196, resmetirom

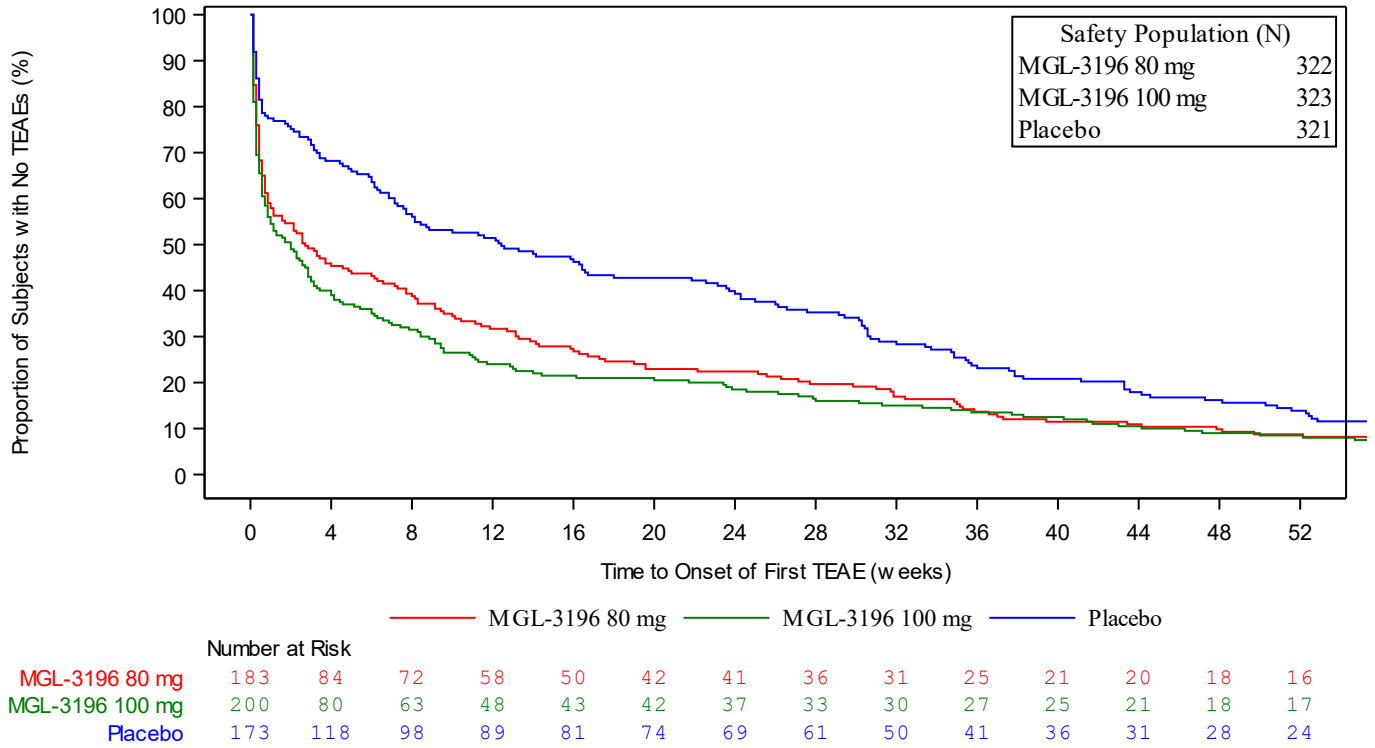
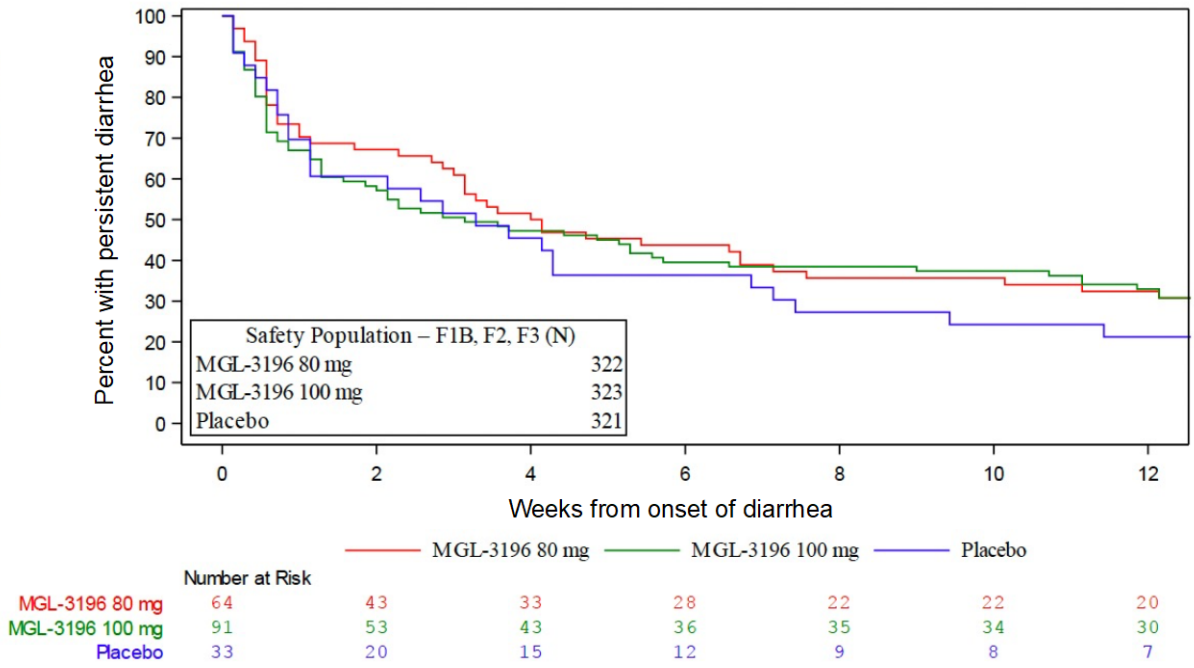


Figure S14. Duration of Diarrhea

MGL-3196, resmetirom, duration of diarrhea reported in first 12 weeks of randomization.



Supplementary Tables

Table S1. Complete List of Endpoints/Objectives for 52 Week Analyses

Type of Endpoint	Endpoint	Reported	Rationale
Dual Primary (#1)	Proportion of NASH Resolution Responders as assessed by two pathologists at Week 52.	Included	N/A
Dual Primary (#2)	Proportion of Fibrosis Responders as assessed by two pathologists at Week 52.	Included	N/A
Key Secondary	Percent change from baseline in directly measured LDL-C at Week 24.	Included	N/A
Secondary	Proportion of patients meeting each of the criteria below at Week 52: • At least a 2-point improvement in NAS with at least 1-point improvement in ballooning or lobular inflammation with no worsening of fibrosis. • At least a 2-point improvement in NAS with at least 1-point improvement in ballooning or lobular inflammation and at least a 1-point improvement in fibrosis. • An improvement in each histologic NAS component (ballooning, inflammation, steatosis) by at least 1 point; or improvement by at least 1 point in both ballooning and inflammation with an MRI-PDFF response (≥30% relative fat reduction) at Week 16 or at Week 52 if no data available at Week 16. • The resolution of fibrosis (reduction to F0). • A 2-stage Fibrosis Responders (a ≥2-point reduction in fibrosis patients with no worsening of NAS) in patients with baseline fibrosis F2 or more severe. • A composite of NASH Resolution Responder and Fibrosis Responder. • No worsening of fibrosis is defined as no progression ≥1-stage (for patients with an F1B baseline stage, a change to F2 is not considered worsening). This analysis includes patients with paired biopsies (i.e., baseline and Week 52).	Included	N/A
Secondary	Change from baseline to week 52 in NAS component score for each of the NAS components (improvement, worsening, no change [change is defined by at least a 1-point change]), NAS (total of the 3 components), and individual components, lobular inflammation, ballooning steatosis and fibrosis and the change from baseline in fibrosis (assuming a value of 1.8 for F1B). A change from F2 to F1B is not considered a decrease and from F1B to F2 is not considered	Included	N/A

Type of Endpoint	Endpoint	Reported	Rationale
	an increase. Based on the paired liver biopsy population for each pathologist (based on glass slide [primary analysis]). For fibrosis, only the subset of patients with baseline F1B or F2 (combined) to be evaluated.		
Secondary	Absolute change and percent change from baseline to Week 52 in MRI-PDFF in all patients with baseline and a Week 52 assessment.	Included	N/A
Secondary	Proportion of patients at Week 16 and Week 52 with ≥30% or ≥50% relative reduction from baseline in MRI-PDFF (analysis includes patients with paired data; that is a baseline assessment and an assessment at Week 16 and/or at baseline and at Week 52).	Included	N/A
Secondary	Absolute change and percent change from baseline to Week 48 in liver parameters, including ALT, AST, and GGT, in patients with baseline ALT ≥30 IUL/ml.	Included	N/A
Secondary	Percent change from baseline at Week 24 and Week 52 in directly measured LDL-C , ApoB, triglycerides in patients with baseline triglycerides >150 mg/dL, ApoCIII, non-HDL-C, and Lp(a) in patients with baseline Lp(a) >10 nmol/L.	Included	N/A
Secondary	Absolute change from baseline in directly measured LDL-C in patients with baseline LDL-C ≥100 mg/dL, ApoB, ApoB in patients with baseline LDL-C ≥100 mg/dL, triglycerides in patients with baseline triglycerides >150 mg/dL, Lp(a) in patients with baseline Lp(a) >10 nmol/L. Note: Percent change from baseline in LDL-C at Week 24 is a key secondary endpoint.	Not included	%CFB only is included as the more relevant measurement
Secondary	Proportion of patients at Week 24 and Week 52 with directly measured LDL-C >100 mg/dL at Baseline who achieve <100 mg/dL.	Not included	Endpoint not relevant to study objectives
Secondary	Proportion of patients at Week 24 and Week 52 with >70 mg/dL directly measured LDL-C at Baseline who achieve <70 mg/dL.	Not included	Endpoint not relevant to study objectives
Secondary	Percent change from baseline at Week 24 and Week 52 in HDL-C, ApoCIII, and lipoprotein particles.	Included	NA
Secondary	Absolute change and percent change from baseline to Week 52 in NASH inflammation and fibrosis biomarkers, including: adiponectin, reverse T3, CK-18, and ELF test with 3 direct components (ELF baseline ≥9.8; PIIINP ≥9 µg/L; TIMP-1 ≥240 µg/L; HA ≥50 µg/L).	Included	NA
Secondary	Absolute change from baseline to Week 52 for NAFLD/NASH CLDQ, SF-LDQOL, and WPAI-NASH (QOL assessments).	Not included	This subanalysis will be presented separately

Type of Endpoint	Endpoint	Reported	Rationale
Secondary	Proportion of patients with baseline MRE as ≥2.9 kPa receiving serial (baseline and at least one of the following: Week 16, Week 52) MRE with ≥19% reduction from baseline and the proportion with ≥19% increase from baseline. F3, F1B, and F2 assessed separately.	Included	N/A
Secondary	Proportion of patients receiving serial (baseline and Week 52) FibroScan with ≥25% and ≥30% reduction from baseline in FibroScan VCTE over time. F3, F1B, and F2 assessed separately.	Included	N/A
Secondary	Absolute change from baseline at Week 52 in FibroScan CAP.	Included	N/A
Secondary	Percent change from baseline at Week 16 and Week 52 in liver volume in patients with baseline MRI-PDFF and at least one post baseline MRI-PDFF.	Included	N/A
Exploratory	Absolute change from baseline at Week 52 in other metabolic, liver, and cardiovascular assessments, including: • Markers of insulin resistance and glucose homeostasis such as adiponectin, HOMA-IR, HbA1c, glucose, and insulin. • Body weight, BMI. • Systolic and diastolic blood pressure. • Heart rate as determined by ECG.	Included	N/A
Exploratory	Agreement between: • The 2 pathologists in the assessment of NASH CRN scores (ballooning, inflammation, steatosis, fibrosis), NAS score, NASH Resolution Response, and Fibrosis Response. • The eligibility and Baseline liver biopsy assessments of NASH CRN scores (ballooning, inflammation, steatosis, fibrosis), NAS score, NASH Resolution Response, and Fibrosis Response for each pathologist.	Included	N/A
Exploratory	Proportion of patients at Week 52 defined as any improvement, no change, or any worsening in fibrosis stage.	Included	N/A
Exploratory	Change from baseline in fracture risk as assessed by T and Z scores at each of 3 sites: hip, femoral neck, and spine based on DEXA scan.	Included	
Exploratory	Adjudicated Events • MACE as adjudicated by an independent EAC in all patients. MACE includes cardiovascular mortality, myocardial infarction, cerebral vascular accident. • Events adjudicated as positive by an adjudication committee related to cardiovascular disease including a composite of 1) heart failure events, 2) hospitalization for unstable angina, and 3) urgent or emergent coronary revascularization.	Included	

Type of Endpoint	Endpoint	Reported	Rationale
	Drug-induced liver injury (DILI) events adjudicated as positive by an adjudication committee.		
Exploratory	DILI events (adjudicated)	Included	

Table S2. Additional Demographic and Baseline Characteristics (Primary Analysis Population)*

	Resmetirom 80 mg (n = 322)	Resmetirom 100 mg (n = 323)	Placebo (n = 321)
HbA1c, %	6.6 ± 1.1	6.6 ± 1.1	6.5 ± 1.0
Total cholesterol, mg/dL	179.6 ± 43.4	176.9 ± 46.0	180.0 ± 50.0
HDL-C, mg/dL	43.8 ± 12.6	44.0 ± 12.9	43.8 ± 13.3
ApoB, mg/dL	98.4 ± 27.8	95.9 ± 27.8	97.8 ± 32.0
Lp(a), nmol/L	44.7 ± 61.1	43.8 ± 60.8	42.2 ± 62.7
Enhanced liver fibrosis score	9.7 ± 0.89	9.8 ± 0.86	9.7 ± 0.86
Triglycerides, mg/dL	189.2 ± 112.5	188.7 ± 153.8	184.1 ± 125.8
Alkaline phosphatase, U/L	74.9 ± 27.1	73.9 ± 23.0	71.5 ± 23.7
Bilirubin, mg/dL†	0.63 ± 0.27	0.66 ± 0.32	0.66 ± 0.31
Platelets, 10^9/L†	236.6 ± 67.9	230.6 ± 59.1	233.6 ± 60.4
Albumin, g/dL†	4.4 ± 0.32	4.3 ± 0.27	4.4 ± 0.29
HOMA-IR	11.9 ± 11.8	10.6 ± 8.3	11.0 ± 12.3
GLP-1 RA, no. (%)	54 (16.8)	41 (12.7)	42 (13.1)
SGLT2i, no. (%)	55 (17.1)	39 (12.1)	36 (11.2)
Insulin, no. (%)	40 (12.4)	41 (12.7)	37 (11.5)
Statin, no. (%)	149 (46.3)	166 (51.4)	158 (49.2)

*Plus-minus signs are mean ±SD. †Safety population.
ApoB, apolipoprotein B; GLP-1 RA, glucagon-like peptide-1 receptor agonist; HbA1c, hemoglobin A1c; HDL-C, high-density lipoprotein cholesterol; HOMA-IR, homeostasis model assessment-estimated insulin resistance; Lp(a), lipoprotein (a); SD, standard deviation; SGLT2i, sodium/glucose cotransporter-2 inhibitor.

Table S3. Representativeness of Study Participants.

Category	Details
Disease	Nonalcoholic steatohepatitis (NASH)
Special considerations related to	
Age	NASH is prevalent in ages 45-65
Gender and sex	NASH is common in both men and women
Geography	NASH is common globally, with estimated 30% of US population has NAFLD
Race or ethnic group	NASH is highly prevalent in Hispanic and Latino communities and is less common in black patients
Other special considerations	NASH is more common in patients with multiple metabolic risk factors: Large waist or Body mass index (BMI) ≥30, Dyslipidemia (raised TGs >150 or receiving treatment for elevated lipids), Dyslipidemia (reduced HDL cholesterol); hypertension (BP >140/90 on two occasions or receiving BP lowering medications), Type 2 diabetes or evidence of insulin resistance derived by HOMA-IR.
Overall total representation in MAESTRO-NASH trial	This trial included participants of the expected age range and ≥3 metabolic comorbid conditions. The study was conducted globally, with the majority of patients from USA. Approximately 55% of participants were women. Representative of the USA population, approximately 20% of participants were Hispanic or Latino.

Table S4. Demographic and Baseline Characteristics by Baseline Fibrosis Stage (Intent-to-Treat Population (n=1050)*

	F1 (n = 84)	Primary (N = 966)	F1B (n = 49)	F2 (n = 319)	F3 (n = 598)
Age, years	57.2 ± 13.3	56.6 ± 10.9	55.4 ± 12.8	55.0 ± 11.5	57.6 ± 10.3
Sex, male, no. (%)†	39 (46.4)	424 (43.9)	25 (51.0)	134 (42.0)	265 (44.3)
Body mass index, kg/m²	35.5 ± 6.6	35.7 ± 6.8	36.2 ± 7.6	36.0 ± 6.7	35.4 ± 6.7
Type 2 diabetes, no. (%)	45 (53.6)	647 (67.0)	33 (67.3)	189 (59.2)	425 (71.1)
Hypertension, no. (%)	62 (73.8)	754 (78.1)	39 (79.6)	225 (70.5)	490 (81.9)
Dyslipidemia, no. (%)	57 (67.9)	689 (71.3)	37 (75.5)	212 (66.5)	440 (73.6)
Hypothyroidism, no. (%)‡	12 (14.3)	130 (13.5)	6 (12.2)	48 (15.0)	76 (12.7)
History of ASCVD, no. (%)	3 (3.6)	57 (5.9)	6 (12.2)	15 (4.7)	36 (6.0)
FibroScan VCTE/LSM, kPa	10.4 ± 3.5	13.3 ± 6.5	11.3 ± 4.8	11.6 ± 5.6	14.4 ± 6.8
MRE, kPa	2.7 ± 0.44	3.6 ± 1.0	3.0 ± 0.58	3.1 ± 0.71	3.9 ± 1.0
FIB-4	1.2 ± 0.59	1.4 ± 0.70	1.2 ± 0.54	1.3 ± 0.64	1.5 ± 0.72
Enhanced liver fibrosis score	9.6 ± 0.93	9.8 ± 0.87	9.4 ± 0.78	9.5 ± 0.76	9.9 ± 0.88
Statin, no. (%)	33 (39.3)	473 (49.0)	32 (65.3)	128 (40.1)	313 (52.3)

*Plus-minus signs are mean ±SD.

†Sex was self-reported by the patient.

‡Patients on thyroxine replacement therapy at screening.

Note: "Primary" column denotes the primary analysis population and includes F1B (by consensus); F2 and F3 populations. The F1 population was scored as baseline F1A or F1C by both pathologists or F1A/F1B or F1C/F1B. The F1 population was an exploratory population and excluded from the primary analysis population because their NASH is not felt to be at risk for progression to advanced fibrosis.

ASCVD, atherosclerotic cardiovascular disease; FIB-4, fibrosis-4 index; LSM, liver stiffness measurement; MRE, magnetic resonance elastography; SD, standard deviation; VCTE, vibration-controlled transient elastography.

Table S5A. Sensitivity Analysis of Primary Endpoints: Consensus and Multiple Imputation

	% Response Resmetirom 80 mg (n = 316)	% Response Resmetirom 100 mg (n = 321)	% Response Placebo (n = 318)	% Difference Resmetirom 80 mg from PBO (95% CI)	% Difference Resmetirom 100 mg from PBO (95% CI)
NASH resolution					
Consensus (sensitivity)	24.4	27.7	7.9	16.8(11.3, 22.4)	20.7(15.0, 26.3)
Multiple imputation (sensitivity)	31.0	36.0	13.5	17.7 (11.3, 24.1)	23.2 (16.8, 29.6)
Fibrosis improvement					
Consensus (sensitivity)	24.4	25.5	12.3	12.2(6.3, 18.2)	13.4(7.4, 19.3)
Multiple imputation (sensitivity)	29.2	31.9	17.9	11.4 (5.1, 17.8)	14.1 (7.9, 20.4)

Data for patients with missing response data at Week 52 were imputed under the missing at random assumption by running a simulation (100 times) which produces a correlated pair of binary 0/1 data for each patient that represents the patient's response status for each pathologist. The statistics were then calculated for each imputation using the same approach as for the primary endpoint. The normalized results from each dataset were combined using Rubin's rule. Patients that were F3 at eligibility and re-evaluated as F4 at baseline by either pathologist are included in this analysis.

CI, confidence interval; NASH, nonalcoholic steatohepatitis; PBO, placebo. Unless otherwise stated "NASH resolution" means ballooning 0,1 with at least a 2-pt reduction in NAS and no worsening of fibrosis; "Fibrosis improvement" means at least 1-stage reduction in fibrosis with no worsening of NAS. Confidence interval widths have not been adjusted for multiplicity and may not be used for hypothesis testing.

Table S5B. Sensitivity Analysis of Primary Endpoint (NASH Resolution): Generalized Estimating Equation Model

	Resmetirom 80 mg (n = 316)	**Resmetirom 100 mg (n = 321)**	**Placebo (n = 318)**
NASH resolution at Week 52, no. (%)*			
(1,1)	62 (19.6)	62 (19.3)	18 (5.7)
(1,0)	24 (7.6)	55 (17.1)	22 (6.9)
(0,1)	16 (5.1)	13 (4.0)	4 (1.3)
(0,0)	214 (67.7)	191 (59.5)	274 (86.2)
OR of resmetirom group to placebo†	3.5	4.2	
95% CI of the OR	(2.3, 5.2)	(2.8, 6.2)	
Difference in percentage of responders, CMH, resmetirom group – placebo (SD)†	17.5 (3.0)	21.3 (3.0)	
95% CI of the difference	(11.7, 23.3)	(15.4, 27.2)	

*The ordered pair indicates (result according to pathologist A, result according to pathologist B). 1 indicates yes/responder; 0 indicates no/non-responder.
†The OR and difference in percentage of responders, 95% CIs, and p-values were obtained by fitting a GEE model with treatment arm, baseline type 2 diabetes status, baseline fibrosis stage, and pathologist by treatment interaction as factors, logit link function and a compound symmetric covariance structure, under the assumption that the scores from the two pathologists are repeated measurements for the same patient with some within-patient correlation structure.
Patients with no valid biopsy within the Week 52 window are considered non-responders for the Week 52 analysis, as are patients who experience a composite clinical endpoint (e.g., liver transplant, death) prior to their Week 52 biopsy.
Patients that were F3 at eligibility and re-evaluated as F4 at baseline by either pathologist are included in this analysis.
CI, confidence interval; GEE, generalized estimating equation; NASH, nonalcoholic steatohepatitis; OR, odds ratio; SD, standard deviation. Unless otherwise stated "NASH resolution" means ballooning 0,1 with at least a 2-pt reduction in NAS and no worsening of fibrosis; "Fibrosis improvement" means at least 1-stage reduction in fibrosis with no worsening of NAS. Confidence interval widths have not been adjusted for multiplicity and may not be used for hypothesis testing.

Table S5C. Sensitivity Analysis of Primary Endpoint (Fibrosis Improvement): Generalized Estimating Equation Model

	Resmetirom 80 mg (n = 316)	Resmetirom 100 mg (n = 321)	Placebo (n = 318)
Fibrosis improvement at Week 52, no. (%)*			
(1,1)	52 (16.5)	51 (15.9)	28 (8.8)
(1,0)	21 (6.6)	34 (10.6)	21 (6.6)
(0,1)	28 (8.9)	30 (9.3)	13 (4.1)
(0,0)	215 (68.0)	206 (64.2)	256 (80.5)
OR of resmetirom group to placebo†	2.0	2.1	
95% CI of the OR	(1.4, 2.9)	(1.5, 3.0)	
Difference in percentage of responders, CMH, resmetirom group - placebo (SD)†	9.7 (2.7)	11.0 (2.7)	
95% CI of the difference	(4.4, 15.0)	(5.7, 16.3)	

*The ordered pair indicates (result according to pathologist A, result according to pathologist B). 1 indicates yes/responder; 0 indicates no/non-responder.
†The OR and difference in percentage of responders, 95% CIs, and p-values were obtained by fitting a GEE model with treatment arm, baseline type 2 diabetes status, baseline fibrosis stage, and pathologist by treatment interaction as factors, logit link function and a compound symmetric covariance structure, under the assumption that the scores from the two pathologists are repeated measurements for the same patient with some within-patient correlation structure.
Patients with no biopsy within the Week 52 window or experience a composite clinical endpoint (e.g., liver transplant, death) prior to their Week 52 biopsy are considered non-responders for the Week 52 analysis.
Patients that were F3 at eligibility and re-evaluated as F4 at baseline by either pathologist are included in this analysis.
CI, confidence interval; GEE, generalized estimating equation; OR, odds ratio; SD, standard deviation. Unless otherwise stated "NASH resolution" means ballooning 0,1 with at least a 2-pt reduction in NAS and no worsening of fibrosis; "Fibrosis improvement" means at least 1-stage reduction in fibrosis with no worsening of NAS. Confidence interval widths have not been adjusted for multiplicity and may not be used for hypothesis testing.

Table S6. Sensitivity Analysis of Primary Endpoints: Tipping Point

	% Response Resmetirom 80 mg (n = 316)	% Response Resmetirom 100 mg (n = 321)	% Response Placebo (n = 318)	% Difference Resmetirom 80 mg from PBO (95% CI)†	p-value	% Difference Resmetirom 100 mg from PBO (95% CI)†	p-value
NASH resolution							
Placebo – 0% imputed responders (Primary Analysis)	25.9	29.9	9.7	16.4 (11.0, 21.8)*	<0.001	20.7 (15.3, 26.2)*	<0.001
Placebo – 30% imputed responders	25.9	29.9	13.7	12.4 (6.5, 18.3)	<0.001	16.7 (10.9, 22.6)	<0.001
Placebo – 60% imputed responders	25.9	29.9	17.8	8.3 (2.2, 14.4)	0.0154	12.7 (6.6, 18.8)	<0.001
Placebo – 67% imputed responders	25.9	29.9	18.8	7.4 (1.2, 13.5)	0.0377	11.7 (5.6, 17.9)	<0.001
Placebo – 90% imputed responders	25.9	29.9	21.9	4.2 (-2.0, 10.4)	0.3693	8.6 (2.4, 14.8)	0.0130
Placebo – 100% imputed responders	25.9	29.9	23.4	2.6 (-3.5, 8.8)	0.4047	7.0 (0.9, 13.2)	0.0246
Fibrosis improvement							
Placebo – 0% imputed responders (Primary Analysis)	24.2	25.9	14.2	10.2 (4.8, 15.7)*	<0.001	11.8 (6.4, 17.2)*	<0.001
Placebo – 10% imputed responders	24.2	25.9	15.6	8.8 (3.2, 14.4)	0.0048	10.4 (4.8, 15.9)	<0.001
Placebo – 20% imputed responders	24.2	25.9	16.8	7.5 (1.8, 13.3)	0.0201	9.1 (3.4, 14.8)	0.0038
Placebo – 25% imputed responders	24.2	25.9	17.5	6.9 (1.1, 12.7)	0.0398	8.5 (2.7, 14.2)	0.0084
Placebo – 30% imputed responders	24.2	25.9	18.1	6.2 (0.3, 12.1)	0.0762	7.8 (1.9, 13.6)	0.0181
Placebo – 35% imputed responders	24.2	25.9	18.9	5.5 (-0.5, 11.5)	0.1410	7.1 (1.2, 13.0)	0.0378

*Calculated using a stratified CMH approach. Patients with missing response at Week 52 are considered non-responders in both active and placebo groups.
†Active patients with missing response at Week 52 are considered non-responders. Placebo patients with missing response at Week 52 had their response imputed by running a simulation (100 times) which produces a correlated pair of binary 0/1 data that represents the patient's response for each pathologist. The

statistics were then calculated for each imputation using the same approach as in preceding footnote. The normalized results from each dataset were combined using Rubin's rule.
Note: Patients who experience a composite clinical endpoint (e.g., liver transplant, death) prior to their Week 52 biopsy. Patients that were F3 at eligibility and re-evaluated as F4 at baseline by either pathologist are included in this analysis. To account for multiplicity, 0.037 is considered to be the threshold for significance in this table.
CI, confidence interval; NASH, nonalcoholic steatohepatitis; PBO, placebo. Unless otherwise stated "NASH resolution" means ballooning 0,1 with at least a 2-pt reduction in NAS and no worsening of fibrosis; "Fibrosis improvement" means at least 1-stage reduction in fibrosis with no worsening of NAS. Confidence interval widths have not been adjusted for multiplicity and may not be used for hypothesis testing.

Table S7. Baseline Characteristics, F1 Patients

	Resmetirom 80mg (N=30)	Resmetirom 100mg (N=26)	Placebo (N=28)
Age, years	55.4 ± 13.0	57.8 ± 13.0	58.5 ± 14.3
Sex, male, no. (%)	14 (46.7%)	11 (42.3%)	14 (50%)
White	28 (93.3%)	23 (88.5%)	27 (96.4%)
Ethnicity, Hispanic or Latino, no. (%)	8 (26.7%)	8 (30.8%)	2(7.1%)
Body weight, kg	165.5 ± 9.2	97.6 ± 17.0	99.8 ± 25.2
Body mass index, kg/m^2	36.2 ± 6.9	35.5 ± 6.6	34.8 ± 6.4
Type 2 diabetes, no. (%)	13 (43.3%)	16 (61.5%)	16 (57.1%)
Hypertension, no. (%)	23 (76.6%)	18 (69.2%)	21 (75.0%)
Dyslipidemia, no. (%)	22 (73.3%)	15 (57.7%)	20 (71.4%)
Hypothyroidism, no. (%)[§]	3 (10%)	3 (11.5%)	6 (21.4%)
History of ASCVD, no. (%)	0	1 (3.8%)	2 (7.1%)
10-year ASCVD risk score	12.7 ± 15.9	14.7 ± 12.9	17.1 ± 13.0
FibroScan VCTE/LSM, kPa	9.8 ± 3.0	11.0 ± 7.1	12.9 ± 5.5
Median (Q1, Q3)	9.8 (8.8,11.7)	10.0 (9.1, 11.1)	9.7 (9.1, 10.9)
FibroScan CAP, dB/m	364.4 ± 34.3	355.2 ± 40.4	340.2 ± 41.0
MRI-PDFF, % fat fraction	18.6 ± 7.2	24.2 ± 8.9	20.2 ± 6.8
MRE, kPa	2.7 ± 0.42	2.9 ± 0.30	2.6 ± 0.52
FIB-4	1.2 ± 0.49	1.3 ± 0.74	1.2 ± 0.53
LDL-C, mg/dL	113.0 ± 48.1	113.3 ± 33.9	124.8 ± 48.8
Alanine aminotransferase, U/L	47.3 ± 28.4	58.1 ± 50.4	62.5 ± 48.2
Aspartate aminotransferase, U/L	30.4 ± 13.6	37.3 ± 20.2	38.0 ± 24.6
Gamma-glutamyl transferase, U/L	67.7 ± 54.8	68.3 ± 55.3	66.6 ± 65.7
Screening NAS ≥5, no. (%)	18 (60%)	17 (65.4%)	19(67.9%)

Table S8. Endpoints and Safety Data, F1 Patients

	Resmetirom 80 mg (N=30) n (%)	Resmetirom 100 mg (N=26) n (%)	Placebo (N=28) n (%)
Biopsy endpoints			
Nash resolution endpoint	38.3	25.0	7.4
relative to placebo	31.9 (11.2, 52.5)	17.6 (-0.9,36.2)	
Fibrosis improvement endpoint	21.7	15.4	11.1
	9.9 (-7.3, 27.0)	(-12.2, 21.1)	
SAFETY			
Any TEAE	28 (93.3)	21 (80.8)	26 (92.9)
Any Serious TEAE	6 (20.0)	6 (23.1)	4 (14.3)
Any TEAEs leading to Study Discontinuation	4 (13.3)	3 (11.5)	1 (3.6)
Grade 1	6 (20.0)	5 (19.2)	6 (21.4)
Grade 2	16 (53.3)	12 (46.2)	17 (60.7)
Grade 3	5 (16.7)	4 (15.4)	3 (10.7)
AEs >5%			
Diarrhea	8 (26.7)	5 (19.2)	7 (25.0)
Constipation	4 (13.3)	2 (7.7)	3 (10.7)
Nausea	5 (16.7)	2 (7.7)	2 (7.1)
Abdominal pain upper	2 (6.7)	1 (3.8)	4 (14.3)
Vomiting	1 (3.3)	3 (11.5)	2 (7.1)
Abdominal pain	1 (3.3)	2 (7.7)	2 (7.1)
Arthralgia	6 (20.0)	2 (7.7)	6 (21.4)
Back pain	4 (13.3)	3 (11.5)	4 (14.3)
Osteopenia	0	2 (7.7)	3 (10.7)
COVID-19	6 (20.0)	6 (23.1)	3 (10.7)
Urinary tract infection	2 (6.7)	4 (15.4)	1 (3.6)
Pruritus	2 (6.7)	4 (15.4)	1 (3.6)
Fatigue	3 (10.0)	2 (7.7)	2 (7.1)
Pyrexia	2 (6.7)	0	3 (10.7)
Headache	4 (13.3)	0	2 (7.1)
Dizziness	2 (6.7)	2 (7.7)	1 (3.6)
Rash	3 (10.0)	2 (7.7)	2 (7.1)

Type 2 diabetes mellitus	1 (3.3)	1 (3.8)	5 (17.9)
Weight decreased	1 (3.3)	3 (11.5)	1 (3.6)
Cough	0	1 (3.8)	4 (14.3)
Hypertension	1 (3.3)	0	4 (14.3)

Table S9. Histologic Response in Patients with Eligible Biopsies at Baseline and Week 52

	% Response Resmetirom 80 mg (N = 316)	% Response Resmetirom 100 mg (N = 321)	% Response Placebo (N = 318)	% Difference Resmetirom 80 mg from PBO (95% CI)	% Difference Resmetirom 100 mg from PBO (95% CI)
	(N = 258)	(N = 248)	(N = 276)		
NASH resolution* (in window including a baseline and Week 52 biopsy)					
% Response	31.8	38.7	11.2	20.9 (14.4, 27.2)	28.6 (22.2, 35.0)
Fibrosis improvement* (in window including a baseline and Week 52 biopsy)					
% Response	29.7	33.5	16.3	13.6 (7.3, 19.9)	17.2 (10.9, 23.6)
NASH resolution* OR fibrosis improvement* (in-window including a baseline and Week 52 biopsy)					
% Response	42.2	50.4	19.2	23.1 (15.4, 30.7)	31.2 (23.4, 39.0)

* NASH resolution or Fibrosis Improvement used the same definition as the primary endpoints (NASH resolution with ballooning 0, inflammation 0,1, with at least a 2-point improvement in NAS and no worsening of fibrosis stage; At least 1 stage fibrosis improvement with no worsening of NAS). Confidence interval widths have not been adjusted for multiplicity and may not be used for hypothesis testing.

Table S10. Additional Subgroups, Primary Endpoints

Additional Prespecified Subgroups				
	Relative to Placebo			
Assessment (%) (CI)	**Resmetirom 80 mg**		**Resmetirom 100 mg**	
Nash Resolution	**N**	**Result**	**N**	**Result**
≥5% Weight Gain from Baseline at Week 52	18	10.5 (-14.9, 36.0)	19	14.5 (-8.5, 37.4)
<5% Weight Gain from Baseline at Week 52	258	19.5 (13.4, 25.7)	248	26.9 (20.5, 33.2)
≥30% PDFF Reduction at Week 16	135	20.0 (12.6, 27.4)	157	26.2 (18.9, 33.5)
<30% PDFF Reduction at Week 16	90	13.8 (6.2, 21.3)	60	17.3 (8.6, 26.1)
Region: US	201	15.8 (9.0, 22.6)	224	19.9 (13.2, 26.7)
Region: Non-US	115	18.7 (9.4, 28.0)	97	22.9 (13.7, 32.1)
F2/F3	300	16.7 (11.1, 22.3)	306	20.9 (15.3, 26.5)
F1B	16	11.3 (-11.3, 34.0)	15	16.7 (-6.9, 40.2)
BMI <35 kg/m^2	164	22.4 (14.2, 30.6)	162	21.0 (13.4, 28.7)
BMI≥35 kg/m^2	152	9.8 (3.1, 16.6)	159	20.0 (12.3, 27.7)
BW≤200 pounds	120	17.4 (7.8, 26.9)	108	20.6 (11.0, 30.1)
BW>200 pounds	196	16.2 (9.6, 22.7)	212	21.1 (14.5, 27.8)
Fibrosis Improvement				
≥5% Weight Gain from Baseline at Week 52	18	0.9 (-22.6, 24.4)	19	16.0 (-7.9, 39.9)
<5% Weight Gain from Baseline at Week 52	258	12.8 (6.5, 19.0)	248	15.0 (8.8, 21.3)
≥30% PDFF Reduction at Week 16	135	13.5 (5.5, 21.5)	157	15.5 (8.0, 23.1)
<30% PDFF Reduction at Week 16	90	3.3 (-4.7, 11.4)	60	2.7 (-6.8, 12.2)
Region: US	201	7.8 (0.9, 14.6)	224	10.4 (3.6, 17.1)
Region: Non-US	115	14.6 (5.3, 23.9)	97	14.4 (5.5, 23.4)
F2/F3	300	10.2 (4.6, 15.9)	306	11.4 (5.8, 16.9)
F1B	16	10.5 (-11.4, 32.4)	15	19.6 (-2.6, 41.9)
BMI <35 kg/m^2	164	13.0 (5.2, 20.8)	162	11.8 (4.2, 19.4)
BMI≥35 kg/m^2	152	7.0 (-0.7, 14.7)	159	11.1 (3.4, 18.9)

BW≤200 pounds	120	15.6 (6.1, 25.1)	108	13.1 (3.9, 22.3)
BW>200 pounds	196	7.0 (0.4, 13.7)	213	10.8 (4.1, 17.6)
Body Weight Impact on Biopsy Responses SHBG and MRI-PDFF by Dose (post hoc)				
	Resmetirom 80 mg		**Resmetirom 100 mg**	
Assessment (%) (CI)	**≤100 kg**	**>100 kg**	**≤100 kg**	**>100 kg**
N	178	174	167	156
≥120% Increase in SHBG at Week 52	60.9 (52.9, 68.5)	29.6 (21.4, 38.8)	64.2 (55.4, 72.3)	55.0 (46.0, 63.8)
≥30% Reduction in MRI-PDFF at Week 52	66.9 (58.3, 74.7)	56.7 (46.3, 66.7)	71.7 (62.4, 79.8)	72.5 (63.1, 80.6)
Week 52 Primary Population *				
Consensus Fibrosis Improvement1	29.2 (22.7, 36.5)	17.4 (11.6, 24.6)	25.1 (18.8, 32.4)	25.6 (19.0, 33.2)
Consensus NASH Resolution1	26.4 (20.1, 33.5)	20.8 (14.5, 28.4)	26.9 (20.4, 34.4)	28.2 (21.3, 36.0)
Week 52 Paired Biopsies†				
N	147	111	125	123
Consensus Fibrosis Improvement	35.4 (27.7, 43.7)	22.5 (15.1, 31.4)	33.6 (25.4, 42.6)	32.5 (24.4, 41.6)
Consensus NASH Resolution	32.0 (24.5, 40.2)	27.0 (19.0, 36.3)	36.0 (27.6, 45.1)	35.8 (27.3, 44.9)

CI = 95% Clopper-Pearson confidence interval

* For NASH Resolution and Fibrosis Responder status, missing responses and Week 52 biopsies out of window are considered non-responders.

† Patients with a baseline and valid Week 52 biopsy. PDFF reduction in resmetirom groups are compared to all placebo patients with any Week 16 PDFF.

[1] Unless otherwise stated "NASH resolution" means ballooning 0,1 with at least a 2-pt reduction in NAS and no worsening of fibrosis; "Fibrosis improvement" means at least 1-stage reduction in fibrosis with no worsening of NAS. Confidence interval widths have not been adjusted for multiplicity and may not be used for hypothesis testing.

Table S11. Change From Baseline in Lipids, Lipoproteins, and Lipid Particles at Weeks 24 and 52 (Primary Analysis Population)

	LS Mean %CFB (SE) Resmetirom 80 mg (n = 322)	LS Mean %CFB (SE) Resmetirom 100 mg (n = 323)	LS Mean %CFB (SE) Placebo (n = 321)	LS Mean %CFB Difference Resmetirom 80 mg from PBO (95% CI)	LS Mean %CFB Difference Resmetirom 100 mg from PBO (95% CI)
LDL-C, mg/dL (baseline LDL-C >100 mg/dL)					
Week 24 – no.	148	133	150		
Baseline mean (SD)	135.6 (26.5)	134.4 (27.3)	136.8 (34.0)		
Week 24 (%CFB)	-21.3 (2.0)	-20.6 (2.0)	-5.9 (1.9)	-15.4 (-19.3, -11.6)	-14.7 (-18.6, -10.8)
Week 52 – no.	147	125	144		
Baseline mean (SD)	135.2 (26.0)	133.3 (27.4)	136.9 (34.2)		
Week 52 (%CFB)	-25.3 (2.2)	-27.1 (2.3)	-9.6 (2.1)	-15.7 (-20.0, -11.4)	-17.5 (-22.0, -13.1)
ApoB, mg/dL (baseline LDL-C >100 mg/dL)					
Week 24 – no.	148	133	150		
Baseline mean (SD)	117.9 (22.2)	116.9 (24.0)	118.3 (29.8)		
Week 24 (%CFB)	-21.9 (1.7)	-22.1 (1.8)	-3.7 (1.6)	-18.1 (-21.4, -14.9)	-18.4 (-21.8, -15.0)
Week 52 – no.	147	125	144		
Baseline mean (SD)	117.4 (20.9)	115.7 (24.1)	118.1 (29.9)		
Week 52 (%CFB)	-25.0 (2.0)	-26.6 (2.0)	-6.5 (1.9)	-18.5 (-22.4, -14.6)	-20.1 (-24.1, -16.1)
ApoCIII					
Week 24 – no.	282	272	288		
Baseline mean (SD)	10.9 (4.7)	10.7 (5.3)	10.5 (5.6)		
Week 24 (%CFB)	-10.6 (3.6)	-14.1 (3.1)	8.1 (3.1)	-18.7 (-27.1, -10.4)	-22.2 (-29.0, -15.4)
Week 52 – no.	272	255	279		
Baseline mean (SD)	10.9 (4.7)	10.7 (5.5)	10.3 (5.5)		
Week 52 (%CFB)	-10.0 (3.8)	-17.1 (3.3)	9.8 (3.3)	-19.8 (-28.4, -11.1)	-26.9 (-34.1, -19.6)
Non-HDL-C					
Week 24 – no.	285	280	294		
Baseline mean (SD)	135.6 (43.2)	131.8 (44.7)	135.7 (50.6)		
Week 24 (%CFB)	-15.2 (1.5)	-17.7 (1.6)	0.16 (1.5)	-15.4 (-18.8, -12.0)	-17.9 (-21.2, -14.5)

	LS Mean %CFB (SE) Resmetirom 80 mg (n = 322)	LS Mean %CFB (SE) Resmetirom 100 mg (n = 323)	LS Mean %CFB (SE) Placebo (n = 321)	LS Mean %CFB Difference Resmetirom 80 mg from PBO (95% CI)	LS Mean %CFB Difference Resmetirom 100 mg from PBO (95% CI)
Week 52 – no.	276	262	284		
Baseline mean (SD)	136.1 (43.1)	132.4 (46.7)	135.2 (50.6)		
Week 52 (%CFB)	-15.7 (1.7)	-22.1 (1.7)	-0.40 (1.6)	-15.3 (-19.0, -11.7)	-21.7 (-25.2, -18.1)
HDL-C					
Week 24 – no.	285	280	294		
Baseline mean (SD)	43.9 (12.5)	43.5 (12.8)	43.8 (13.5)		
Week 24 (%CFB)	2.7 (1.5)	2.9 (1.5)	1.7 (1.5)	0.98 (-2.3, 4.2)	1.2 (-2.1, 4.5)
Week 52 – no.	276	262	284		
Baseline mean (SD)	44.1 (12.5)	43.6 (12.7)	44.1 (13.3)		
Week 52 (%CFB)	4.7 (1.4)	4.6 (1.5)	2.7 (1.4)	2.1 (-0.99, 5.1)	2.0 (-1.2, 5.1)
HDL particles, umol/L					
Week 52 – no.	266	252	275		
Baseline mean (SD)	30.5 (6.5)	31.0 (6.0)	30.8 (6.2)		
Week 52 (%CFB)	5.2 (1.3)	2.3 (1.3)	1.6 (1.2)	3.6 (0.80, 6.3)	0.69 (-2.1, 3.5)
HDL particle size, nm					
Week 52 – no.	266	252	275		
Baseline mean (SD)	8.9 (0.46)	9.0 (0.46)	9.0 (0.45)		
Week 52 (%CFB)	0.90 (0.28)	0.40 (0.28)	0.30 (0.27)	0.60 (-0.01, 1.2)	0.10 (-0.51, 0.71)
LDL particles, nmol/L					
Week 52 – no.	266	252	275		
Baseline mean (SD)	1322.9 (438.5)	1273.5 (440.1)	1282.5 (476.1)		
Week 52 (%CFB)	-16.8 (1.7)	-20.0 (1.7)	-0.69 (1.6)	-16.1 (-19.8, -12.4)	-19.4 (-23.1, -15.6)
LDL particle size, nm					
Week 52 – no.	260	251	273		
Baseline mean (SD)	20.4 (0.59)	20.3 (0.59)	20.4 (0.60)		
Week 52 (%CFB)	-0.32 (0.17)	-0.35 (0.17)	0.04 (0.16)	-0.36 (-0.73, 0.01)	-0.39 (-0.76, -0.02)
Large HDL particles, umol/L					
Week 52 – no.	265	251	275		
Baseline mean (SD)	5.5 (2.9)	5.5 (2.9)	5.8 (2.9)		
Week 52 (%CFB)	12.3 (5.5)	4.5 (5.6)	12.2 (5.3)	0.06 (-11.9, 12.0)	-7.7 (-19.8, 4.4)

	LS Mean %CFB (SE) Resmetirom 80 mg (n = 322)	LS Mean %CFB (SE) Resmetirom 100 mg (n = 323)	LS Mean %CFB (SE) Placebo (n = 321)	LS Mean %CFB Difference Resmetirom 80 mg from PBO (95% CI)	LS Mean %CFB Difference Resmetirom 100 mg from PBO (95% CI)
Large LDL particles, nmol/L					
Week 52 – no.	234	213	239		
Baseline mean (SD)	366.8 (263.0)	330.2 (265.6)	355.7 (257.3)		
Week 52 (%CFB)	47.8 (25.3)	55.6 (26.0)	86.4 (24.3)	-38.6 (-93.6, 16.3)	-30.8 (-87.0, 25.4)
Large VLDL and chylomicron particles, nmol/L					
Week 52 – no.	266	252	275		
Baseline mean (SD)	8.2 (5.8)	8.2 (8.9)	8.3 (8.7)		
Week 52 (%CFB)	-5.7 (4.9)	-9.1 (5.0)	11.2 (4.7)	-16.8 (-27.5, -6.2)	-20.2 (-31.1, -9.4)
Medium HDL particles, umol/L					
Week 52 – no.	260	245	271		
Baseline mean (SD)	5.5 (4.1)	5.7 (4.6)	5.1 (4.1)		
Week 52 (%CFB)	88.2 (28.0)	130.2 (28.7)	65.1 (27.1)	23.1 (-38.2, 84.4)	65.2 (2.9, 127.5)
Medium VLDL particles, nmol/L					
Week 52 – no.	253	240	260		
Baseline mean (SD)	25.0 (25.2)	23.2 (25.6)	23.1 (27.6)		
Week 52 (%CFB)	-18.7 (52.0)	-27.4 (53.5)	115.8 (50.6)	-134.5 (-249.6, -19.5)	-143.2 (-259.8, -26.7)
Small LDL particles, nmol/L					
Week 52 – no.	266	252	275		
Baseline mean (SD)	880.7 (336.3)	882.0 (360.3)	850.0 (356.9)		
Week 52 (%CFB)	-3.3 (9.2)	6.6 (9.5)	21.7 (8.9)	-25.0 (-45.2, -4.7)	-15.0 (-35.5, 5.5)
Small VLDL particles, nmol/L					
Week 52 – no.	260	250	270		
Baseline mean (SD)	25.8 (15.0)	27.1 (17.0)	27.4 (15.3)		
Week 52 (%CFB)	62.9 (38.2)	26.3 (39.0)	41.6 (36.9)	21.3 (-62.8, 105.4)	-15.3 (-100.2, 69.5)
VLDL and chylomicron particles, nmol/L					
Week 52 – no.	266	252	275		
Baseline mean (SD)	57.2 (32.4)	57.1 (38.8)	57.1 (38.2)		
Week 52 (%CFB)	-11.2 (4.3)	-20.2 (4.4)	8.2 (4.1)	-19.4 (-28.8, -10.0)	-28.4 (-37.9, -18.9)

	LS Mean %CFB (SE) Resmetirom 80 mg (n = 322)	LS Mean %CFB (SE) Resmetirom 100 mg (n = 323)	LS Mean %CFB (SE) Placebo (n = 321)	LS Mean %CFB Difference Resmetirom 80 mg from PBO (95% CI)	LS Mean %CFB Difference Resmetirom 100 mg from PBO (95% CI)
VLDL particle size, nm					
Week 52 – no.	266	252	274		
Baseline mean (SD)	55.5 (6.5)	55.3 (6.9)	54.7 (7.3)		
Week 52 (%CFB)	1.9 (0.92)	3.4 (0.94)	2.0 (0.89)	-0.05 (-2.1, 2.0)	1.4 (-0.65, 3.4)
VLDL and chylomicron triglycerides, mg/dL					
Week 52 – no.	266	252	275		
Baseline mean (SD)	114.1 (68.6)	113.5 (92.1)	113.6 (91.7)		
Week 52 (%CFB)	-14.7 (3.1)	-17.9 (3.1)	4.9 (3.0)	-19.6 (-26.3, -12.9)	-22.8 (-29.6, -16.0)

ApoB, apolipoprotein B; ApoCIII, apolipoprotein CIII; CFB, change from baseline; CI, confidence interval; HDL, high-density lipoprotein; LDL, low-density lipoprotein; LS, least squares; PBO, placebo; RLP, remnant-like protein; SD, standard deviation; SE, standard error; VLDL, very low-density lipoprotein. Confidence interval widths have not been adjusted for multiplicity and may not be used for hypothesis testing.

Table S12. Additional Secondary Endpoints (Primary Analysis population).

	LS Mean %CFB (SE) Resmetirom 80 mg (n = 321)	LS Mean %CFB (SE) Resmetirom 100 mg (n = 323)	LS Mean %CFB (SE) Placebo (n = 321)	LS Mean %CFB Difference Resmetirom 80 mg from PBO (95% CI)	LS Mean %CFB Difference Resmetirom 100 mg from PBO (95% CI)
LDL-C (mg/dL)					
Week 24* – no.	285	280	294		
Baseline mean (SD)	106.6 (37.8)	102.9 (37.6)	106.2 (41.4)		
Week 24 (%CFB)	-13.6 (1.7)	-16.3 (1.7)	0.11 (1.7)	-13.7 (-17.5, -10.0)	-16.4 (-20.1, -12.6)
p value				<0.001	<0.001
Week 52 – no.	276	262	284		
Baseline mean (SD)	106.9 (37.9)	102.9 (36.8)	106.1 (41.7)		
Week 52 (%CFB)	-13.7 (1.8)	-19.5 (1.8)	-0.4 (1.7)	-13.3 (-17.3, -9.3)	-19.0 (-23.0, -15.1)
ApoB (U/L)					
Week 24 – no.	285	280	294		
Baseline mean (SD)	98.4 (28.1)	95.9 (28.3)	97.5 (32.1)		
Week 24 (%CFB)	-16.8 (1.3)	-19.8 (1.3)	0.39 (1.3)	-17.2 (-20.0, -14.4)	-20.2 (-22.9, -17.4)
Week 52 – no.	276	262	284		
Baseline mean (SD)	98.5 (27.6)	95.6 (27.8)	97.1 (32.1)		
Week 52 (%CFB)	-16.2 (1.5)	-22.3 (1.5)	0.59 (1.4)	-16.8 (-20.0, -13.7)	-22.9 (-26.0, -19.7)
Triglycerides (mg/dL) (baseline triglycerides >150 mg/dL)					
Week 24 – no.	170	146	144		
Baseline mean (SD)	237.9 (120.7)	244.7 (132.8)	261.5 (146.0)		
Week 24 (%CFB)	-22.7 (4.0)	-21.7 (4.3)	-2.6 (4.1)	-20.1 (-28.3, -11.8)	-19.1 (-27.8, -10.3)
Week 52 – no.	165	134	140		
Baseline mean (SD)	241.0 (122.4)	252.1 (160.2)	256.5 (145.7)		
Week 52 (%CFB)	-22.5 (4.2)	-28.4 (4.4)	-3.5 (4.2)	-19.0 (-27.9, -10.1)	-24.9 (-34.1, -15.7)
Lp(a) (nmol/L) (baseline Lp(a) >10 nmol/L)					
Week 24 – no.	200	187	200		
Baseline mean (SD)	62.0 (67.5)	58.7 (64.6)	57.9 (70.2)		

	LS Mean %CFB (SE) Resmetirom 80 mg (n = 321)	LS Mean %CFB (SE) Resmetirom 100 mg (n = 323)	LS Mean %CFB (SE) Placebo (n = 321)	LS Mean %CFB Difference Resmetirom 80 mg from PBO (95% CI)	LS Mean %CFB Difference Resmetirom 100 mg from PBO (95% CI)
Week 24 (%CFB)	-30.4 (3.8)	-35.9 (4.0)	-0.84 (3.5)	-29.5 (-37.6, -21.5)	-35.1 (-43.5, -26.6)
Week 52 – no.	193	172	194		
Baseline mean (SD)	64.5 (68.4)	57.6 (62.7)	57.7 (70.0)		
Week 52 (%CFB)	-34.0 (4.9)	-37.5 (5.6)	-5.0 (4.6)	-29.5 (-39.4, -19.6)	-32.4 (-43.1, -21.8)
MRI-PDFF, % fat fraction					
Week 16 – no.	228	218	224		
Baseline mean (SD)	18.2 (6.8)	17.2 (6.5)	17.9 (6.6)		
Week 16 (%CFB)	-37.8 (2.2)	-42.1 (2.2)	-6.4 (2.2)	-31.4 (-36.3, -26.4)	-35.7 (-40.7, -30.7)
Week 52 – no.	233	222	230		
Baseline mean (SD)	18.2 (6.8)	17.2 (6.7)	17.9 (6.6)		
Week 52 (%CFB)	-35.4 (2.8)	-46.6 (2.8)	-8.7 (2.7)	-26.7 (-32.9, -20.6)	-37.9 (-44.2, -31.7)
ALT (U/L)[#]					
Baseline – no.	265	264	244		
Baseline mean (SD)	59.1 (25.4)	65.4 (33.5)	63.1 (26.7)		
Week 48 (CFB)	-20.4 (2.2)	-24.8 (2.3)	-8.5 (2.3)		
Week 48 (%CFB)	-26.6 (3.7)	-33.2 (3.9)	-6.9 (3.8)	-19.7 (-27.7, -11.6)	-26.3 (-34.5, -18.1)
AST (U/L)					
Baseline – no.	265	264	244		
Baseline mean (SD)	41.5 (18.6)	48.5 (26.0)	44.9 (20.8)		
Week 48 (CFB)	-13.9 (1.7)	-15.8 (1.7)	-6.0 (1.7)		
Week 48 (%CFB)	-22.1 (3.9)	-28.3 (3.9)	-2.9 (3.8)	-19.3 (-27.2, -11.3)	-25.4 (-33.5, -17.4)
GGT (U/L)					
Baseline – no.	265	264	244		
Baseline mean (SD)	87.3 (122.2)	92.4 (109.4)	80.1 (87.2)		
Week 48 (CFB)	-31.3 (3.9)	-32.7 (4.0)	-7.3 (3.9)		
Week 48 (%CFB)	-25.0 (5.5)	-31.9 (6.3)	3.3 (5.2)	-28.3 (-37.3, -19.3)	-35.2 (-45.5, -25.0)

	LS Mean %CFB or CFB (SE) Resmetirom 80 mg (n = 322)	LS Mean %CFB or CFB (SE) Resmetirom 100 mg (n = 323)	LS Mean %CFB or CFB (SE) Placebo (n = 321)	LS Mean Difference or Resmetirom 80 mg from PBO (95% CI)	LS Mean Difference or Resmetirom 100 mg from PBO (95% CI)
FibroScan CAP, dB/m					
Week 52 – no.	256	252	267		
Baseline mean (SD)	346.7 (37.4)	348.4 (40.3)	347.0 (36.9)		
Week 52 (CFB)	-39.6 (4.3)	-41.3 (4.4)	-14.5 (4.1)	-25.2 (-34.5, -15.9)	-26.9 (-36.2, -17.5)
FibroScan VCTE/LSM, kPa					
F1B – no.	12	10	17		
Baseline mean (SD)	11.1 (5.5)	12.5 (5.7)	10.9 (4.4)		
Week 52 (CFB)	-3.7 (1.0)	-3.7 (1.3)	-0.62 (0.87)	-3.1 (-5.8, -0.33)	-3.1 (-6.1, -0.02)
Responder Analysis – no.	12	10	17		
Improving ≥25% – no. (%)	6 (50.0)	7 (70.0)	4 (23.5)	3.7 (0.70, 19.1)	10.0 (1.3, 75.3)
Improving ≥30% – no. (%)	5 (41.7)	6 (60.0)	4 (23.5)	2.7 (0.50, 14.0)	4.5 (0.78, 25.5)
Worsening≥25% – no. (%)	1 (8.3)	0	5 (29.4)	0.17 (0.02, 2.0)	0
Worsening ≥30% – no. (%)	1 (8.3)	0	5 (29.4)	0.17 (0.02, 2.0)	0
F2 – no.	84	76	89		
Baseline mean (SD)	11.6 (5.5)	10.9 (3.3)	10.7 (3.3)		
Week 52 (CFB)	-2.3 (0.44)	-2.4 (0.46)	-1.3 (0.43)	-1.1 (-2.3, 0.12)	-1.2 (-2.4, 0.04)
Responder Analysis – no.	84	76	89		
Improving ≥25% – no. (%)	37 (44.0)	38 (50.0)	31 (34.8)	1.5 (0.80, 2.7)	1.9 (1.0, 3.6)
Improving ≥30% – no. (%)	28 (33.3)	31 (40.8)	23 (25.8)	1.4 (0.75, 2.8)	2.0 (1.0, 3.8)
Worsening≥25% – no. (%)	10 (11.9)	8 (10.5)	13 (14.6)	0.80 (0.33, 1.9)	0.68 (0.26, 1.8)
Worsening ≥30% – no. (%)	7 (8.3)	5 (6.6)	12 (13.5)	0.57 (0.21, 1.5)	0.47 (0.16, 1.4)
F3 – no.	163	167	162		
Baseline mean (SD)	14.2 (6.2)	14.7 (8.4)	13.8 (5.1)		
Week 52 (CFB)	-2.0 (0.41)	-3.2 (0.41)	-1.1 (0.41)	-0.86 (-2.0, 0.25)	-2.1 (-3.2, -1.0)
Responder Analysis – no.	163	167	162		
Improving ≥25% – n (%)	67 (41.1)	80 (47.9)	43 (26.5)	2.0 (1.2, 3.1)	2.5 (1.6, 4.0)
Improving ≥30% – n (%)	56 (34.4)	65 (38.9)	29 (17.9)	2.5 (1.5, 4.2)	2.9 (1.7, 4.7)
Worsening≥25% – n (%)	24 (14.7)	18 (10.8)	30 (18.5)	0.77 (0.43, 1.4)	0.53 (0.28, 1.0)
Worsening ≥30% – n (%)	21 (12.9)	17 (10.2)	21 (13.0)	1.0 (0.53, 1.9)	0.76 (0.39, 1.5)

	LS Mean %CFB or CFB (SE) Resmetirom 80 mg (n = 322)	LS Mean %CFB or CFB (SE) Resmetirom 100 mg (n = 323)	LS Mean %CFB or CFB (SE) Placebo (n = 321)	LS Mean Difference or Resmetirom 80 mg from PBO (95% CI)	LS Mean Difference or Resmetirom 100 mg from PBO (95% CI)
MRE, kPa					
F1B – no.	7	10	11		
Baseline mean (SD)	3.1 (0.53)	2.8 (0.44)	3.2 (0.71)		
Week 52 (%CFB)	-5.7 (7.2)	2.4 (6.2)	7.5 (5.5)	-13.2 (-31.3, 4.8)	-5.1 (-21.9, 11.8)
Responder Analysis – no.	5	5	6		
≥19% increase from BL – no. (%)	0	1 (20.0)	2 (33.3)	0	0.50 (0.03, 9.0)
≥19% reduction from BL – no. (%)	1 (20.0)	0	0	NA	--
F2 – no.	47	42	61		
Baseline mean (SD)	3.0 (0.69)	3.0 (0.71)	3.0 (0.61)		
Week 52 (%CFB)	-1.9 (3.1)	1.1 (3.2)	2.1 (2.7)	-4.0 (-12.0, 4.0)	-0.99 (-9.2, 7.3)
Responder Analysis – no.	20	22	33		
≥19% increase from BL – no. (%)	0	0	2 (6.1)	0	0
≥19% reduction from BL – no. (%)	7 (35.0)	5 (22.7)	8 (24.2)	1.6 (0.45, 5.6)	0.89 (0.25, 3.2)
F3 – no.	94	96	91		
Baseline mean (SD)	3.8 (0.92)	4.0 (1.1)	3.9 (1.1)		
Week 52 (%CFB)	-8.9 (1.9)	-5.2 (1.9)	-0.38 (1.9)	-8.6 (-13.7, -3.4)	-4.8 (-10.0, 0.30)
Responder Analysis – no.	84	86	80		
≥19% increase from BL – no. (%)	3 (3.6)	7 (8.1)	13 (16.3)	0.19 (0.05, 0.70)	0.46 (0.17, 1.2)
≥19% reduction from BL – no. (%)	22 (26.2)	27 (31.4)	10 (12.5)	2.5 (1.1, 5.5)	3.2 (1.4, 7.2)
Enhanced liver fibrosis score (baseline enhanced liver fibrosis score ≥9.8)					
no.	122	123	122		
Baseline mean (SD)	10.5 (0.59)	10.5 (0.53)	10.5 (0.59)		
Week 52 (CFB)	-0.34 (0.092)	-0.35 (0.094)	-0.11 (0.091)	-0.22 (-0.40, -0.05)	-0.24 (-0.41, -0.07)

	LS Mean %CFB or CFB (SE) Resmetirom 80 mg (n = 322)	LS Mean %CFB or CFB (SE) Resmetirom 100 mg (n = 323)	LS Mean %CFB or CFB (SE) Placebo (n = 321)	LS Mean Difference or Resmetirom 80 mg from PBO (95% CI)	LS Mean Difference or Resmetirom 100 mg from PBO (95% CI)
PIIINP, ng/mL (baseline PIIINP ≥9 ng/mL)					
no.	193	200	200		
Baseline mean (SD)	14.0 (4.4)	13.6 (4.4)	13.4 (3.9)		
Week 52 (%CFB)	-10.7 (2.8)	-10.3 (2.7)	-0.99 (2.6)	-9.7 (-15.5, -3.8)	-9.3 (-15.1, -3.5)
Week 52 (CFB)	-2.0 (0.38)	-2.1 (0.37)	-0.66 (0.36)	-1.4 (-2.2, -0.56)	-1.4 (-2.2, -0.61)
TIMP-1, ng/mL (baseline TIMP-1 ≥240 ng/mL)					
no.	174	172	183		
Baseline mean (SD)	304.0 (64.7)	309.3 (73.4)	303.7 (64.0)		
Week 52 (%CFB)	-9.3 (2.0)	-10.1 (2.0)	-3.6 (1.9)	-5.7 (-9.6, -1.8)	-6.5 (-10.4, -2.6)
Week 52 (CFB)	-31.0 (6.2)	-35.7 (6.2)	-13.3 (6.0)	-17.7 (-30.1, -5.3)	-22.3 (-34.8, -9.9)
Hyaluronic acid, ug/L (baseline hyaluronic acid ≥50 ng/mL)					
no.	156	152	155		
Baseline mean (SD)	114.9 (99.2)	122.9 (88.4)	125.3 (111.0)		
Week 52 (%CFB)	17.2 (7.8)	11.8 (8.1)	24.8 (7.8)	-7.6 (-23.1, 7.8)	-13.0 (-28.5, 2.5)
Week 52 (CFB)	1.1 (9.7)	-2.9 (10.1)	0.55 (9.7)	0.58 (-18.7, 19.9)	-3.5 (-22.9, 16.0)
CK-18, U/L					
no.	279	264	277		
Baseline mean (SD)	834.8 (476.3)	849.2 (520.3)	857.9 (519.0)		
Week 52 (CFB)	-278.2 (29.2)	-309.5 (29.8)	-143.9 (28.5)	-134.3 (-198.1, -70.5)	-165.7 (-230.3, -101.0)
Adiponectin, ug/mL					
no.	277	266	279		
Baseline mean (SD)	4.1 (2.6)	4.2 (2.5)	3.9 (2.3)		
Week 52 (CFB)	0.86 (0.19)	1.1 (0.19)	-0.10 (0.18)	0.97 (0.56, 1.4)	1.3 (0.83, 1.7)
rT3, ng/dL					
no.	278	269	282		
Baseline mean (SD)	18.5 (5.4)	19.2 (6.2)	18.4 (5.6)		
Week 52 (CFB)	-4.6 (0.31)	-5.1 (0.32)	0.17 (0.30)	-4.7 (-5.4, -4.1)	-5.2 (-5.9, -4.6)

	LS Mean %CFB or CFB (SE) Resmetirom 80 mg (n = 322)	LS Mean %CFB or CFB (SE) Resmetirom 100 mg (n = 323)	LS Mean %CFB or CFB (SE) Placebo (n = 321)	LS Mean Difference or Resmetirom 80 mg from PBO (95% CI)	LS Mean Difference or Resmetirom 100 mg from PBO (95% CI)
Liver volume					
Week 16 – no.	229	217	226		
Baseline mean (SD)	2447.0 (615.7)	2376.0 (642.8)	2404.5 (666.0)		
Week 16 (%CFB)	-18.8 (0.90)	-21.4 (0.91)	-0.29 (0.87)	-18.5 (-20.5, -16.5)	-21.1 (-23.1, -19.1)
Week 52 – no.	235	225	235		
Baseline mean (SD)	2410.1 (598.9)	2368.8 (645.5)	2406.6 (683.9)		
Week 52 (%CFB)	-21.6 (1.1)	-25.8 (1.1)	-1.0 (1.0)	-20.5 (-22.9, -18.2)	-24.8 (-27.2, -22.4)
Spleen volume					
Week 16 – no.	229	217	226		
Baseline mean (SD)	373.4 (170.6)	357.2 (164.9)	361.8 (193.5)		
Week 16 (%CFB)	-2.4 (0.95)	-3.6 (0.96)	1.5 (0.92)	-3.9 (-6.0, -1.8)	-5.0 (-7.2, -2.9)
Week 52 – no.	235	225	235		
Baseline mean (SD)	364.0 (168.3)	355.8 (163.8)	360.2 (189.8)		
Week 52 (%CFB)	-5.9 (1.1)	-6.1 (1.1)	3.2 (1.1)	-9.0 (-11.5, -6.5)	-9.3 (-11.8, -6.7)

*Key secondary endpoint. # multiply imputed values. Direct LDL-C was measured.

CI, confidence interval; NAS, nonalcoholic fatty liver disease activity score; NASH, nonalcoholic steatohepatitis; PBO, placebo. BL, baseline; CAP, controlled attenuation parameter; CFB, change from baseline; CI, confidence interval; CK-18, cytokeratin 18; FIB-4, fibrosis-4 index; LS, least squares; LSM, liver stiffness measurement; MRE, magnetic resonance elastography; MRI-PDFF, magnetic resonance imaging-proton density fat fraction; OR, odds ratio; rT3, reverse triiodothyronine; VCTE, vibration-controlled transient elastography; SD, standard deviation; SE, standard error. Confidence interval widths have not been adjusted for multiplicity and may not be used for hypothesis testing.

Table S13. Adverse Events Reported in ≥5% of Patients (Safety Population)

Preferred Term, no. (%)	Resmetirom 80 mg (n = 322)	Resmetirom 100 mg (n = 323)	Placebo (n = 321)
Abdominal pain upper	23 (7.1)	27 (8.4)	29 (9.0)
Headache	30 (9.3)	25 (7.7)	27 (8.4)
Vomiting	28 (8.7)	35 (10.8)	17 (5.3)
Type 2 diabetes	24 (7.5)	26 (8.0)	25 (7.8)
Abdominal pain	23 (7.1)	29 (9.0)	18 (5.6)
Constipation	21 (6.5)	28 (8.7)	17 (5.3)
Muscle spasms	14 (4.3)	22 (6.8)	22 (6.9)
Hypertension	16 (5.0)	13 (4.0)	25 (7.8)
Dizziness	20 (6.2)	19 (5.9)	11 (3.4)
Nasopharyngitis	14 (4.3)	20 (6.2)	14 (4.4)
Pain in extremity	12 (3.7)	12 (3.7)	23 (7.2)
Upper respiratory tract infection	23 (7.1)	8 (2.5)	17 (5.3)
Rash	12 (3.7)	21 (6.5)	12 (3.7)
Cough	14 (4.3)	18 (5.6)	12 (3.7)
Abdominal distension	14 (4.3)	13 (4.0)	17 (5.3)
Procedural pain	16 (5.0)	9 (2.8)	19 (5.9)
Gastrooesophageal reflux disease	16 (5.0)	7 (2.2)	8 (2.5)
Decreased appetite	5 (1.6)	16 (5.0)	4 (1.2)
Sinusitis	10 (3.1)	13 (4.0)	17 (5.3)

Table S14. Serious Adverse Events

System Organ Class*	Resmetirom 80 mg (N=322) n (%)	Resmetirom 100 mg (N=323) n (%)	Placebo (N=321) n (%)
Patients with at least one TE-SAEs	35 (10.9)	41 (12.7)	37 (11.5)
Infections and infestations	12 (3.7)	5 (1.5)	13 (4.0)
COVID-19	1 (0.3)	1 (0.3)	2 (0.6)
COVID-19 pneumonia	2 (0.6)	0	3 (0.9)
Gastrointestinal disorders	3 (0.9)	7 (2.2)	7 (2.2)
Acute Gallstone-related disorders[#]	3 (0.9)	3 (0.9)	1 (0.3)
Injury, poisoning and procedural complications	3 (0.9)	8 (2.5)	4 (1.2)
Post procedural haemorrhage	1 (0.3)	1 (0.3)	2 (0.6)
Cardiac disorders	3 (0.9)	3 (0.9)	7 (2.2)
Respiratory, thoracic and mediastinal disorders	3 (0.9)	4 (1.2)	6 (1.9)
Musculoskeletal and connective tissue disorders	7 (2.2)	1 (0.3)	4 (1.2)
Nervous system disorders	3 (0.9)	2 (0.6)	6 (1.9)
General disorders and administration site conditions	2 (0.6)	5 (1.5)	3 (0.9)

* Included SAEs where category was not unblinding [#] Includes a combination of acute cholecystitis, gallstone related pancreatitis, or choledolithiasis. Individual SAEs that were unblinding for treatment of individual patients are not shown because MAESTRO-NASH is a blinded ongoing 54-month outcome study.

Table S15. Malignancies

	Resmetirom 80 mg (N=322) n (%)	Resmetirom 100 mg (N=323) n (%)	Placebo (N=321) n (%)
Patients with any malignancy	4 (1.2)	11 (3.4)	12 (3.7)
Basal cell carcinoma	2 (0.6)	3 (0.9)	2 (0.6)
Breast carcinoma	1 (0.3)	2 (0.6)	1 (0.3)
Malignant melanoma	0	1 (0.3)	1 (0.3)
Skin cancer (squamous cell)	1 (0.3)	1 (0.3)	3 (0.9)

* Malignancies that were unblinding for treatment of individual patients are not shown because MAESTRO-NASH is a blinded ongoing 54-month outcome study.

Table S16. Change From Baseline in Metabolic Factors at Week 52 (Primary Analysis Population)

	LS Mean %CFB (SE) Resmetirom 80 mg (n = 322)	LS Mean %CFB (SE) Resmetirom 100 mg (n = 323)	LS Mean %CFB (SE) Placebo (n = 321)	LS Mean %CFB Difference Resmetirom 80 mg from PBO (95% CI)	LS Mean %CFB Difference Resmetirom 100 mg from PBO (95% CI)
Body weight, kg					
no.	281	265	286		
Baseline mean (SD)	98.6 (21.9)	102.1 (22.6)	99.5 (22.4)		
Week 52 (%CFB)	-1.2 (0.37)	-1.8 (0.38)	-0.87 (0.36)	-0.35 (-1.2, 0.45)	-0.88 (-1.7, -0.07)
SBP, mmHg					
no.	281	265	286		
Baseline mean (SD)	129.7 (13.6)	129.8 (15.1)	130.2 (14.4)		
Week 52 (%CFB)	-1.9 (0.70)	-2.1 (0.71)	0.74 (0.67)	-2.7 (-4.2, -1.2)	-2.9 (-4.4, -1.4)
DBP, mmHg					
no.	281	265	286		
Baseline mean (SD)	79.4 (8.8)	79.1 (9.7)	80.8 (9.5)		
Week 52 (%CFB)	-1.6 (0.71)	-2.0 (0.72)	-0.17 (0.69)	-1.5 (-3.0, 0.08)	-1.9 (-3.4, -0.32)
Heart rate, beats/min (based on electrocardiogram)					
no.	278	267	280		
Baseline mean (SD)	70.4 (10.7)	69.0 (10.4)	69.7 (10.9)		
Week 52 (%CFB)	-1.7 (0.83)	-2.6 (0.84)	-0.19 (0.81)	-1.5 (-3.3, 0.28)	-2.4 (-4.2, -0.57)
Glucose (mg/dL)					
no.	277	261	285		
Baseline mean (SD)	131.7 (40.8)	129.4 (35.0)	128.4 (38.8)		
Week 52 (%CFB)	-0.72 (1.8)	-3.4 (1.9)	2.2 (1.7)	-2.9 (-6.8, 1.0)	-5.6 (-9.5, -1.6)
Insulin (mIU/L)					
no.	277	261	285		
Baseline mean (SD)	34.5 (24.5)	31.5 (20.1)	33.2 (30.6)		
Week 52 (%CFB)	2.8 (5.0)	3.3 (5.1)	5.1 (4.8)	-2.3 (-13.0, 8.4)	-1.8 (-12.6, 9.1)
HOMA-IR					
no.	276	261	285		
Baseline mean (SD)	11.8 (11.1)	10.3 (8.0)	11.0 (12.7)		
Week 52 (%CFB)	9.1 (7.1)	3.4 (7.2)	14.4 (6.8)	-5.3 (-20.5, 9.9)	-11.0 (-26.4, 4.4)

	LS Mean %CFB (SE) Resmetirom 80 mg (n = 322)	LS Mean %CFB (SE) Resmetirom 100 mg (n = 323)	LS Mean %CFB (SE) Placebo (n = 321)	LS Mean %CFB Difference Resmetirom 80 mg from PBO (95% CI)	LS Mean %CFB Difference Resmetirom 100 mg from PBO (95% CI)
HbA1c (%)					
no.	277	262	285		
Baseline mean (SD)	6.6 (1.1)	6.6 (1.1)	6.5 (1.0)		
Week 52 (%CFB)	1.7 (0.88)	1.5 (0.89)	1.5 (0.85)	0.19 (-1.7, 2.1)	0.05 (-1.9, 1.9)

CFB, change from baseline; CI, confidence interval; ECG, electrocardiogram; PBO, placebo; SD, standard deviation; SE, standard error. Confidence interval widths have not been adjusted for multiplicity and may not be used for hypothesis testing.

Table S17. Change From Baseline in Sex Hormones at Week 52 (Safety Population)

	LS Mean %CFB or CFB (SE) Resmetirom 80 mg (n = 322)	LS Mean %CFB or CFB (SE) Resmetirom 100 mg (n = 323)	LS Mean %CFB or CFB (SE) Placebo (n = 321)	LS Mean Difference Resmetirom 80 mg from PBO (95% CI)	LS Mean Difference Resmetirom 100 mg from PBO (95% CI)
Estradiol, ng/L (female)					
no.	160	147	155		
Baseline mean (SD)	28.6 (37.0)	32.1 (56.3)	32.8 (65.9)		
Week 52 CFB (SE)	17.7 (8.0)	30.6 (8.3)	1.8 (8.0)	15.9 (-1.3, 33.1)	28.8 (11.3, 46.3)
Estradiol, ng/L (male)					
no.	118	118	128		
Baseline mean (SD)	28.0 (11.6)	27.6 (10.9)	29.3 (12.1)		
Week 52 CFB (SE)	8.9 (1.3)	11.0 (1.3)	-0.15 (1.2)	9.0 (6.3, 11.8)	11.2 (8.4, 13.9)
FSH, mIU/mL (female)					
no.	160	148	155		
Baseline mean (SD)	39.2 (25.7)	39.3 (22.6)	39.8 (23.2)		
Week 52 CFB (SE)	-0.54 (0.89)	0.63 (0.92)	-1.3 (0.89)	0.79 (-1.1, 2.7)	2.0 (0.02, 3.9)
FSH, mIU/mL (male)					
no.	118	119	128		
Baseline mean (SD)	8.1 (7.7)	7.8 (9.7)	7.2 (6.5)		
Week 52 CFB (SE)	1.1 (0.24)	1.7 (0.24)	0.01 (0.22)	1.1 (0.57, 1.6)	1.7 (1.1, 2.2)
LH, mIU/mL (female)					
no.	160	148	155		
Baseline mean (SD)	23.5 (14.1)	24.2 (13.2)	23.3 (12.0)		
Week 52 CFB (SE)	-0.93 (0.69)	0.80 (0.72)	-0.60 (0.70)	-0.33 (-1.8, 1.2)	1.4 (-0.11, 2.9)
LH, mIU/mL (male)					
no.	118	119	128		
Baseline mean (SD)	6.3 (4.1)	6.0 (4.6)	6.1 (4.0)		
Week 52 CFB (SE)	1.7 (0.30)	1.9 (0.30)	-0.10 (0.28)	1.8 (1.1, 2.4)	2.0 (1.4, 2.7)
Testosterone, ug/L (female)					
no.	160	147	156		
Baseline mean (SD)	0.2 (0.17)	0.2 (0.16)	0.1 (0.24)		
Week 52 CFB (SE)	0.15 (0.019)	0.19 (0.020)	0 (0.019)	0.15 (0.10, 0.19)	0.19 (0.14, 0.23)

	LS Mean %CFB or CFB (SE) Resmetirom 80 mg (n = 322)	LS Mean %CFB or CFB (SE) Resmetirom 100 mg (n = 323)	LS Mean %CFB or CFB (SE) Placebo (n = 321)	LS Mean Difference Resmetirom 80 mg from PBO (95% CI)	LS Mean Difference Resmetirom 100 mg from PBO (95% CI)
Testosterone, ug/L (male)					
no.	118	118	128		
Baseline mean (SD)	3.5 (1.6)	3.7 (2.0)	3.3 (1.5)		
Week 52 CFB (SE)	2.6 (0.26)	3.5 (0.25)	0.44 (0.24)	2.2 (1.6, 2.8)	3.0 (2.5, 3.6)
Free testosterone, nmol/L (female)					
no.	121	110	106		
Baseline mean (SD)	0 (0.01)	0 (0.01)	0 (0.01)		
Week 52 CFB (SE)	0 (0.001)	0 (0.001)	0 (0.001)	0	0
Free testosterone, nmol/L (male)					
no.	116	108	127		
Baseline mean (SD)	0.2 (0.09)	0.2 (0.11)	0.2 (0.07)		
Week 52 CFB (SE)	0.04 (0.009)	0.03 (0.009)	0.02 (0.008)	0.02 (0, 0.04)	0.01 (-0.01, 0.03)
SHBG, nmol/L					
no.	275	262	283		
Baseline mean (SD)	48.8 (56.7)	45.6 (37.9)	47.2 (43.7)		
Week 52 %CFB (SE)	157.5 (10.2)	217.4 (10.3)	9.0 (9.8)	148.5 (126.6, 170.4)	208.4 (186.3, 230.5)
Week 52 CFB (SE)	60.2 (4.0)	80.7 (4.1)	1.3 (3.9)	58.9 (50.3, 67.5)	79.4 (70.7, 88.1)
SHBG, nmol/L (female)					
no.	159	145	155		
Baseline mean (SD)	58.2 (71.7)	48.9 (44.5)	55.7 (54.6)		
Week 52 %CFB (SE)	193.0 (15.3)	251.8 (16.0)	15.6 (15.5)	177.4 (144.3, 210.5)	236.2 (202.3, 270.0)
Week 52 CFB (SE)	74.1 (6.0)	94.0 (6.3)	0.82 (6.1)	73.3 (60.3, 86.2)	93.2 (79.9, 106.5)
SHBG, nmol/L (male)					
no.	116	117	128		
Baseline mean (SD)	36.0 (17.2)	41.4 (27.1)	37.0 (20.8)		
Week 52 %CFB (SE)	108.0 (11.2)	174.0 (10.9)	0.74 (10.1)	107.3 (83.4, 131.2)	173.3 (149.4, 197.2)
Week 52 CFB (SE)	41.9 (4.2)	60.8 (4.1)	1.5 (3.8)	40.4 (31.4, 49.4)	59.3 (50.3, 68.3)

CFB, change from baseline; CI, confidence interval; FSH, follicle-stimulating hormone; LH, luteinizing hormone; LS, least squares; SD, standard deviation; SE, standard error; SHBG, sex hormone binding globulin. Confidence interval widths have not been adjusted for multiplicity and may not be used for hypothesis testing.

Table S18. Change From Baseline in Thyroid Hormones at Week 52 (Safety Population)

	LS Mean %CFB or CFB (SE) Resmetirom 80 mg (n = 322)	LS Mean %CFB or CFB (SE) Resmetirom 100 mg (n = 323)	LS Mean %CFB or CFB (SE) Placebo (n = 321)	LS Mean Difference Resmetirom 80 mg from PBO (95% CI)	LS Mean Difference Resmetirom 100 mg from PBO (95% CI)
FT3, ng/L					
no.	279	265	286		
Baseline mean (SD)	3.0 (0.41)	3.0 (0.48)	3.0 (0.41)		
Week 52 CFB (SE)	-0.01 (0.030)	-0.08 (0.031)	-0.03 (0.029)	0.02 (-0.05, 0.08)	-0.05 (-0.12, 0.01)
FT3, ng/L (not on thyroxine)					
no.	248	229	245		
Baseline mean (SD)	3.0 (0.40)	3.0 (0.42)	3.1 (0.39)		
Week 52 CFB (SE)	-0.01 (0.032)	-0.08 (0.033)	-0.02 (0.031)	0.01 (-0.06, 0.08)	-0.06 (-0.13, 0.01)
FT3, ng/L (thyroxine-treated)					
no.	31	36	41		
Baseline mean (SD)	2.7 (0.38)	2.8 (0.72)	2.8 (0.41)		
Week 52 CFB (SE)	0.04 (0.089)	-0.03 (0.081)	-0.02 (0.079)	0.05 (0.13, 0.23)	-0.01 (-0.19, 0.17)
FT4, ng/dL					
no.	279	265	286		
Baseline mean (SD)	1.1 (0.19)	1.1 (0.21)	1.1 (0.17)		
Week 52 %CFB (SE)	-13.9 (0.96)	-18.1 (0.97)	2.6 (0.92)	-16.6 (-18.6, -14.5)	-20.7 (-22.8, -18.6)
Week 52 CFB (SE)	-0.17 (0.011)	-0.22 (0.011)	0.02 (0.010)	-0.19 (-0.21, -0.16)	-0.24 (-0.26, -0.21)
FT4, ng/dL (not on thyroxine)					
no.	248	229	245		
Baseline mean (SD)	1.1 (0.18)	1.1 (0.18)	1.1 (0.16)		
Week 52 %CFB (SE)	-13.8 (0.97)	-17.6 (1.0)	2.5 (0.95)	-16.3 (-18.4, -14.2)	-20.0 (-22.2, -17.9)
Week 52 CFB (SE)	-0.16 (0.011)	-0.21 (0.011)	0.02 (0.010)	-0.18 (-0.20, -0.16)	-0.23 (-0.25, -0.20)
FT4, ng/dL (thyroxine-treated)					
no.	31	36	41		
Baseline mean (SD)	1.3 (0.23)	1.2 (0.31)	1.2 (0.21)		
Week 52 %CFB (SE)	-14.0 (3.6)	-20.6 (3.3)	3.8 (3.1)	-17.9 (-25.2, -10.5)	-24.4 (-31.6, -17.3)
Week 52 CFB (SE)	-0.18 (0.041)	-0.26 (0.037)	0.02 (0.036)	-0.21 (-0.29, -0.12)	-0.29 (-0.37, -0.21)

	LS Mean %CFB or CFB (SE) Resmetirom 80 mg (n = 322)	LS Mean %CFB or CFB (SE) Resmetirom 100 mg (n = 323)	LS Mean %CFB or CFB (SE) Placebo (n = 321)	LS Mean Difference Resmetirom 80 mg from PBO (95% CI)	LS Mean Difference Resmetirom 100 mg from PBO (95% CI)
rT3, ng/dL					
no.	278	269	282		
Baseline mean (SD)	18.5 (5.4)	19.2 (6.2)	18.4 (5.6)		
Week 52 CFB (SE)	-4.6 (0.31)	-5.1 (0.32)	0.17 (0.30)	-4.73 (-5.4, -4.1)	-5.2 (-5.9, -4.6)
rT3, ng/dL (not on thyroxine)					
no.	247	232	244		
Baseline mean (SD)	18.3 (5.3)	18.7 (5.7)	18.3 (5.6)		
Week 52 CFB (SE)	-4.5 (0.33)	-4.9 (0.34)	0.19 (0.33)	-4.5 (-5.4, -3.9)	-5.1 (-5.9, -4.8)
rT3, ng/dL (thyroxine-treated)					
no.	31	37	38		
Baseline mean (SD)	20.7 (6.1)	22.2 (8.1)	19.1 (5.6)		
Week 52 CFB (SE)	-5.1 (0.94)	-6.3 (0.85)	-0.01 (0.84)	-5.1 (-7.1, -3.1)	-6.3 (-8.2, -4.4)
FT3/rT3					
no.	278	269	282		
Baseline mean (SD)	0.18 (0.058)	0.17 (0.054)	0.18 (0.058)		
Week 52 CFB (SE)	0.06 (0.004)	0.06 (0.004)	0 (0.004)	0.06 (0.05, 0.07)	0.07 (0.06, 0.08)
FT3/rT3 (not on thyroxine)					
no.	247	232	244		
Baseline mean (SD)	0.18 (0.057)	0.18 (0.051)	0.18 (0.057)		
Week 52 CFB (SE)	0.06 (0.005)	0.06 (0.005)	0 (0.005)	0.06 (0.05, 0.07)	0.07 (0.06, 0.08)
FT3/rT3 (thyroxine-treated)					
no.	31	37	38		
Baseline mean (SD)	0.14 (0.051)	0.14 (0.058)	0.16 (0.065)		
Week 52 CFB (SE)	0.05 (0.011)	0.06 (0.010)	-0.01 (0.010)	0.06 (0.04, 0.09)	0.06 (0.04, 0.09)
TSH, mIU/L					
no.	279	265	286		
Baseline mean (SD)	2.0 (1.1)	2.1 (1.2)	2.0 (1.1)		
Week 52 CFB (SE)	-0.28 (0.059)	-0.20 (0.060)	-0.10 (0.057)	-0.18 (-0.31, -0.05)	-0.10 (-0.23, 0.03)

	LS Mean %CFB or CFB (SE) Resmetirom 80 mg (n = 322)	LS Mean %CFB or CFB (SE) Resmetirom 100 mg (n = 323)	LS Mean %CFB or CFB (SE) Placebo (n = 321)	LS Mean Difference Resmetirom 80 mg from PBO (95% CI)	LS Mean Difference Resmetirom 100 mg from PBO (95% CI)
TSH, mIU/L (not on thyroxine)					
no.	248	229	245		
Baseline mean (SD)	2.0 (1.0)	2.0 (1.1)	1.9 (0.98)		
Week 52 CFB (SE)	-0.23 (0.056)	-0.20 (0.058)	-0.08 (0.055)	-0.15 (-0.27, -0.03)	-0.12 (-0.24, 0)
TSH, mIU/L (thyroxine-treated)					
no.	31	36	41		
Baseline mean (SD)	2.0 (1.9)	2.6 (1.5)	2.2 (1.8)		
Week 52 CFB (SE)	-0.63 (0.27)	-0.13 (0.25)	-0.22 (0.24)	-0.41 (-0.97, 0.15)	0.09 (-0.46, 0.63)
TBG, mg/L					
no.	269	261	283		
Baseline mean (SD)	24.4 (7.3)	24.5 (8.1)	24.3 (8.1)		
Week 52 CFB (SE)	-0.32 (0.45)	-0.60 (0.45)	1.6 (0.43)	-1.9 (-2.8, -0.91)	-2.2 (-3.1, -1.2)
TBG, mg/L (not on thyroxine)					
no.	238	225	244		
Baseline mean (SD)	24.4 (7.5)	24.1 (7.7)	23.8 (7.7)		
Week 52 CFB (SE)	-0.19 (0.47)	-0.21 (0.49)	1.6 (0.45)	-1.8 (-2.8, -0.74)	-1.8 (-2.8, -0.74)
TBG, mg/L (thyroxine-treated)					
no.	31	36	39		
Baseline mean (SD)	24.9 (5.8)	26.6 (9.8)	27.5 (9.7)		
Week 52 CFB (SE)	-1.1 (1.4)	-3.1 (1.3)	1.5 (1.3)	-2.6 (-5.5, 0.43)	-4.6 (-7.5, -1.8)

CFB, change from baseline; CI, confidence interval; LS, least squares; rT3, reverse triiodothyronine; SE, standard error; FT3, free triiodothyronine; FT4, free thyroxine; TBG, thyroxine binding globulin; TSH, thyroid-stimulating hormone. Confidence interval widths have not been adjusted for multiplicity and may not be used for hypothesis testing.

Table S19. Shift Table of Bone Mineral Density T Score Risk Category

	Baseline								
	Resmetirom 80 mg (N = 78)			Resmetirom 100 mg (N = 61)			Placebo (N = 82)		
	Normal N (%)	Low Density N (%)	Possible Osteoporosis N (%)	Normal N (%)	Low Density N (%)	Possible Osteoporosis N (%)	Normal N (%)	Low Density N (%)	Possible Osteoporosis N (%)
Week 52									
Femoral Neck									
Normal	40 (51.3)	2 (2.6)	0	32 (52.5)	3 (4.9)	0	41 (50.0)	2 (2.4)	0
Low Density	6 (7.7)	25 (32.1)	0	2 (3.3)	21 (34.4)	0	4 (4.9)	31 (37.8)	0
Possible Osteoporosis	0	1 (1.3)	1 (1.3)	0	1 (1.6)	0	0	0	0
Hip									
Normal	67 (85.9)	1 (1.3)	0	47 (77.0)	0	0	64 (78.0)	1 (1.2)	0
Low Density	1 (1.3)	6 (7.7)	0	2 (3.3)	10 (16.4)	0	2 (2.4)	11 (13.4)	0
Possible Osteoporosis	0	0	0	0	0	0	0	0	0
Spine									
Normal	54 (69.2)	2 (2.6)	0	45 (73.8)	2 (3.3)	0	52 (63.4)	2 (2.4)	0
Low Density	2 (2.6)	14 (17.9)	0	1 (1.6)	8 (13.1)	0	4 (4.9)	18 (22.0)	0
Possible Osteoporosis	0	1 (1.3)	1 (1.3)	0	1 (1.6)	3 (4.9)	0	0	2 (2.4)

Note: Category Criteria: Normal = T Score ≥-1.0; Low Density = T Score ≥-2.5 and <-1.0; Possible Osteoporosis = T Score <-2.5
Observed Data (Primary Analysis Population, Subgroup: Female Subjects Not Taking Thyroxine at Baseline, Estradiol <30 ng/L at Baseline, and Weight Loss <5% at Week 52; Spine Adjusted Total; L1-L4Note: Column headers are baseline status and row headers are status at the post-baseline visit.