Title

MRI Quantification of Placebo Effect in Nonalcoholic Steatohepatitis Clinical Trials.

Journal

Radiology, 306(3)

Authors

Nedrud, Marybeth

Chaudhry, Mohammad

Middleton, Michael

et al.

Publication Date

2023-03-01

DOI

10.1148/radiol.220743

Peer reviewed

Radiology

ORIGINAL RESEARCH • GASTROINTESTINAL IMAGING

MRI Quantification of Placebo Effect in Nonalcoholic Steatohepatitis Clinical Trials

*Marybeth A. Nedrud, MD, PhD** • *Mohammad Chaudhry, MBBS** • *Michael S. Middleton, MD, PhD* • *Cynthia A. Moylan, MD, MHS* • *Reginald Lerebours, MA* • *Sheng Luo, PhD* • *Alfredo Farjat, PhD* • *Cynthia Guy, MD* • *Rohit Loomba, MD, MHSc* • *Manal F. Abdelmalek, MD, MPH* • *Claude B. Sirlin, MD* • *Mustafa R. Bashir, MD*

From the Department of Radiology (M.A.N., M.R.B.), Division of Gastroenterology, Department of Medicine (C.A.M., M.R.B.), Department of Biostatistics & Bioinformatics (R. Lerebours, S.L., A.F.), Department of Pathology (C.G.), and Center for Advanced Magnetic Resonance Development (M.R.B.), Duke University Medical Center, Department of Radiology, Box 3808, Durham, NC 27710; Rutgers University Hospital, School of Medicine, Newark, NJ (M.C.); Liver Imaging Group, Department of Radiology (M.S.M., C.B.S.), and Division of Gastroenterology, Department of Medicine (R. Loomba), University of California at San Diego School of Medicine, San Diego, Calif; Department of Medicine, Durham Veterans Affairs Medical Center, Durham, NC (C.A.M.); and Division of Gastroenterology and Hepatology, Mayo Clinic, Rochester, Minn (M.F.A.). Received April 22, 2022; revision requested June 1; revision received July 21; accepted September 9. **Address correspondence to** M.A.N. (email: *marybeth.nedrud@duke.edu*).

* M.A.N. and M.C. contributed equally to this work.

Conflicts of interest are listed at the end of this article.

See also the editorial by Yoon in this issue.

Radiology 2023; 306(3):e220743 • https://doi.org/10.1148/radiol.220743 • **Content codes:** GI MR

Background: Several early-phase clinical trials for the treatment of nonalcoholic steatohepatitis (NASH) use liver fat content as measured with the MRI-derived proton density fat fraction (PDFF) for a primary outcome. These trials have shown relative reductions in liver fat content with placebo treatment alone, a phenomenon termed "the placebo effect." This phenomenon confounds the results and limits generalizability to future trials.

Purpose: To quantify the effect of placebo treatment on change in the absolute PDFF value and to identify variables associated with this observed change.

Materials and Methods: This is a secondary analysis of prospectively collected data from seven early phase clinical trials that included participants with a diagnosis of NASH based on MRI and/or liver biopsy who received placebo treatment. The primary outcome was a greater than or equal to 30% relative reduction in PDFF after placebo treatment. Normalization of PDFF, relative change in alanine aminotransferase (ALT) level, and normalization of ALT level were also examined. An exploratory linear mixed-effects model was used to estimate an overall change in absolute PDFF and to explore parameters associated with this response.

Results: A total of 187 participants (median age, 52 years [IQR, 43–60 years]; 114 women) who received placebo treatment were evaluated. A greater than or equal to 30% relative reduction in baseline PDFF was seen in 20% of participants after 12 weeks of placebo treatment (10 of 49), 9% of participants after 16 weeks (two of 22), and 28% of participants after 24 weeks (34 of 122). A repeated-measures linear mixed-effects model estimated a decrease of 2.3 units (median relative reduction of 13%) in absolute PDFF values after 24 weeks of placebo treatment (95% CI: 3.2, 1.4; $P < .001$).

Conclusion: In this analysis of 187 participants, a clinically relevant decrease in PDFF was observed with placebo treatment. Based on the study model, assuming an absolute PDFF decrease of approximately 3 units (upper limit of 95% CI) to account for this "placebo effect" in sample size calculations for future clinical trials is suggested.

Clinical trial registration nos. NCT01066364, NCT01766713, NCT01963845, NCT02443116, NCT02546609, NCT02316717, and NCT02442687

Supplemental material is available for this article.

Nonalcoholic fatty liver disease, including the progressive liver disease nonalcoholic steatohepatitis (NASH), is the most common liver disorder in Western countries and affects 25% of the worldwide population (1–3). There are no current U.S. Food and Drug Administration–approved medications for the treatment of NASH, and numerous placebo-controlled clinical trials are underway in search of potential therapies (4). Many early phase trials use noninvasive imaging markers as primary outcomes to assess treatment response. In particular, the proton density fat fraction (PDFF) as measured using MRI accurately quantifies liver triglyceride content and assesses steatosis grade (5–9). Because this imaging marker is highly reproducible and robust across scanner types, its use is gaining popularity in clinical trials that assess changes in liver fat content as a primary end point (10–14).

Prior trials for potential NASH treatments have demonstrated substantial changes in primary end points after placebo treatment alone, a phenomenon termed "the placebo effect." In some trials, the placebo effect is equal to or greater than the effect of the drug under investigation, significantly confounding the results of the study and limiting generalizability to larger populations (15). The groups who received placebo in these trials were typically small and

Abbreviations

ALT = alanine aminotransferase, NASH = nonalcoholic steatohepatitis, PDFF = proton density fat fraction

Summary

In nonalcoholic steatohepatitis clinical trials, placebo treatment resulted in decreased liver fat as assessed using the MRI-derived proton density fat fraction.

Key Results

- This secondary analysis of pooled data from seven clinical trials included 187 participants with a diagnosis of nonalcoholic steatohepatitis who received placebo treatment.
- A greater than or equal to 30% relative reduction in MRI proton density fat fraction (PDFF) was observed in 28% of participants after 24 weeks of placebo treatment.
- After 24 weeks of placebo treatment, a median relative reduction in baseline PDFF of 13% was observed ($P < .001$).

imaging marker assessments were limited to a few time points, thus limiting the ability to assess the magnitude of the placebo response for generalizations that can inform future trial designs.

The purpose of this study was to quantify the effect of placebo treatment on change in absolute PDFF values for 187 pooled participants who received placebo in seven different early-phase NASH treatment trials. An additional goal was to identify variables associated with change in absolute PDFF over time.

Materials and Methods

Study Design

This was a Health Insurance Portability and Accountability Act–compliant, institutional review board–approved secondary analysis of prospectively collected data in placebo-treated participants pooled from seven early-phase NASH clinical trials from 2012 to 2018 (registered with ClinicalTrials.gov and hereafter termed "parent trials"). The published parent trials include NCT01066364 with 25 placebo participants (16), NCT01766713 with 25 placebo participants (9), NCT01963845 with 25 placebo participants (17), NCT02443116 with 27 placebo participants (18), and NCT02546609 with 22 of 24 placebo participants (two were excluded due to missing data) (19). In addition, unpublished trials NCT02316717 and NCT02442687 had 41 and 22 placebo participants, respectively. Trials were included because primary data were accessible. Participants gave written informed consent. Trial sponsors gave permission to analyze data and had no role in study design, analysis, or manuscript preparation.

Participants

The sample included 187 participants who received placebo treatment and underwent MRI for PDFF estimation at baseline and at least one additional time point (per parent trial protocol). Participants met individual inclusion and exclusion criteria for each parent trial. Common inclusion criteria included adults who gave informed consent and had a diagnosis of NASH within the past 12 months based on MRI (baseline PDFF of 6%–15%), alanine aminotransferase (ALT) level (maximum 19 U/L for women and 30 U/L for men), and/or liver biopsy. Common exclusion criteria were the presence of chronic liver disease other than NASH, including hepatitis B or C, autoimmune hepatitis, hemochromatosis, Wilson disease, and alpha-1 antitrypsin deficiency, as well as previous and/or current excessive alcohol use. Additional exclusion criteria unique to each trial are provided in Table S1. In brief, multiple trials excluded participants with decompensated liver cirrhosis, recent weight loss, prior bariatric surgery, type 1 diabetes, recent change in diabetic medications, and use of agents known to affect steatosis.

Data Collection

Baseline clinical data included age, sex, and self-reported race and ethnicity according to funding agency requirements. Ad-

Table 1: Baseline Demographics of Participants across the Seven Parent Trials

Characteristic	Total (n = 187)	Trial 1 (n = 25)	Trial 2 (n = 41)	Trial 3 (n = 25)	Trial 4 (n = 27)	Trial 5 (n = 22)	Trial 6 (n = 25)	Trial 7 (n = 22)	P Value
Age (y)*	52 (43–60)	53 (45–58)	50 (41–57)	49 (41–61)	56 (46–52)	46 (41–55)	57 (46–64)	50 (40–60)	.3†
Sex									
F	114 (61)	13 (52)	20 (49)	17 (68)	20 (74)	11 (50)	17 (68)	16 (73)	.2‡
M	73 (39)	12 (48)	21 (51)	8 (32)	7 (26)	11 (50)	8 (32)	6 (27)	
Race and ethnicity									
Hispanic	50 (27)	7 (28)	3 (7)	9 (36)	12 (44)	5 (23)	9 (36)	5 (23)	<.001‡
White, non-Hispanic	105 (56)	8 (32)	32 (78)	12 (48)	13 (48)	14 (64)	10 (40)	16 (73)	
Body mass index (kg/m^2)*	32.8 (30.1–37.7)	31.9 (26.4–36.2)	33.2 (30.8–37.9)	32.5 (30.4–35.0)	34.7 (30.7–39.5)	32.1 (30.1–36.8)	30.7 (28.4–33.5)	35.6 (31.4–40.0)	.06†
Type 2 diabetes	77 (41)	10 (40)	17 (42)	7 (28)	18 (67)	0 (0)	13 (52)	12 (55)	<.001†

Note.—Except where indicated, data are numbers of participants, with percentages in parentheses.

* Data are medians, with IQRs in parentheses.

† Kruskal-Wallis test.

‡ χ^2 test.

Table 2: Baseline Biochemical, Histologic, and Imaging Information of Participants across the Seven Parent Trials

Variable	Total (n = 187)	Trial 1 (n = 25)	Trial 2 (n = 41)	Trial 3 (n = 25)	Trial 4 (n = 27)	Trial 5 (n = 22)	Trial 6 (n = 25)	Trial 7 (n = 22)	P Value
ALT (U/L)*	57.0 (40.0–84.6)	73.0 (45.0–92.0)	59.0 (42.0–94.0)	47.0 (36.0–62.0)	52.8 (41.9–84.6)	63.5 (42.0–80.0)	40.0 (25.0–1.0)	70.0 (57.0–128.0)	<.001†
Total bilirubin (mg/dL)*	0.5 (0.3–0.6)	0.5 (0.4–0.6)	0.5 (0.4–0.7)	0.4 (0.3–0.5)	0.3 (0.2–0.4)	0.5 (0.5–0.7)	0.5 (0.3–0.7)	0.5 (0.4–0.7)	<.001†
Hemoglobin (g/dL)*	14.1 (13.5–15.2)	NA	14.2 (13.8–15.3)	13.8 (13.0–15.5)	NA	14.9 (12.9–15.8)	14.0 (13.3–14.9)	13.6 (12.5–14.4)	.07†
HbA_{1c} (%)*	6.1 (5.6–6.7)	6.1 (5.9–6.8)	5.9 (5.4–6.8)	6.1 (5.6–6.6)	6.6 (5.8–6.9)	5.7 (5.3–5.9)	6.2 (5.8–6.6)	6.1 (5.7–6.8)	.003†
HOMA-IR‡	8.8 ± 8.9	8.3 ± 9.8	11.5 ± 9.3	9.8 ± 7.8	7.8 ± 10.5	3.2 ± 1.1	7.4 ± 6.5	11.1 ± 10.0	<.001†
Liver inflammation									.1§
1	56 (40)	16 (64)	15 (37)	8 (32)	11 (41)	NA	NA	6 (27)	
2	74 (53)	7 (28)	23 (56)	17 (68)	14 (52)	NA	NA	13 (59)	
3	10 (7)	2 (8)	3 (7)	0 (0)	2 (7)	NA	NA	3 (14)	
Ballooning									<.001§
0	8 (6)	5 (20)	0 (0)	3 (12)	0 (0)	NA	NA	0 (0)	
1	69 (49)	15 (60)	23 (56)	12 (48)	14 (52)	NA	NA	5 (23)	
2	63 (45)	5 (20)	18 (44)	10 (40)	13 (48)	NA	NA	17 (77)	
NAFLD activity score									.2§
3 and 4	44 (31)	13 (52)	13 (32)	7 (28)	9 (33)	NA	NA	2 (9)	
5	38 (27)	7 (28)	9 (22)	5 (20)	8 (30)	NA	NA	9 (41)	
6	36 (26)	4 (16)	11 (27)	8 (32)	6 (22)	NA	NA	7 (32)	
7	21 (15)	0 (0)	8 (20)	5 (20)	4 (15)	NA	NA	4 (18)	
8	1 (0.7)	1 (4)	0 (0)	0 (0)	0 (0)	NA	NA	0 (0)	
Fibrosis									<.001§
0 and 1	67 (48)	18 (72)	18 (44)	16 (64)	11 (41)	NA	NA	4 (18)	
2	34 (24)	2 (8)	16 (39)	2 (8)	7 (26)	NA	NA	7 (32)	
3	36 (26)	3 (12)	7 (17)	6 (24)	9 (33)	NA	NA	11 (50)	
4	3 (2)	2 (8)	0 (0)	1 (4)	0 (0)	NA	NA	0 (0)	
MRI proton density fat fraction (%)*	18.0 (11.5–22.9)	16.4 (11.3–21.7)	17.3 (14.0–21.4)	19.6 (13.2–24.4)	15.4 (11.0–20.7)	23.1 (21.3–26.3)	15.8 (10.8–21.7)	13.2 (8.0–21.5)	

Note.—Except where indicated, data are numbers of participants, with percentages in parentheses. Histologic data were available for 140 of 187 participants. ALT = alanine aminotransferase, HbA_{1c} = hemoglobin A_{1c}, HOMA-IR = homeostatic model assessment for insulin resistance, NA = not available, NAFLD = nonalcoholic fatty liver disease.

* Data are medians, with IQRs in parentheses.

† Kruskal-Wallis test.

‡ Data are means ± SDs.

§ χ^2 test.

ditional parameters included body mass index, baseline hemoglobin A_{1c} level, histologic characteristics (if biopsy was required), ALT level, aspartate aminotransferase results, total bilirubin level, homeostatic model assessment for insulin resistance results, and PDFF.

MRI PDFF Technique

Imaging was performed using 1.5-T or 3-T clinical MRI systems (Siemens Healthcare, GE Healthcare, Philips Healthcare, or Hitachi Medical Corporation) at trial sites in the United States, Australia, and Israel. Protocols were composed by one of two MRI core laboratories (Bashir Laboratory for Liver Imaging Research, Duke University and Liver Imaging Group, and University of California San Diego) and distributed to sites in manuals. Protocols used either a six-echo or dual-dual echo chemical shift-encoded gradient-echo technique, depending on sequence availability, with low flip angles to minimize T1 bias. Imaging parameters included repetition time, greater than or equal to 150 msec; flip angle, less than or equal to 15° at 1.5 T and less than or equal to 10° at 3 T; image matrix, 128–256 × 92–256. For the six-echo sequences, nominal in-phase and opposed-phase echo times were used based on field strength. For the dual-dual echo sequences, two acquisitions were performed; the first used the two earliest in-phase echo times, while the second used the first opposed-phase and first in-phase echo times.

To ensure protocol adherence, deidentified digital imaging and communications in medicine, or DICOM, images were sent to a core laboratory for quality control and analysis. PDFF measurements were performed using algorithms within

an image viewer (OsiriX version 7.5.1; Pixmeo) or in MATLAB (MathWorks). Both techniques have been validated against MR spectroscopy and incorporate corrections for T2* effects and spectral fat complexity (8,20,21). For dual-dual echo sequences, the initial pair of in-phase echoes was used to estimate T2* based on monoexponential decay, and the regional T2* estimate was used along with the in-phase and opposed-phase data to estimate PDFF.

Outcome Measures

A primary outcome of greater than or equal to 30% relative reduction in PDFF after 12, 16, and/or 24 weeks of placebo treatment was selected, as this metric corresponds to histologic NASH resolution (22). Secondary outcomes after 12, 16, and/or 24 weeks of placebo treatment include *(a)* PDFF normalization, defined as a less than or equal to 5% absolute PDFF value; *(b)* greater than 50% relative reduction in ALT level; and *(c)* ALT normalization, defined as less than or equal to 30 U/L for men and less than or equal to 19 U/L for women.

Statistical Analysis

Baseline data were analyzed using medians with IQRs for continuous variables or frequency and percentages for categorical variables. Differences were compared using the Kruskal-Wallis test for continuous variables and χ^2 test for categorical variables. Changes in relative PDFF and ALT values were examined with the Student *t* test or Wilcoxon signed-rank test. A repeated-measures linear mixed-effects model with a compound symmetry covariance was used to evaluate the relationship between absolute PDFF and time (the placebo effect). The multiple imputation by chained equations (MICE)

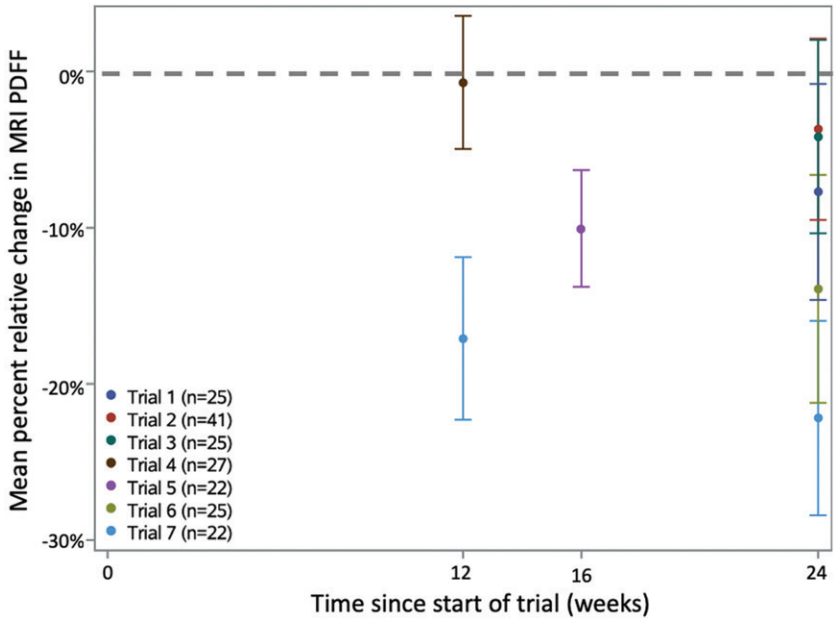

Figure 1: Dot plot shows the mean and standard error for relative change in baseline MRI proton density fat fraction (PDFF) values for all participants (*n* = 187) in the study at all time points across the seven parent trials.

Table 3: Relative Change Rate in MRI PDFF and ALT Level across the Seven Parent Trials

% Change from Baseline	Total (*n* = 187)	Trial 1 (*n* = 25)	Trial 2 (*n* = 41)	Trial 3 (*n* = 25)	Trial 4 (*n* = 27)	Trial 5 (*n* = 22)	Trial 6 (*n* = 25)	Trial 7 (*n* = 22)	*P* Value
Mean PDFF at 12 weeks	–8.1 ± 24.4	…	…	…	–0.7 ± 22.2	…	…	–17.1 ± 24.4	.03*
Mean PDFF at 16 weeks	–10.1 ± 17.5	…	…	…	…	–10.1 ± 17.5	…	…	.01*
Mean PDFF at 24 weeks	–9.0 ± 32.8	–7.7 ± 32.4	–3.7 ± 37.2	–4.2 ± 29.0	…	…	–13.9 ± 30.1	–22.2 ± 27.8	.003*
No. of missing data points	14†	3	…	3	…	…	8	2	
Median ALT at 12 weeks	–13.0 (–28.8 to 12.0)	…	…	…	7.5 (–13.2 to 21.6)	…	…	–24.2 (–41.5 to –13.0)	.10‡
Median ALT at 16 weeks)	–0.1 (–31.4 to 18.8)	…	…	…	…	–0.1 (–31.4 to 18.8)	…	…	.86‡
Median ALT at 24 weeks	–15.0 (–35.7 to 8.3)	–28.1 (–42.6 to 5.7)	–16.7 (–35.7 to 5.8)	–2.4 (–21.8 to 18.8)	…	…	–4.5 (–32.3 to 15.4)	–28.1 (–50.9 to –7.0)	<.001‡
No. of missing data points	13	2	1	2	4	…	3	1	

Note.—Except were indicated, data are means ± SDs or medians, with IQRs in parentheses. *P* values test the null hypothesis that the mean or median of the variable of interest is equal to 0 for all participants at that time point (regardless of trial). ALT = alanine aminotransferase, PDFF = proton density fat fraction.

* Student *t* test.

† Two participants with missing data at 24 weeks in trial 7 had available data at 12 weeks.

‡ Wilcoxon signed-rank test.

method was used for missing information (23,24). Because missingness can be attributed to trial collection, data were assumed missing at random (25). The imputation was aggregated over 10 sets. Candidate variables were added to a linear mixed-effects model along with a priori selected explanatory variables. Analyses were conducted using R version 3.4.3 (The R Foundation) and SAS version 9.4 (SAS Institute). Two-sided tests were employed ($P < .05$), and a threshold for assessing a statistically significant difference was set at $\alpha = .05$.

Table 4: Observed Change in MRI PDFF and ALT Level across Time Points

Variable	Week 12	Week 16	Week 24
MRI PDFF characteristics			
≥30% relative reduction in PDFF	10/49 (20)	2/22 (9)	34/122 (28)
≥30% relative increase in PDFF	2/49 (4)	0/22 (0)	11/122 (9)
PDFF normalization (<5% at end of study)	2/49 (4)	0/22 (0)	6/122 (5)
Biochemical characteristics			
≥50% relative reduction in ALT level	3/45 (7)	2/22 (9)	15/129 (12)
ALT normalization (ALT ≤19 U/L for women, ≤30 U/L for men)	0/45 (0)	1/22 (5)	8/129 (6)

Note.—Except were indicated, data are numbers of participants, with percentages in parentheses. ALT = alanine aminotransferase, PDFF = proton density fat fraction.

Results

Participant Characteristics

As shown in Table 1, a total of 187 participants met inclusion criteria for individual parent trials and received placebo treatment during each parent trial. The median age of participants was 52 years (IQR, 43–60 years), with 114 women. Fifty-six percent (105 of 187) of participants self-identified as non-Hispanic White and 27% (50 of 187) self-identified as Hispanic (Table 1). The median body mass index was 32.8 kg/m^2 (IQR, 30.1–37.7 kg/m^2; $n = 186$) and 41% (77 of 187) of participants had diabetes, with a median baseline hemoglobin A_{1c} level of 6.1% (IQR, 5.6%–6.7%; $n = 184$) (Table 2). The study sample median baseline ALT level was 57.0 U/L (IQR, 40.0–84.6 U/L) and median baseline MRI PDFF was 18.0% (IQR, 11.5%–22.9%). There was no evidence of differences in participant age, sex, and body mass index across parent trials. Differences were observed across the parent trials for race and ethnicity, diabetes diagnosis, baseline ALT level, total bilirubin level, hemoglobin A_{1c} level, homeostatic model assessment for insulin resistance result, baseline ballooning, and baseline fibrosis (Tables 1, 2). Two of seven parent trials collected data after 12 weeks of placebo treatment in 49 of 187 participants. One parent trial collected data at 16 weeks in 22 of 187 participants. Five of the seven trials collected data at 24 weeks in 122 of 187 participants. A single trial included data at both the 12-week and 24-week time points in 20 participants. Actual data collected, including missing data points for each trial, are included in Table 3.

Relative Reduction in MRI PDFF

As shown in Figure 1, we observed a relative reduction in the mean MRI proton density fat fraction (PDFF) at all time points examined among the parent trials. Table 3 demonstrates the relative change in PDFF observed for each trial at all their specified end points. A statistically significant decrease in baseline PDFF

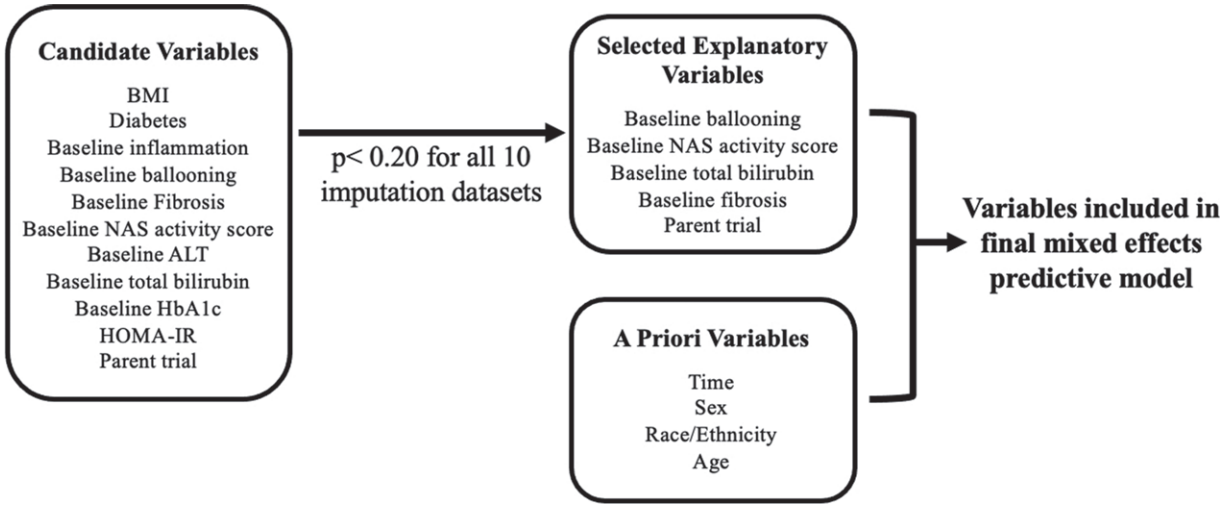

Figure 2: Schematic shows the model selection process and variables for the linear mixed-effects model used to evaluate the percentage change in baseline MRI proton density fat fraction (PDFF). ALT = alanine aminotransferase, BMI = body mass index, HbA1c = hemoglobin A_{1c}, HOMAR-IR = homeostatic model assessment for insulin resistance, NAS = nonalcoholic fatty liver disease activity score.

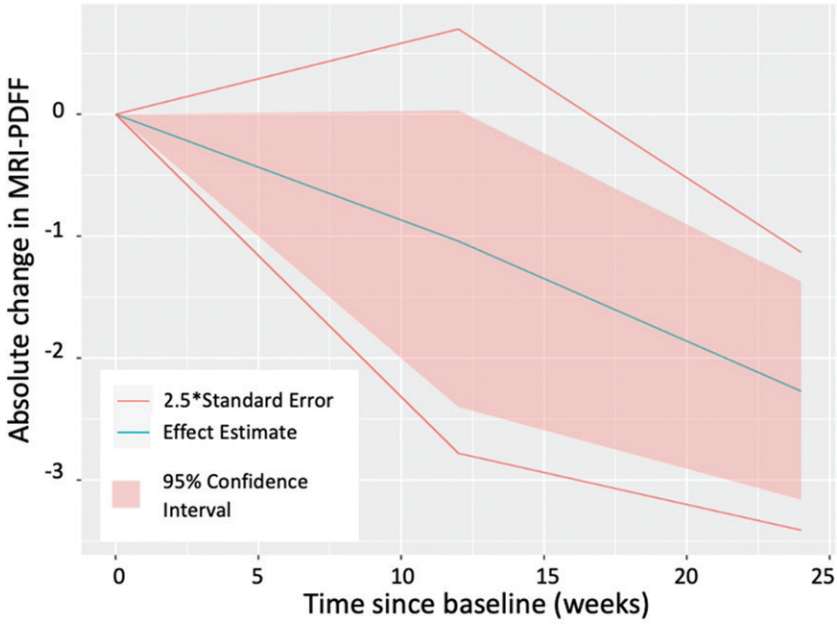

Figure 3: Line graph shows absolute change in MRI proton density fat fraction (PDFF) as a function of duration of placebo treatment (weeks) for our repeated-measures linear mixed-effects model, which estimates a mean absolute reduction in baseline MRI PDFF of 2.3 units after 24 weeks of placebo treatment.

was observed with 12 weeks of placebo treatment (mean relative change, −8.1% ± 24.4 [SD]; P = .03), 16 weeks of placebo treatment (mean relative change, −10.1% ± 17.5; P = .01), and 24 weeks of placebo treatment (mean relative change, −9.0% ± 32.8; P = .003). A greater than or equal to 30% relative reduction in MRI PDFF correlates with a change in liver fat on histologic reports and is, therefore, an important benchmark indicating clinically meaningful changes in hepatic steatosis and resolution of NASH (22). A greater than or equal to 30% relative reduction in PDFF was observed in 20% (10 of 49) of participants after 12 weeks, 9% (two of 22) of participants after 16 weeks, and 28% (34 of 122) of participants after 24 weeks of placebo treatment (Table 4). By comparison, a relative increase in PDFF of greater than or equal to 30% was observed in 4% (two of 49) of participants after 12 weeks, 0% (0 of 22) of participants after 16 weeks, and 9% (11 of 122) of participants after 24 weeks of placebo treatment. As shown in Table 4, PDFF normalization, defined as a PDFF value less than or equal to 5% at follow-up, was observed in 4% (two of 49) of participants at 12 weeks, 0% (0 of 22) of participants at 16 weeks, and 5% (six of 122) of participants at 24 weeks. Figure S1 shows the relative change in baseline PDFF values for all participants at each time point and demonstrates marked variability in participant PDFF values as a response to placebo treatment over time.

Relative Reduction in ALT

Clinically relevant reductions in ALT, defined as a greater than or equal to 50% relative reduction in the ALT level from baseline, were observed in 7% (three of 45) of participants after 12 weeks, 9% (two of 22) of participants after 16 weeks, and 12% (15 of 129) of participants after 24 weeks of placebo treatment (Table 4). There was no evidence of a relative percent change in ALT level from baseline after 12 weeks (median change, −13.0% [IQR, −28.8% to 12.0%]; P = .10) or 16 weeks (median change, −0.1% [IQR, −31.4% to 18.8%]; P = .86) of placebo treatment (Table 3). However, there was a relative percent change after 24 weeks, with a calculated median ALT reduction from baseline of −15.0% (IQR, −35.7% to 8.3%; P < .001). Normalization of ALT, defined as an ALT level less than or equal to 30 U/L for men and less than or equal to 19 U/L for women at clinical follow-up, was observed in 5% (one of 22) of participants at 16 weeks and 6% (eight of 129) of participants at 24 weeks. No participants demonstrated ALT normalization at 12 weeks.

Association between Absolute Change in MRI PDFF and Explanatory Variables

After applying the variable selection criteria, the parameters of baseline ballooning, baseline nonalcoholic fatty liver disease activity score, baseline total bilirubin level, baseline fibrosis, and parent trial were selected as variables for the final exploratory mixed-effects model along with the a priori determined explanatory variables of time, sex, race and ethnicity, and age (Fig 2). Correlated participant observations nested within trials were treated as repeated measures. The model estimated a 2.3-unit decrease in absolute PDFF after 24 weeks following the start of the trial (95% CI: 3.2, 1.4; P < .001) (Fig 3). In the current pool of participants, with a median baseline PDFF of 18.0%, this corresponds to a relative reduction in the median PDFF after placebo treatment of approximately 13% (95% CI: 18, 8). As shown in Table 5, key histologic parameters that included high levels of ballooning in the histologic report (level 2 ballooning: effect estimate, −6.2; 95% CI: −12.5, 0.01; P = .05), higher levels of fibrosis in the histologic report (level 3 or 4 fibrosis: effect estimate, −3.8; 95% CI: −6.9, −0.7; P = .02), total bilirubin level (effect estimate, −3.5; 95% CI: −6.7, −0.04; P = .03), and a nonalcoholic fatty liver disease activity score greater than 5 (effect estimate, 5.7; 95% CI: 3.0, 8.4; P < .001) were predicative of baseline PDFF. None of these parameters were associated with a change in baseline PDFF over time (data not shown). Adjusted linear estimates for all selected variables are provided in Table 5.

Discussion

Given the noninvasive nature of imaging, clinical trials for nonalcoholic steatohepatitis (NASH) are moving toward use of the MRI-derived proton density fat fraction (PDFF) to quantify changes in liver fat content as a response to medical intervention (26). Multiple early-phase trials have demonstrated a significant change in relative PDFF values with placebo treatment alone, a phenomenon termed "the placebo effect." Unfortunately, early

Table 5: Estimated Effect from Linear Mixed Model on Absolute Percentage Change in MRI PDFF with Placebo Treatment

Covariate	Effect Estimate	*P* Value
Intercept	23.5 (15.9, 31.1)	<.001
Time since start of trial [*P* < .001]		
Baseline	Reference	
12 weeks	−1.0 (−2.4, 0.3)	.13
24 weeks	−2.3 (−3.2, −1.4)	<.001
Sex		
M	Reference	
F	−1.2 (−3.4, 1.1)	.31
Race and ethnicity [*P* = .51]		
White, non-Hispanic	Reference	
Hispanic	−0.8 (−3.3, 1.6)	.50
Other*	−1.4 (−4.2, 1.3)	.31
Age (y)	−0.01 (−0.1, 0.08)	.77
Baseline ballooning [*P* = .01]		
Level 0	Reference	
Level 1	−1.5 (−6.9, 3.9)	.59
Level 2	−6.2 (−12.5, 0.1)	.05
Baseline NAFLD activity score		
Levels 3 or 4	Reference	
Level 5 or higher	5.7 (3.0, 8.4)	<.001
Baseline total bilirubin level	−3.5 (−6.7, −0.4)	.03
Baseline fibrosis [*P* = .047]		
Level 1	Reference	
Level 2	−2.5 (−5.6, 0.6)	.11
Level 3 or 4	−3.8 (−6.9, −0.7)	.02
Trial [*P* = .28]		
Trial 2	Reference	
Trial 1	−1.5 (−5.1, 2.1)	.41
Trial 3	−0.7 (−4.2, 2.7)	.68
Trial 4	−1.0 (−4.4, 2.4)	.56
Trial 5	3.4 (−0.5, 7.2)	.09
Trial 6	−0.7 (−4.4, 2.9)	.72
Trial 7	−1.7 (−5.1, 1.8)	.36

Note.—Data in parentheses are 95% CIs. *P* values in brackets indicate the type III tests of fixed effects, which allow for multiple degrees of freedom when the model of interest includes categorical variables with three or more levels of interest. NAFLD = nonalcoholic fatty liver disease, PDFF = proton density fat fraction.

* Participants who self-reported non-Hispanic ethnicity (or those who did not specify ethnicity) who identified as either Asian, Black, or mixed race.

trials are often not sufficiently powered to accurately quantify the placebo effect to inform future clinical trial designs. To this end, we examined the change in PDFF in 187 participants who received placebo treatment across seven early-phase clinical trials. We observed a clinically significant relative reduction in PDFF (greater than or equal to 30% relative decrease) in 20% and 28% of participants after 12 and 24 weeks of placebo treatment, respectively. We generated a linear mixed-effects model to estimate an absolute decrease in PDFF of 2.3 units (95% CI: 3.2, 1.4), which corresponds to a relative decrease of approximately 13% in the median PDFF (95% CI: 18, 8) after 24 weeks of placebo treatment in our pool of participants. As a conservative estimate for future clinical trials, we recommend assuming an absolute decrease of approximately 3 units in baseline PDFF (upper 95% CI limit) as a result of placebo treatment for sample size calculations.

Our findings are consistent with a recent meta-analysis that showed a placebo response in NASH clinical trials with a variety of end points, including serum aspartate aminotransferase and ALT levels, histologic steatosis, MR spectroscopy, and PDFF (27). The current study extends these results by estimating the magnitude of the change in PDFF over time, a value that can be used to inform sample size calculations in future clinical trials. Our study also provides a glimpse into the large intrinsic heterogeneity, as measured with the PDFF, among participants over time. We found that 28% of participants had a relative decrease in PDFF values after 24 weeks of placebo treatment, while 9% had a relative increase in PDFF values. We did not find a specific clinical or biochemical parameter that was associated with change in the absolute PDFF over time in our mixed-effects model. Given this, the observed heterogeneity is likely due to the high sensitivity of PDFF for quantification of liver fat, as well as multiple confounding variables such as changes in participant diet and lifestyle during the observation period.

Our study had multiple limitations. First, while the pooling of data yielded increased statistical power, differing inclusion and exclusion criteria among the parent trials led to significant differences in sample demographics, as well as multiple clinical, biochemical, and histologic parameters. For example, hemoglobin A_{1c} levels varied among the parent trials (median, 6.1% [IQR, 5.6%–6.7%]; *P* = .003) as a result of differences in exclusion criteria of participants with type 2 diabetes. One trial excluded participants with a hemoglobin A_{1c} level greater than 8.0%, others excluded participants with recent changes in diabetic medications, and one trial eliminated those taking any form of metformin, effectively excluding all participants with type 2 diabetes. This specific example limits the generalizability of our study, as type 2 diabetes and NASH typically coexist in Western populations. To address the effect of varying trial inclusion and exclusion criteria on our results, correlated observations for participants nested within a trial were accounted for, along with the inclusion of the parent trial as an explanatory variable in our mixed-effects model. Although this level of variation is addressed, the proposed model specification must be tested in external data sets for confirmation.

Second, an important limitation is the absence of data regarding participant diet, caloric intake, and exercise or weight-loss regimens. It is thought that a relative change in biochemical imaging markers in response to placebo treatment results from these factors, which are outside medical intervention and not routinely assessed in data collection. To minimize this, multiple parent trials excluded participants with recent weight loss and those taking drugs known to encourage weight loss or influence hepatic steatosis, such as dipeptidyl peptidase–4 inhibitors and vitamin

E, respectively. Third, trial length varied across the included studies with short-interval follow-up at 12 and 16 weeks being underrepresented in our cohort. Given this, our model may not be generalizable to clinical trials with a shorter duration. Finally, parent trials did not correlate biochemical or imaging markers with histologic findings. While our primary end point of a greater than or equal to 30% relative reduction in PDFF is known to correlate with histologic resolution of NASH, we do not have biopsy data to confirm this in our sample (22). Lack of histologic correlation makes it unclear whether the observed changes in PDFF values are truly a placebo effect or if they represent disease resolution due to interventions outside of medical intervention.

In conclusion, this secondary analysis of pooled data from seven clinical trials included participants with nonalcoholic steatohepatitis who received placebo treatment. A greater than or equal to 30% relative reduction in MRI-derived proton density fat fraction (PDFF) values was observed in 28% of participants after 24 weeks of placebo treatment. We suggest future clinical trials aim to account for a "placebo effect" in their sample size estimates by assuming an absolute decrease in baseline PDFF of approximately 3 units with placebo treatment alone. However, to ensure the results of our analysis are unbiased, our model needs validation in additional, external data sets.

Author contributions: Guarantors of integrity of entire study, **M.A.N., C.G., M.R.B.**; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, **M.A.N., M.C., M.R.B.**; clinical studies, **M.S.M., C.A.M., C.G., R. Loomba, M.F.A., M.R.B.**; experimental studies, **C.G.**; statistical analysis, **R. Lerebours, S.L., A.F., M.R.B.**; and manuscript editing, **M.A.N., M.C., M.S.M., C.A.M., R. Lerebours, A.F., R. Loomba, M.F.A., C.B.S., M.R.B.**

Data sharing: Data generated or analyzed during the study are available from the corresponding author by request.

Disclosures of conflicts of interest: **M.A.N.** No relevant relationships. **M.C.** No relevant relationships. **M.S.M.** Institutional lab service agreements from Alexion, AstraZeneca, Bristol-Myers Squibb, Celgene, Enanta, Galmed, Genzyme, Gilead, Guerbet, Intercept, Ionis, Janssen, NuSirt, Organovo, Pfizer, Roche, Sanofi, Shire, Synageva, and Takeda; consulting fees from Alimentiv, Arrowhead, Glympse, Kowa, Median, and Novo Nordisk; stockholder, Pfizer. **C.A.M.** No relevant relationships. **R. Lerebours** No relevant relationships. **S.L.** No relevant relationships. **A.F.** No relevant relationships. **C.G.** Consulting fees from CymaBay, NGM, Madrigal, and 89bio. **R. Loomba** Consulting fees from Aardvark Therapeutics, Altimmune, Anylam/Regeneron, Amgen, Arrowhead Pharmaceuticals, AstraZeneca, Bristol-Myers Squibb, CohBar, Eli Lilly, Galmed, Gilead, Glympse Bio, Hightide, Inipharma, Intercept, Inventiva, Ionis, Janssen, Madrigal, Metacrine, NGM Biopharmaceuticals, Novartis, Novo Nordisk, Merck, Pfizer, Sagimet, Theratechnologies, 89bio, Terns Pharmaceuticals, and Viking Therapeutics; funding support from the National Center for Advancing Translational Sciences, National Institute of Diabetes and Digestive and Kidney Diseases, National Heart Lung and Blood Institute, and National Institute on Alcohol Abuse and Alcoholism; cofounder and stockholder, LipoNexus. **M.F.A.** No relevant relationships. **C.B.S.** Grants or contracts from American College of Radiology, Foundation of National Institutes of Health, Bayer, GE, Gilead, Pfizer, Philips, and Siemens; royalties or licenses from Medscape and Wolters Kluwer; institutional consulting representative, AMRA, BMS, Exact Sciences, IBM Watson, and Pfizer; personal consulting, Blade, Boehringer, and Epigenomics; leadership and stockholder, Livivos; advisory board member, QuantixBio; board of directors, Society of Abdominal Radiology; equipment loan, GE Healthcare; institutional lab service agreements, Enanta Pharmaceuticals, Gilead, ICON, Intercept, Nusirt, Shire, Synageva, and Takeda. **M.R.B.** Member of *Radiology* editorial board, institutional grant support from Madrigal Pharmaceuticals, Metacrine, Carmot, and Corcept.

References

1. Vernon G, Baranova A, Younossi ZM. Systematic review: the epidemiology and natural history of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in adults. Aliment Pharmacol Ther 2011;34(3):274–285.
2. LaBrecque DR, Abbas Z, Anania F, et al. World Gastroenterology Organisation global guidelines: Nonalcoholic fatty liver disease and nonalcoholic steatohepatitis. J Clin Gastroenterol 2014;48(6):467–473.
3. Williams CD, Stengel J, Asike MI, et al. Prevalence of nonalcoholic fatty liver disease and nonalcoholic steatohepatitis among a largely middle-aged population utilizing ultrasound and liver biopsy: a prospective study. Gastroenterology 2011;140(1):124–131.
4. Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of non-alcoholic fatty liver disease: practice Guideline by the American Association for the Study of Liver Diseases, American College of Gastroenterology, and the American Gastroenterological Association. Hepatology 2012;55(6):2005–2023.
5. Mashhood A, Railkar R, Yokoo T, et al. Reproducibility of hepatic fat fraction measurement by magnetic resonance imaging. J Magn Reson Imaging 2013; 37(6):1359–1370.
6. Reeder SB, Cruite I, Hamilton G, Sirlin CB. Quantitative Assessment of Liver Fat with Magnetic Resonance Imaging and Spectroscopy. J Magn Reson Imaging 2011;34(4):729–749.
7. Reeder SB, Hu HH, Sirlin CB. Proton density fat-fraction: a standardized MR-based biomarker of tissue fat concentration. J Magn Reson Imaging 2012;36(5):1011–1014.
8. Zhong X, Nickel MD, Kannengiesser SA, Dale BM, Kiefer B, Bashir MR. Liver fat quantification using a multi-step adaptive fitting approach with multi-echo GRE imaging. Magn Reson Med 2014;72(5):1353–1365.
9. Loomba R, Sirlin CB, Ang B, et al. Ezetimibe for the treatment of nonalcoholic steatohepatitis: assessment by novel magnetic resonance imaging and magnetic resonance elastography in a randomized trial (MOZART trial). Hepatology 2015;61(4):1239–1250.
10. Hernando D, Sharma SD, Aliyari Ghasabeh M, et al. Multisite, multivendor validation of the accuracy and reproducibility of proton-density fat-fraction quantification at 1.5T and 3T using a fat-water phantom. Magn Reson Med 2017;77(4):1516–1524.
11. Sofue K, Zhong X, Nickel MD, Dale BM, Bashir MR. Stability of liver proton density fat fraction and changes in R 2* measurements induced by administering gadoxetic acid at 3T MRI. Abdom Radiol (NY) 2016;41(8):1555–1564.
12. Tang A, Chen J, Le TA, et al. Cross-sectional and longitudinal evaluation of liver volume and total liver fat burden in adults with nonalcoholic steatohepatitis. Abdom Imaging 2015;40(1):26–37.
13. Tang A, Desai A, Hamilton G, et al. Accuracy of MR imaging-estimated proton density fat fraction for classification of dichotomized histologic steatosis grades in nonalcoholic fatty liver disease. Radiology 2015;274(2):416–425.
14. Yokoo T, Serai SD, Pirasteh A, et al. Linearity, Bias, and Precision of Hepatic Proton Density Fat Fraction Measurements by Using MR Imaging: A Meta-Analysis. Radiology 2018;286(2):486–498.
15. Diehl AM, Harrison S, Caldwell S, et al. JKB-121 in patients with nonalcoholic steatohepatitis: A phase 2 double blind randomized placebo control study. J Hepatol 2018;68(Supplement 1):S103.
16. Le TA, Chen J, Changchien C, et al. Effect of colesevelam on liver fat quantified by magnetic resonance in nonalcoholic steatohepatitis: a randomized controlled trial. Hepatology 2012;56(3):922–932.
17. Cui J, Philo L, Nguyen P, et al. Sitagliptin vs. placebo for non-alcoholic fatty liver disease: A randomized controlled trial. J Hepatol 2016;65(2):369–376.
18. Harrison SA, Rinella ME, Abdelmalek MF, et al. NGM282 for treatment of non-alcoholic steatohepatitis: a multicentre, randomised, double-blind, placebo-controlled, phase 2 trial. Lancet 2018;391(10126):1174–1185.
19. Chalasani N, Vuppalanchi R, Rinella M, et al. Randomised clinical trial: a leucine-metformin-sildenafil combination (NS-0200) vs placebo in patients with non-alcoholic fatty liver disease. Aliment Pharmacol Ther 2018;47(12):1639–1651.
20. Bashir MR, Zhong X, Nickel MD, et al. Quantification of hepatic steatosis with a multistep adaptive fitting MRI approach: prospective validation against MR spectroscopy. AJR Am J Roentgenol 2015;204(2):297–306.
21. Hu HH, Yokoo T, Bashir MR, et al. Linearity and Bias of Proton Density Fat Fraction as a Quantitative Imaging Biomarker: A Multicenter, Multiplatform, Multivendor Phantom Study. Radiology 2021;298(3):640–651.

22. Stine JG, Munaganuru N, Barnard A, et al. Change in MRI-PDFF and Histologic Response in Patients With Nonalcoholic Steatohepatitis: A Systematic Review and Meta-Analysis. Clin Gastroenterol Hepatol 2021;19(11):2274–2283.e5.
23. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? Int J Methods Psychiatr Res 2011;20(1):40–49.
24. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. 2011 2011;45:67.
25. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical guide with flowcharts. BMC Med Res Methodol 2017;17(1):162.
26. Caussy C, Reeder SB, Sirlin CB, Loomba R. Noninvasive, Quantitative Assessment of Liver Fat by MRI-PDFF as an Endpoint in NASH Trials. Hepatology 2018;68(2):763–772.
27. Han MAT, Altayar O, Hamdeh S, et al. Rates of and Factors Associated With Placebo Response in Trials of Pharmacotherapies for Nonalcoholic Steatohepatitis: Systematic Review and Meta-analysis. Clin Gastroenterol Hepatol 2019;17(4):616–629.e26.