

Supplementary Appendix

Effect of dapagliflozin on metabolic dysfunction-associated steatohepatitis: multicentre, double-blind, randomised, placebo-controlled trial

Contents		Pages
Supplemental Methods		
Inclusion and Exclusion Criteria		2
Procedure of Pathology Evaluation		3
Supplemental Figures and Tables		
Supplementary figure 1	Primary and Confirmatory Secondary End Points at Week 48 in Patients with F2-F3 Stage	4
Supplementary figure 2	Subgroup Analysis in Proportions of Patients Achieving MASH Improvement Without Worsening of Fibrosis at Week 48	5
Supplementary figure 3	Subgroup Analysis in Proportions of Patients Achieving MASH Resolution Without Worsening of Fibrosis at Week 48	6
Supplementary figure 4	Subgroup Analysis in Proportions of Patients Achieving Fibrosis Improvement Without Worsening of MASH at Week 48	7
Supplementary figure 5	Centre-specific Estimates of the Risk Ratio of the Primary End Point	8
Supplementary figure 6	Proportions of Patients with Worsened, Stable, or Improved Histological Features at Week 48	9
Supplementary figure 7	Proportions of Patients Achieving both MASH Resolution and Fibrosis Improvement at Week 48	10
Supplementary figure 8	Changes in Liver steatosis and Stiffness as Assessed by FibroScan from Baseline to Week 48	11
Supplementary figure 9	Changes in Liver Enzymes from Baseline to Week 48	12
Supplementary figure 10	Changes in Body Weight, and Waist Circumference, and Abdominal Fat Area from Baseline to Week 48	13
Supplementary figure 11	Changes in Glycaemic Parameters in Patients with Type 2 Diabetes from Baseline to Week 48	14
Supplementary figure 12	Changes in Glycaemic Parameters in Patients without Type 2 Diabetes from Baseline to Week 48	15
Supplementary figure 13	Changes in Lipid Profile from Baseline to Week 48	16
Supplementary table 1	Characteristics in Medication Use at Baseline, Week 24, and Week 48	17
Supplementary table 2	Dietary Intake and Physical Activity	19
Supplementary table 3	Sensitivity Analysis for Primary and Confirmatory Secondary End Points at Week 48 in Multiple Imputation Analysis and Per Protocol Analysis	20
Supplementary table 4	Sensitivity Analysis for Primary End Point with Different Definitions Used in Previous Studies.	21
Supplementary table 5	Sensitivity Analysis for Primary and Confirmatory Secondary End Points at Week 48 Adjusted for Sites	22
Supplementary table 6	Mediation Analysis on Weight Change for Primary and Confirmatory Secondary End Points at Week 48.	23
Supplementary table 7	Adverse Events	24

Supplemental Methods

Inclusion criteria

- Aged ≥ 18 years;
- MASH (defined as a Clinical Research Network non-alcoholic fatty liver disease activity score of ≥ 4 , with ≥ 1 point for steatosis, ballooning, and lobular inflammation) as determined by liver biopsies within 6 months;
- Patients with HbA1c $< 9.5\%$ if diagnosed with type 2 diabetes at screening or patient without type 2 diabetes.

Exclusion criteria

- Excessive alcohol consumption in the past six months (more than 20 g/day for women or 30 g/day for men);
- A history of alcoholic liver disease, chronic viral hepatitis, drug-induced hepatitis, autoimmune hepatitis, cirrhosis, and liver cancer;
- Obstructive biliary disease;
- Other medical condition that would affect metabolism (i.e., Cushing syndrome, known hyperthyroidism or hypothyroidism);
- Poor glucose control (defined as HbA1C $\geq 9.5\%$ within 3 months) if diagnosed with type 2 diabetes, or taking any antidiabetic medication that would affect metabolism or body weight (i.e., thiazolidinediones, glucagon-like peptide 1 receptor agonist (GLP-1RA), dipeptidyl peptidase 4 inhibitor (DPP-4i) or initially insulin treatment in the past 3 months);
- Taking any medication (i.e., cortisol, methotrexate) that would affect steatohepatitis for more than two weeks in the past year;
- Chronic kidney disease or severe impaired renal function (serum creatinine ≥ 2.0 mg/dl);
- Serum alanine aminotransferase greater than 300U/L;
- A history of type 1 diabetes;
- A history of bladder cancer;
- Serious medical disease with likely life expectancy less than 5 years;
- Women who are pregnant or plan to become pregnant;
- Patients who cannot be followed for 24 months (due to a health situation or migration);
- Participation in other clinical trial in the 30 days before randomization;
- Patients who are unwilling or unable to give informed consent.

Procedure of Pathology Evaluation

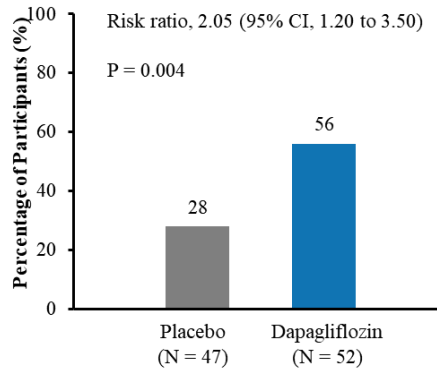
All liver biopsy specimens were obtained by ultrasound-guided percutaneous liver biopsy using a 16G needle in 6 centres. The average biopsy specimen length was ~20 mm. A histological laboratory received locally prepared unstained slices with 4 μ m thickness. These slices were then centrally stained with H&E and Masson's Trichrome, as well as staining for reticular fibre, Fe, Cu, CK19, HBsAg, and HBcAg.

DEAN trial used an adjudicator mode of three pathologists for pathology evaluation. To diminish the authority bias or observer bias from the interaction between pathologists, expert pathologists were invited to independently assess the pathology of the subjects from three top pathology departments in China. Data safety and monitoring board (DSMB) is responsible for reviewing the recruitment progress, database quality control, outcome assessment, and descriptive statistics of the study. An independent data coordinating and analysis centre (DCAC) in the Department of Statistics of the School of Public Health, Southern Medical University, was responsible for masking the slices, monitoring the progress of pathology assessments, and conveying any communications. All pathologists were unaware of the treatment assignments and patient characteristics. No additional training or discussion of the NASH-CRN criteria was done before the pathology assessments of three pathologists. The assessments were based on the clinical practice by him/herself.

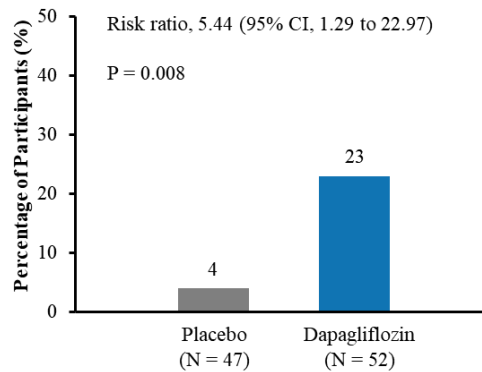
Slides were first read by pathologist A within two weeks of the screening visit or the week 48 biopsy, to determine whether it is a NASH, cirrhosis, or other liver disease for eligibility or AE report. After the study stopped recruiting, pathologist B began to centrally read slides in batches containing randomly mixed slides of baseline and week 48 biopsy from 10~20 patients. Both pathologists would read all the slides, judge the biopsy adequacy, do the differentiation diagnosis based on the histological findings, and scored steatosis, ballooning, lobular inflammation, and fibrosis according to the NASH-CRN criteria, without knowledge of the scores assessed by each other. Both baseline and follow-up biopsies were reread in batches by the same pathologists to minimize the variability of the individual reader over the whole study period after the trial was completed. Consistent scores were recorded as consensus values used in the primary analysis. In cases of discordant scores on any variable (steatosis, hepatocyte ballooning, lobular inflammation, and fibrosis), pathologist C would act as the adjudicator and make the final decision. If the adjudicator believed that a discussion between pathologists was necessary, a consensus call would be held.

Totally, the scores of pathologist A and B were consistent in 67% of the cases for steatosis, 63% for ballooning, 54% for inflammation, and 59% for fibrosis, with a weighted kappa of 0.55-0.66 for steatosis, 0.30-0.48 for ballooning, 0.39-0.47 for lobular inflammation, and 0.51-0.56 for fibrosis, for the baseline and week 48 biopsies, respectively. No consensus call was required for the adjudicator.

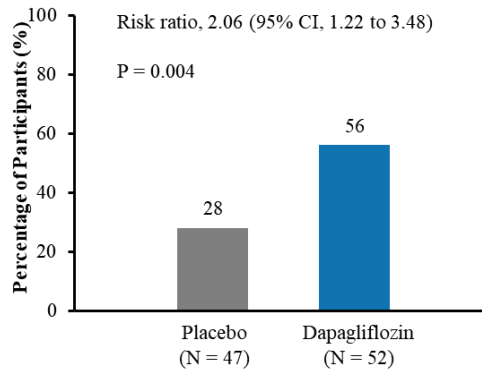
A MASH Improvement without Worsening of Fibrosis in Patients with F2-F3 Fibrosis Stage.



B MASH Resolution without Worsening of Fibrosis in Patients with F2-F3 Fibrosis Stage.



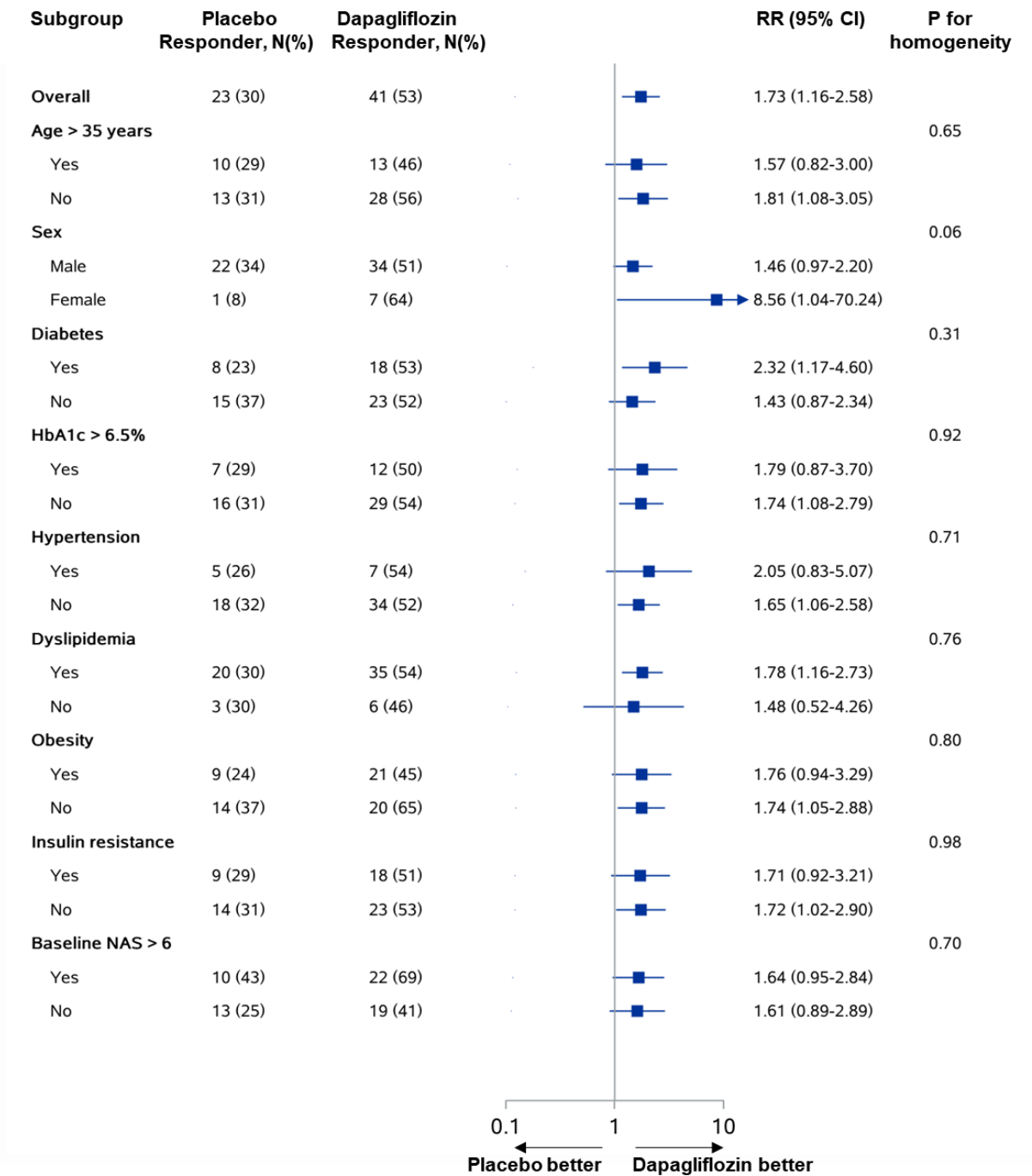
C Fibrosis Improvement without Worsening of MASH in Patients with F2-F3 Fibrosis Stage.



Supplementary figure 1. Primary and Confirmatory Secondary End Points at Week 48 in Patients with F2-F3 Fibrosis Stage.

MASH denotes non-alcoholic steatohepatitis; MASH improvement was defined as a decrease in NAS score of ≥ 2 points or a NAS score of ≤ 3 points after treatment) without worsening of liver fibrosis (defined as increase of ≥ 1 stage) at week 48; MASH resolution was defined as achievement of a hepatocellular ballooning score of 0, and lobular inflammation score of 0 or 1, without worsening of fibrosis at week 48; Fibrosis improvement was defined as reduction in fibrosis of ≥ 1 stage without worsening of MASH (defined as an increase in steatosis, ballooning, or inflammation of ≥ 1 stage) at week 48; Patients with missing biopsies were considered as non-responder.

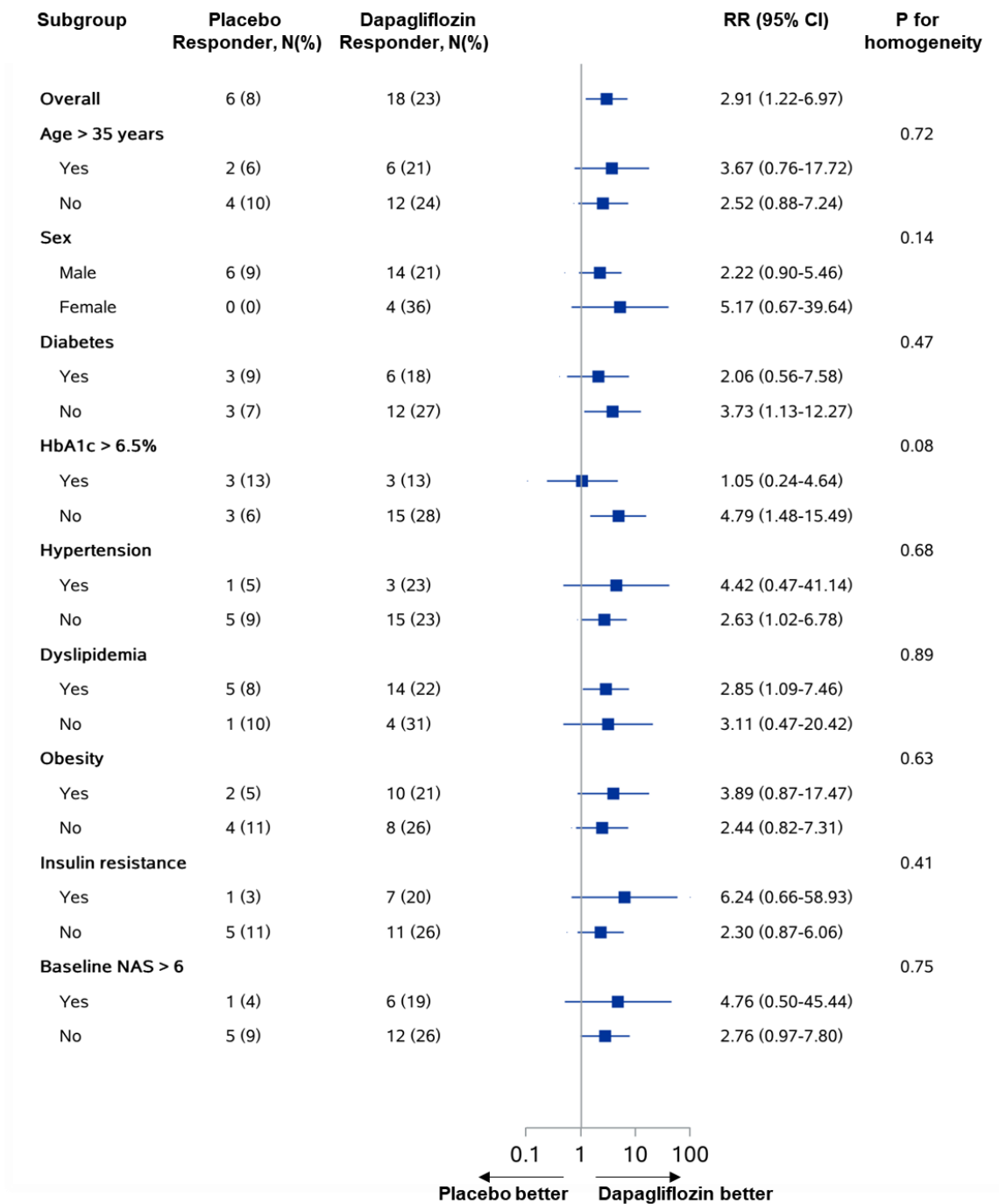
The Risk ratio (RR) and P values were calculated using the Cochran-Mantel-Haenszel (CMH) test, with stratified for baseline diabetes status.



Supplementary figure 2. Subgroup Analysis in Proportions of Patients Achieving MASH Improvement Without Worsening of Fibrosis at Week 48.

MASH improvement was defined as a decrease in NAS score of ≥ 2 points or a NAS score of ≤ 3 points after treatment) without worsening of liver fibrosis (defined as increase of ≥ 1 stage) at week 48; Obesity was defined as body-mass index ≥ 28 kg/m²; Insulin resistance was defined as homeostatic model assessment of insulin resistance (HOMA-IR) ≥ 4.01 ; NAS denotes non-alcoholic fatty liver disease activity score, on a scale of 0 to 8, with higher scores indicating more severe disease.

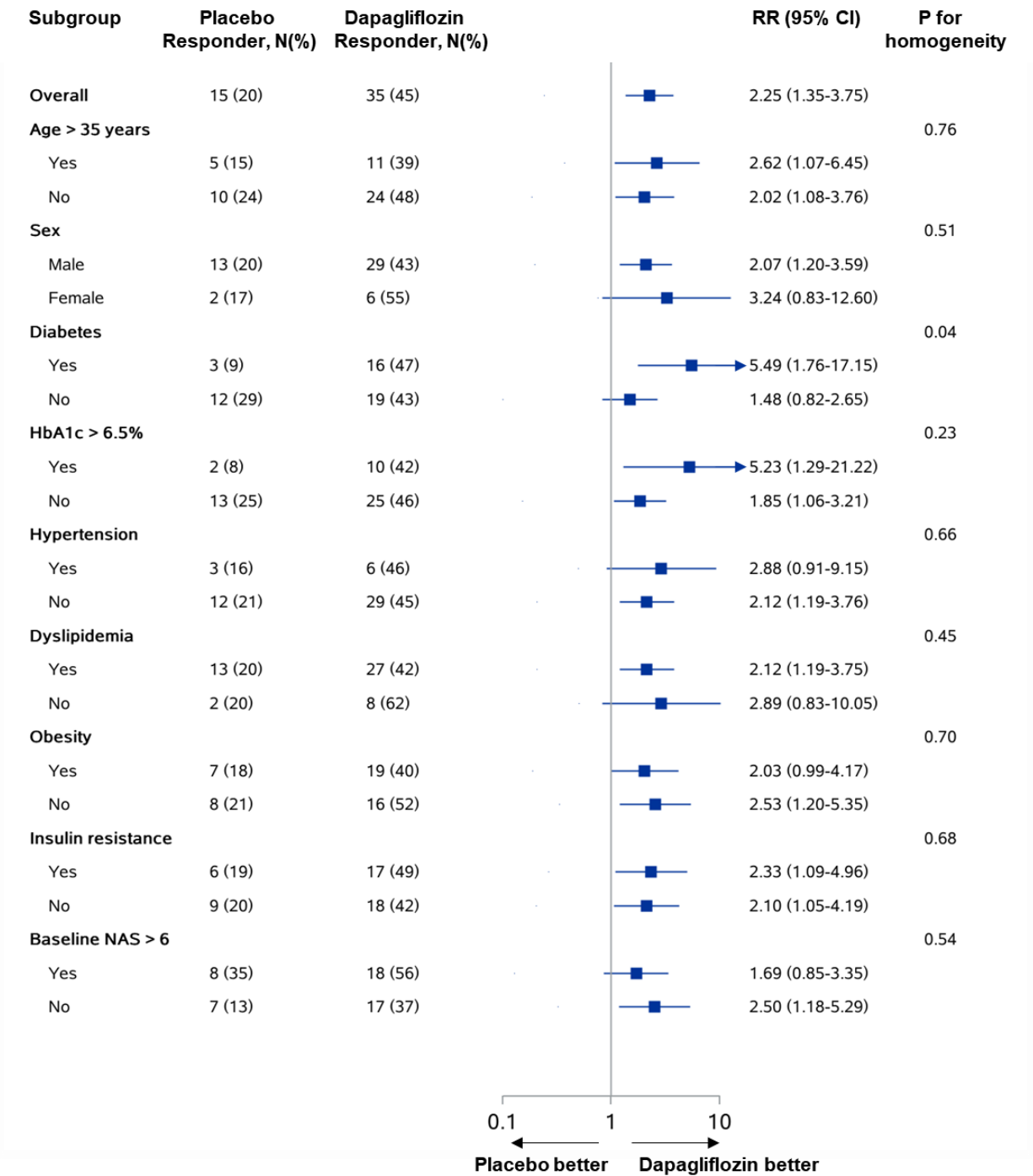
The risk ratios (RRs) were calculated using the Cochran-Mantel-Haenszel (CMH) test. Breslow-Day tests were conducted for analysing homogeneity between subgroups. The 95% confidence intervals (CIs) and P values were not adjusted for multiple comparisons and should not be used to infer definitive treatment effects.



Supplementary figure 3. Subgroup Analysis in Proportions of Patients Achieving MASH Resolution Without Worsening of Fibrosis at Week 48.

MASH resolution was defined as achievement of a hepatocellular ballooning score of 0, and lobular inflammation score of 0 or 1, without worsening of fibrosis at week 48; Obesity was defined as body-mass index ≥ 28 kg/m²; Insulin resistance was defined as homeostatic model assessment of insulin resistance (HOMA-IR) ≥ 4.01 ; NAS denotes non-alcoholic fatty liver disease activity score, on a scale of 0 to 8, with higher scores indicating more severe disease.

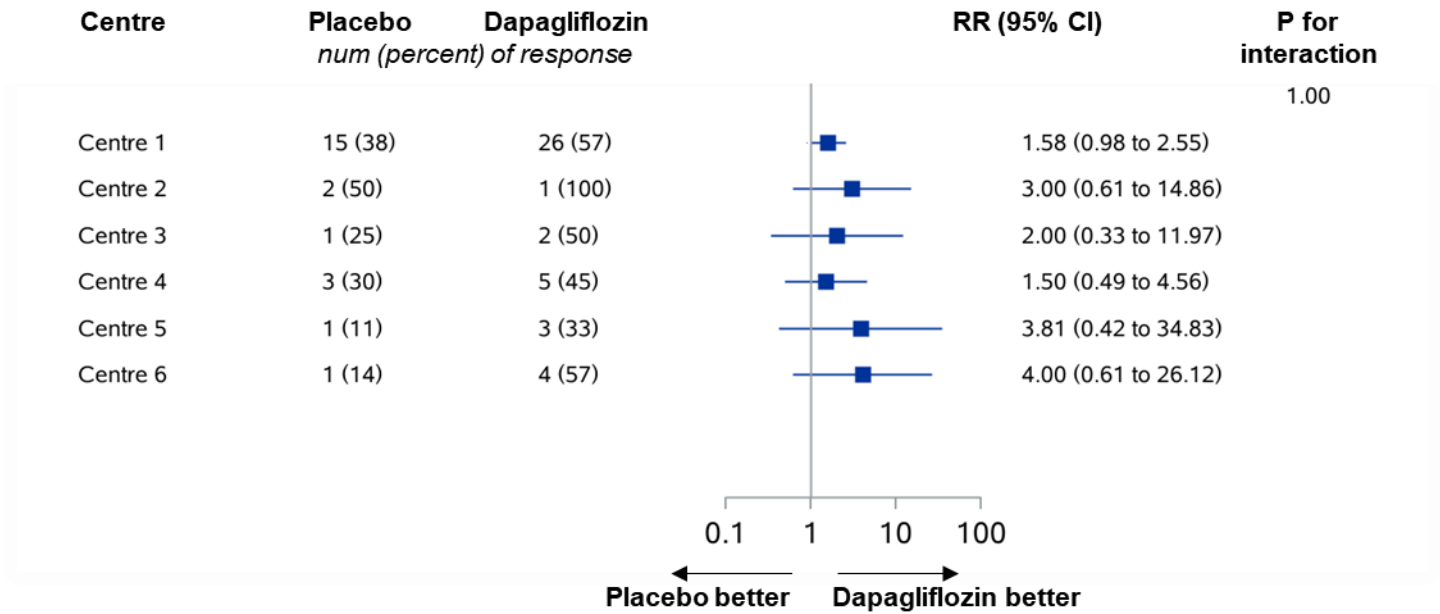
The risk ratios (RRs) were calculated using the Cochran-Mantel-Haenszel (CMH) test. Breslow-Day tests were conducted for analysing homogeneity between subgroups. The 95% confidence intervals (CIs) and P values were not adjusted for multiple comparisons and should not be used to infer definitive treatment effects.



Supplementary figure 4. Subgroup Analysis in Proportions of Patients Achieving Fibrosis Improvement Without Worsening of MASH at Week 48.

Fibrosis improvement was defined as reduction in fibrosis of ≥ 1 stage without worsening of MASH (defined as an increase in steatosis, ballooning, or inflammation of ≥ 1 stage) at week 48; Obesity was defined as body-mass index ≥ 28 kg/m²; Insulin resistance was defined as homeostatic model assessment of insulin resistance (HOMA-IR) ≥ 4.01 ; NAS denotes non-alcoholic fatty liver disease activity score, on a scale of 0 to 8, with higher scores indicating more severe disease.

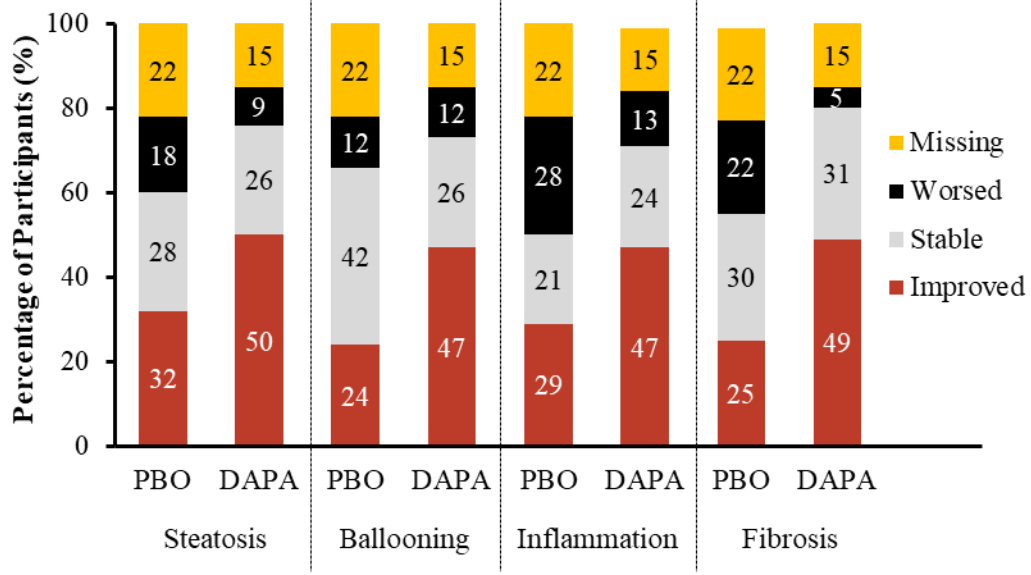
The risk ratios (RRs) were calculated using the Cochran-Mantel-Haenszel (CMH) test. Breslow-Day tests were conducted for analysing homogeneity between subgroups. The 95% confidence intervals (CIs) and P values were not adjusted for multiple comparisons and should not be used to infer definitive treatment effects.



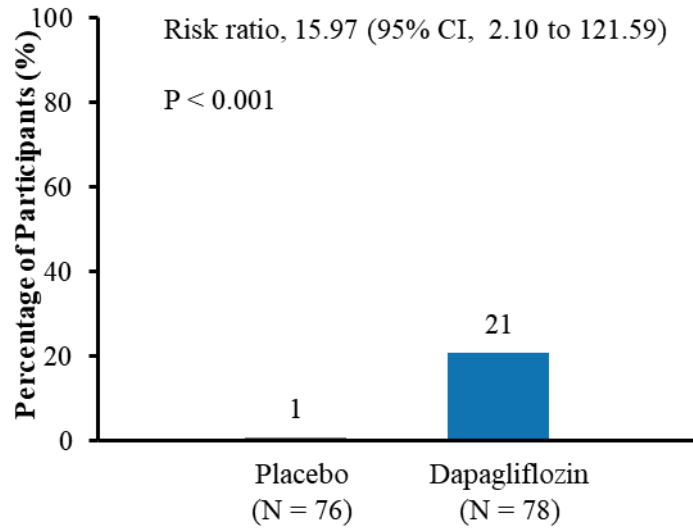
Supplementary figure 5. Centre-specific Estimates of the Risk Ratio of the Primary End Point.

The primary end point was defined as MASH improvement (a decrease in NAS score of ≥ 2 points or a NAS score of ≤ 3 points after treatment) without worsening of liver fibrosis (defined as increase of ≥ 1 stage) at week 48.

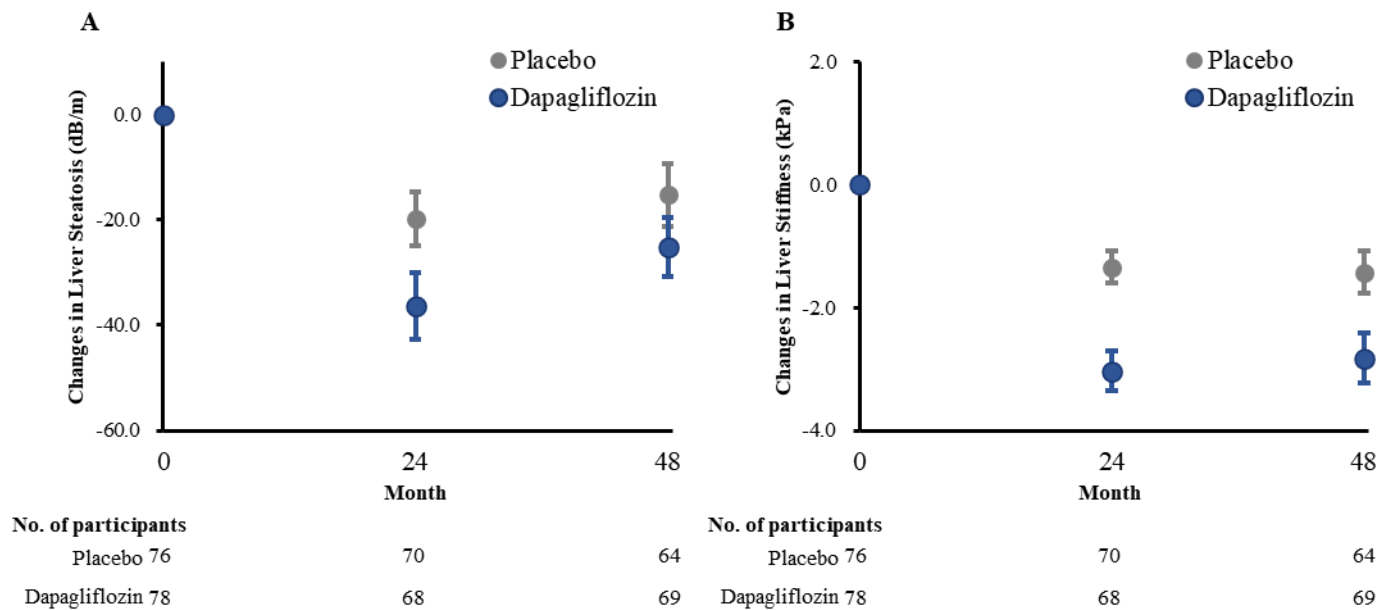
The risk ratios (RRs) were calculated using the Cochran-Mantel-Haenszel (CMH) test, controlling for baseline diabetes status. The 95% confidence intervals (CIs) and P values were not adjusted for multiple comparisons and should not be used to infer definitive treatment effects.



Supplementary figure 6. Proportions of Patients with Worsened, Stable, or Improved Histological Features at Week 48
PBO denotes placebo; DAPA denotes dapagliflozin.



Supplementary figure 7. Proportions of Patients Achieving both MASH Resolution and Fibrosis Improvement at Week 48. MASH resolution with fibrosis improvement is defined as a hepatocellular ballooning score of 0, and lobular inflammation score of 0 or 1, and reduction in fibrosis of ≥ 1 stage. The risk ratios (RR) and P value were calculated using the Cochran-Mantel-Haenszel (CMH) test, with stratified for baseline diabetes status. The 95% confidence interval (CI) and P value was not adjusted for multiple comparisons and should not be used to infer definitive treatment effects.

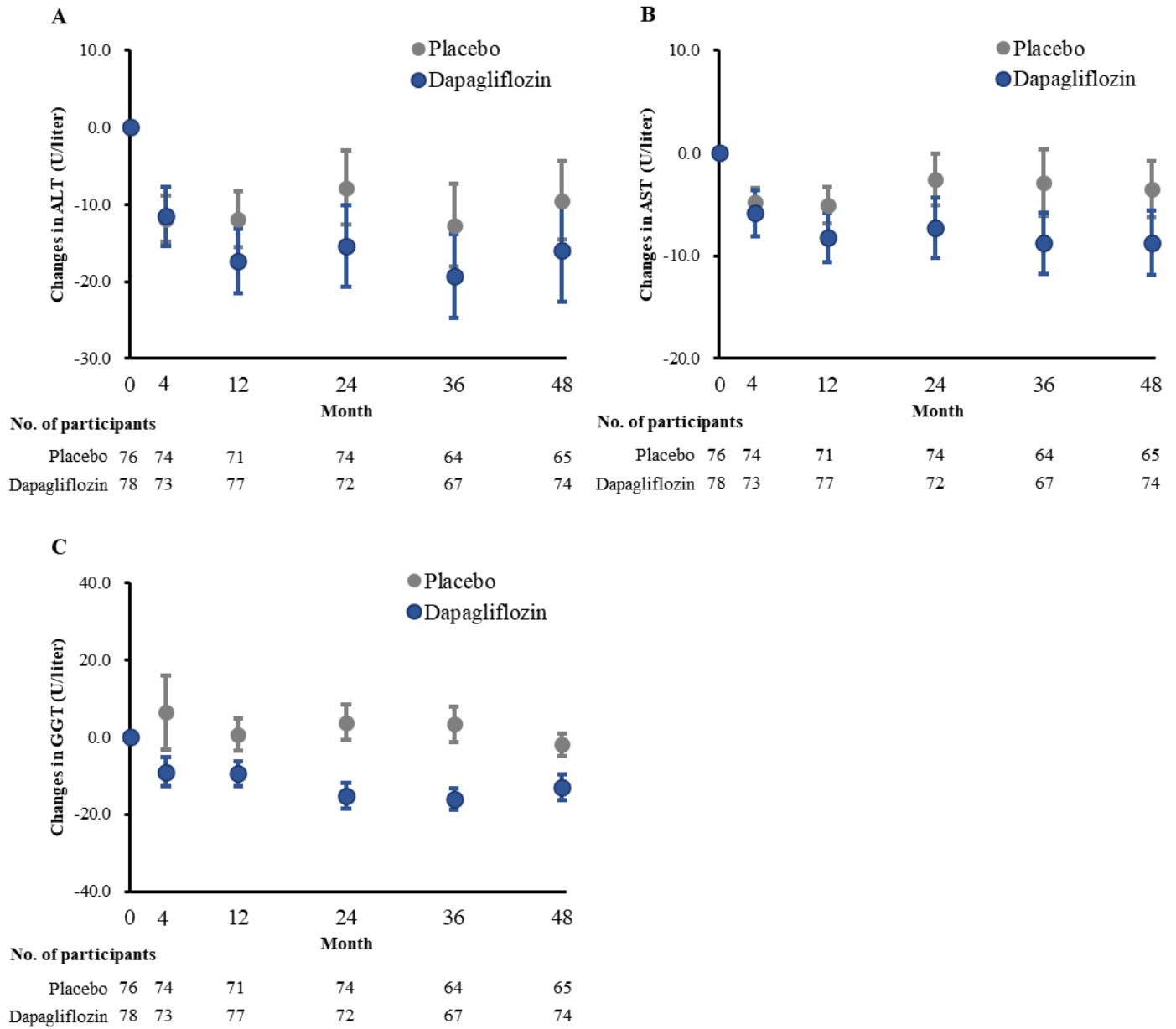


Supplementary figure 8. Changes in Liver steatosis and Stiffness as Assessed by FibroScan from Baseline to Week 48.

Panel A. Change in liver steatosis (as assessed by FibroScan) over 48 weeks;

Panel B. Change in liver stiffness (as assessed by FibroScan) over 48 weeks;

Data presented are observed data for the full analysis set from the in-trial period (the time from baseline to the date of last visit); Error bars represent standard error for the mean; Participant numbers shown denote those contributing to the mean.



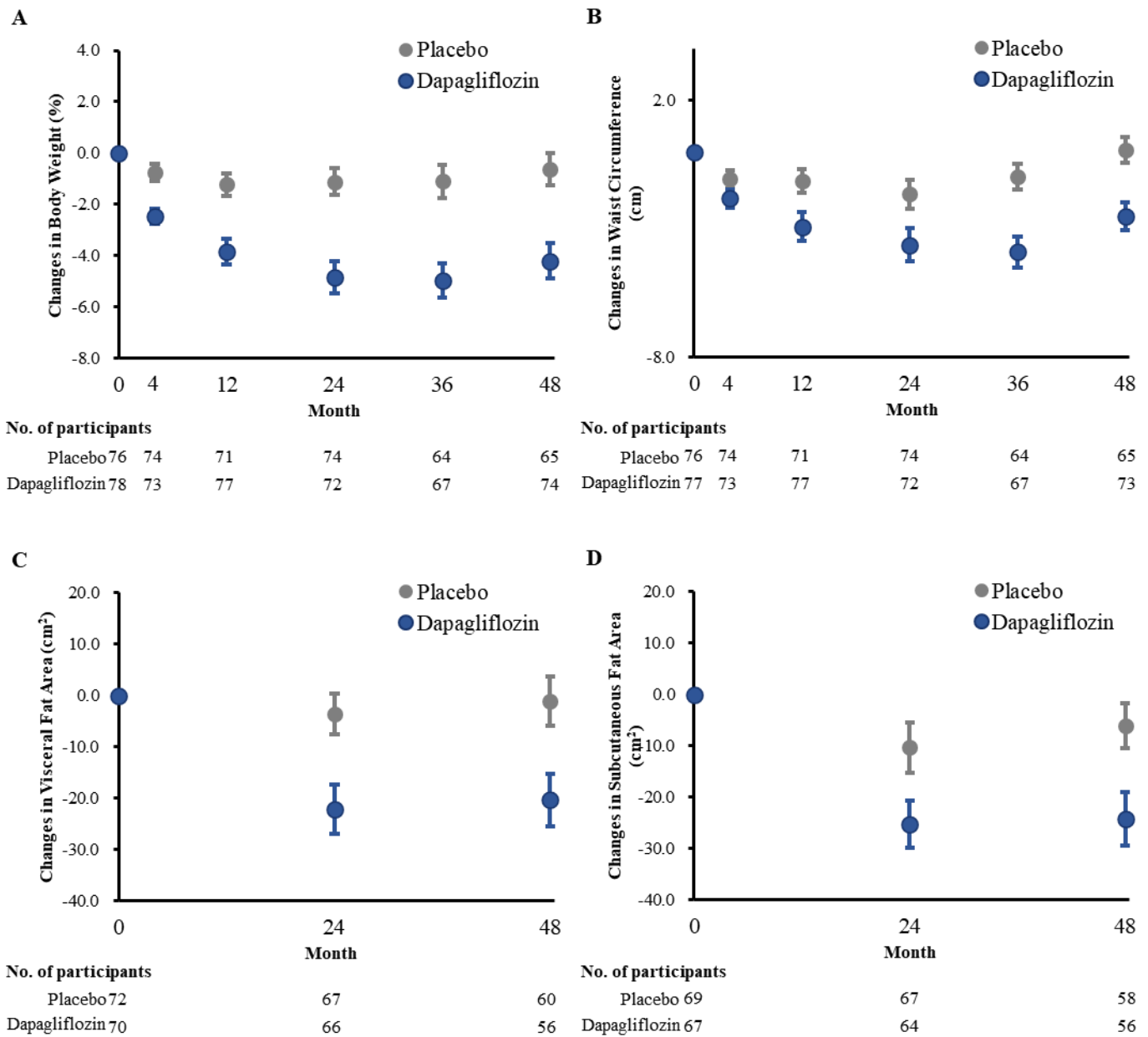
Supplementary figure 9. Changes in Liver Enzymes from Baseline to Week 48.

Panel A. Change in ALT (alanine aminotransferase) over 48 weeks;

Panel B. Change in AST (alanine aminotransferase) over 48 weeks;

Panel C. Change in GGT (γ -Glutamyltransferase) over 48 weeks;

Data presented are observed data for the full analysis set from the in-trial period (the time from baseline to the date of last visit); Error bars represent standard error for the mean; Participant numbers shown denote those contributing to the mean.



Supplementary figure 10. Changes in Body Weight, and Waist Circumference, and Abdominal Fat Area from Baseline to Week 48.

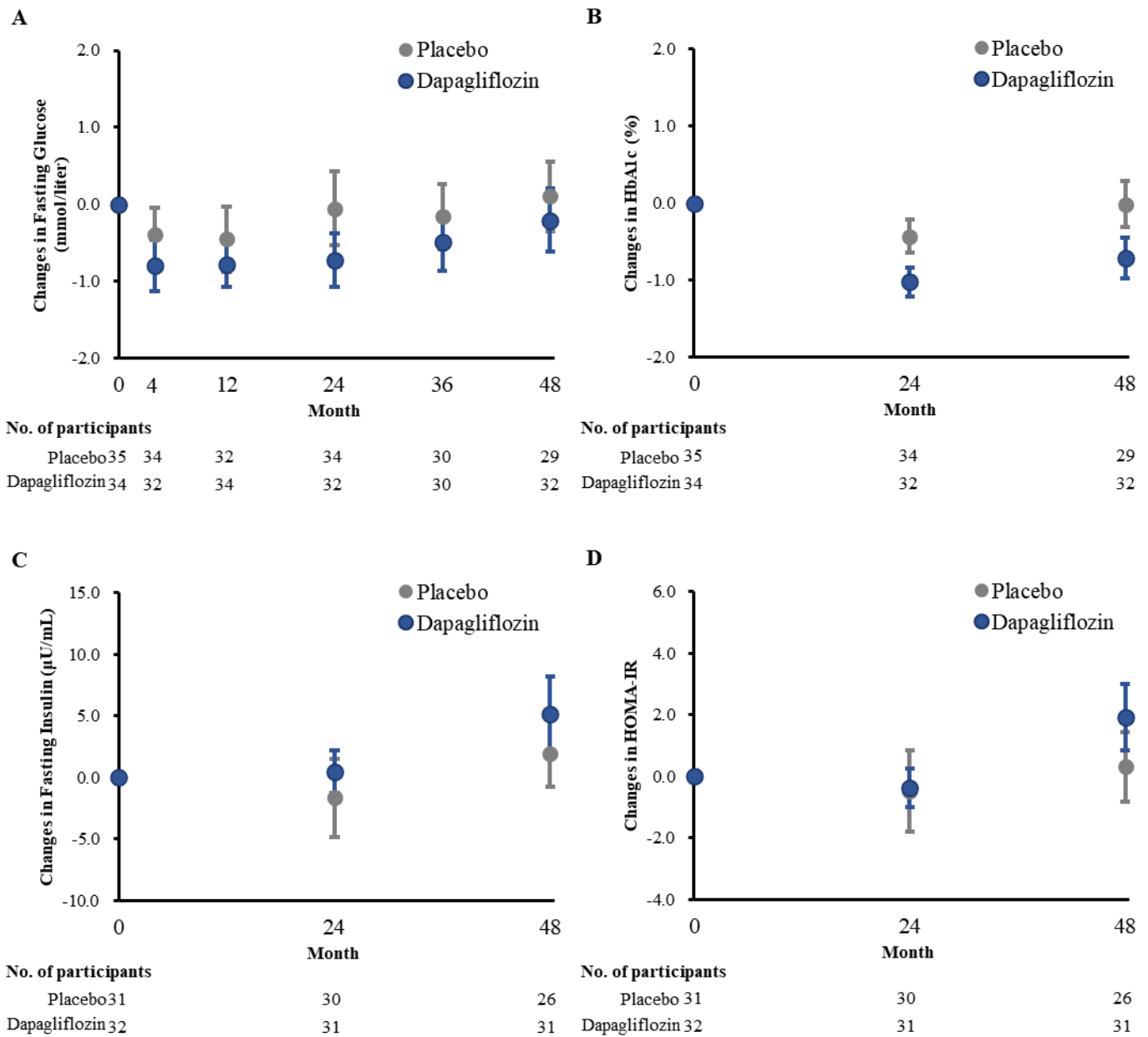
Panel A. Change in body weight percentage over 48 weeks;

Panel B. Change in waist circumference over 48 weeks;

Panel C. Change in visceral fat area over 48 weeks;

Panel D. Change in subcutaneous fat area over 48 weeks;

Data presented are observed data for the full analysis set from the in-trial period (the time from baseline to the date of last visit); Error bars represent standard error for the mean; Participant numbers shown denote those contributing to the mean.



Supplementary figure 11. Changes in Glycaemic Parameters in Patients with Type 2 Diabetes from Baseline to Week 48.

Panel A. Change in fasting glucose over 48 weeks in patients with type 2 diabetes;

Panel B. Change in HbA1c (glycated haemoglobin) over 48 weeks in patients with type 2 diabetes;

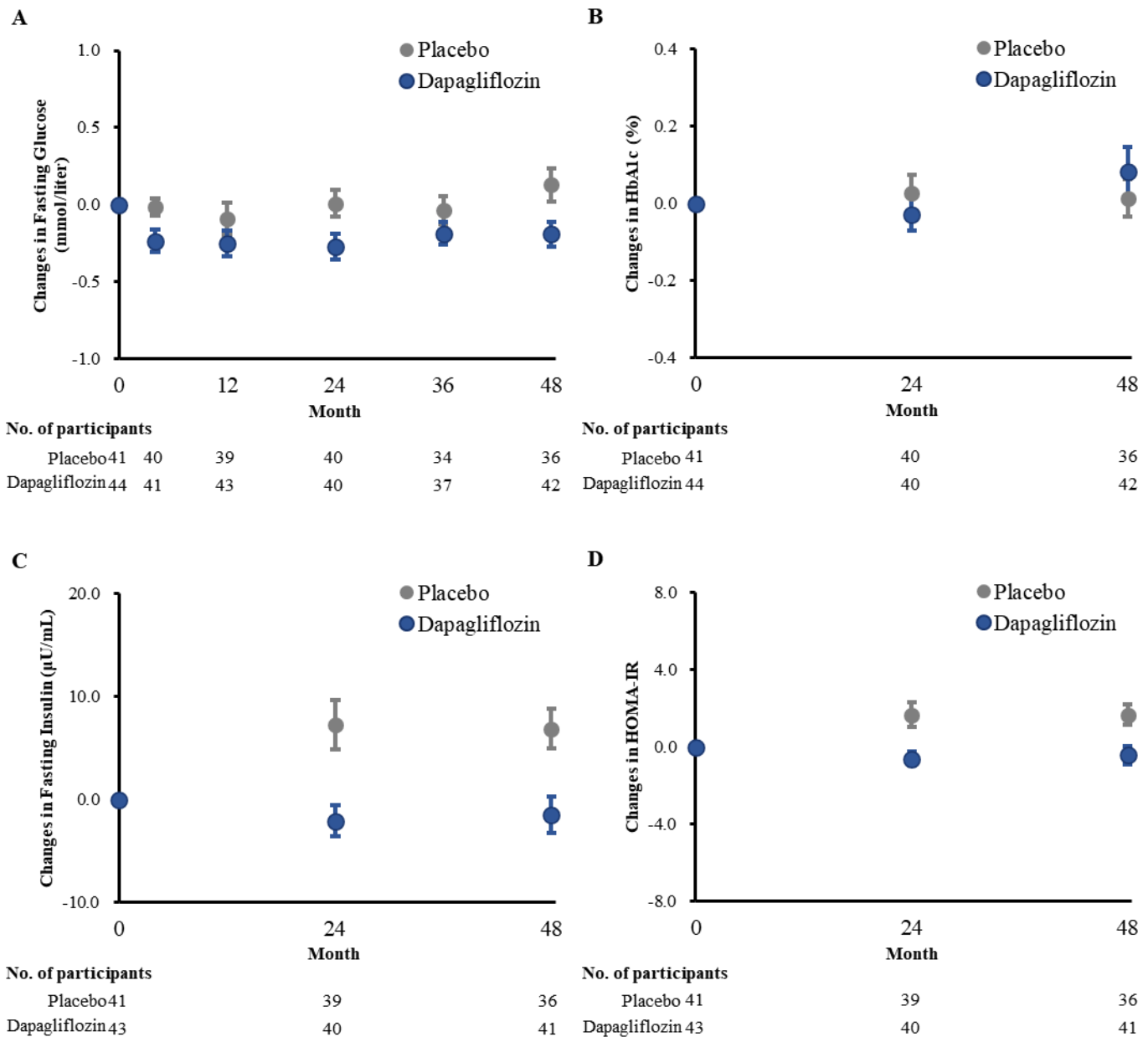
Panel C. Change in fasting insulin over 48 weeks in patients with type 2 diabetes;

Panel D. Change in HOMA-IR (homeostatic model assessment of insulin resistance) over 48 weeks in patients with type 2 diabetes.

HOMA-IR is calculated as the level of fasting glucose (measured in millimoles per litre) times the level of fasting insulin (measured in microunits per millilitre) divided by 22.5; Patients with insulin treatment were excluded from the analyses in Panel C and Panel D.

Data presented are observed data for patients with type 2 diabetes from the in-trial period (the time from baseline to the date of last visit);

Error bars represent standard error for the mean; Participant numbers shown denote those contributing to the mean.



Supplementary figure 12. Changes in Glycaemic Parameters in Patients without Type 2 Diabetes from Baseline to Week 48.

Panel A. Change in fasting glucose over 48 weeks in patients without type 2 diabetes;

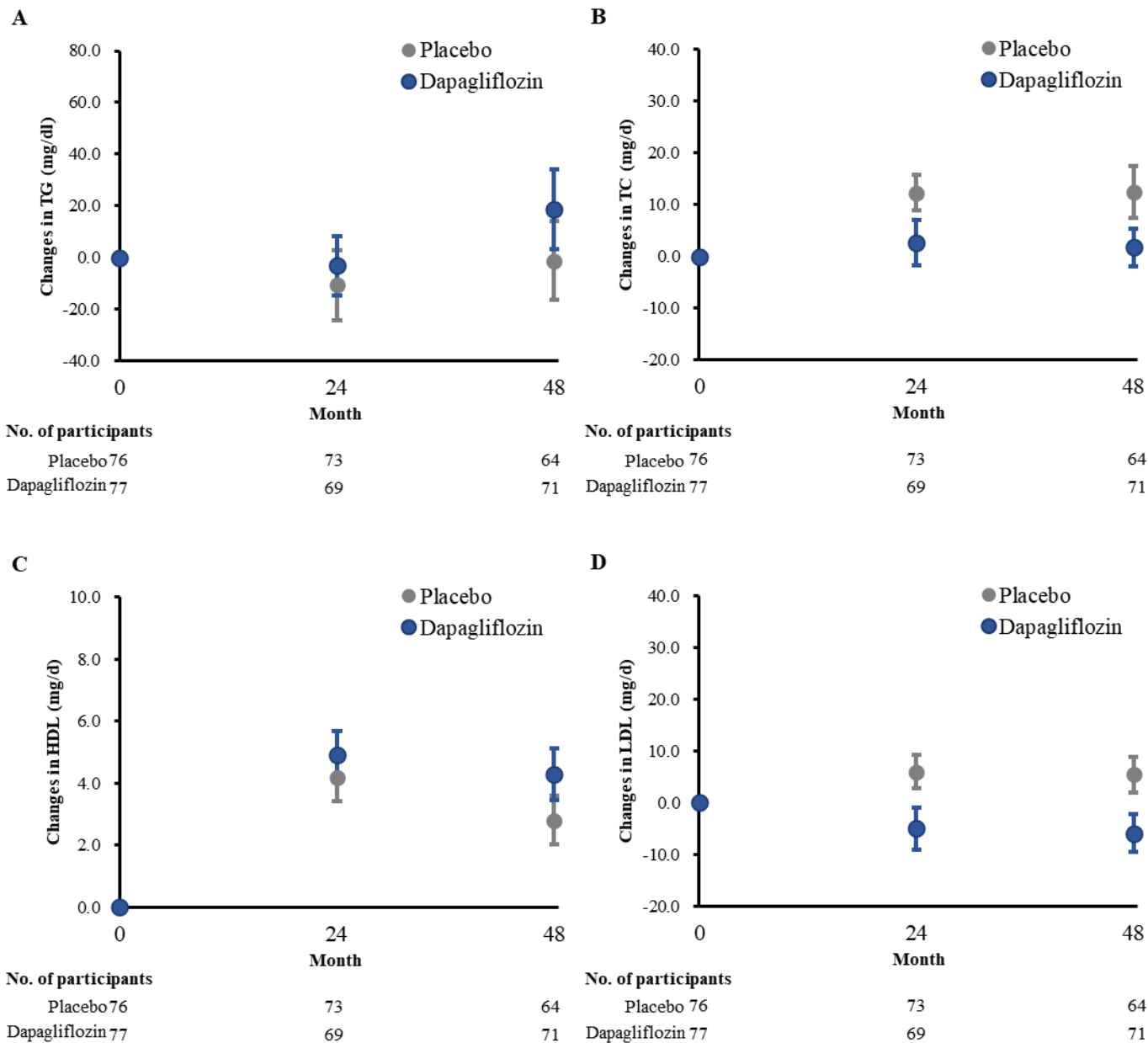
Panel B. Change in HbA1c (glycated haemoglobin) over 48 weeks in patients without type 2 diabetes;

Panel C. Change in fasting insulin over 48 weeks in patients without type 2 diabetes;

Panel D. Change in HOMA-IR (homeostatic model assessment of insulin resistance) over 48 weeks in patients without type 2 diabetes.

HOMA-IR is calculated as the level of fasting glucose (measured in millimoles per litre) times the level of fasting insulin (measured in microunits per millilitre) divided by 22.5.

Data presented are observed data for all the patients without type 2 diabetes from the in-trial period (the time from baseline to the date of last contact with trial site); Error bars represent standard error for the mean; Participant numbers shown denote those contributing to the mean.



Supplementary figure 13. Changes in Lipid Profile from Baseline to Week 48.

Panel A. Change in TG (triglyceride) over 48 weeks;

Panel B. Change in TC (total cholesterol) over 48 weeks;

Panel C. Change in HDL (high-density lipoprotein cholesterol) over 48 weeks;

Panel D. Change in LDL (low-density lipoprotein cholesterol) over 48 weeks;

Data presented are observed data for the full analysis set from the in-trial period (the time from baseline to the date of last visit); Error bars represent standard error for the mean; Participant numbers shown denote those contributing to the mean.

Supplementary table 1. Characteristics in Medication Use at Baseline, Week 24, and Week 48.

Characteristics	Placebo (n=76)	Dapagliflozin (n=78)	P value
Baseline			
Anti-diabetic agent			
Biguanide	29 (38)	24 (31)	0.40
Alpha-glucosidase inhibitor	14 (18)	10 (13)	0.38
Sulfonylurea	12 (16)	11 (14)	0.82
Insulin	4 (5)	2 (3)	0.44
Anti-hypertensive agent			
Angiotensin-converting enzyme inhibitor	1 (1)	1 (1)	1.00
Angiotensin II receptor blocker	3 (4)	2 (3)	0.68
Calcium channel blocker	4 (5)	1 (1)	0.21
Diuretics	2 (3)	2 (3)	1.00
Alpha-blocker	0 (0)	1 (1)	1.00
Beta-blocker	1 (1)	0 (0)	0.49
Anti-dyslipidaemia agent			
Statin	9 (12)	4 (5)	0.16
Fibrate	5 (7)	2 (3)	0.27
Week 24			
Anti-diabetic agent			
Biguanide	28 (37)	27 (36)	0.87
Alpha-glucosidase inhibitor	12 (16)	4 (5)	0.04
Sulfonylurea	8 (11)	9 (12)	1.00
Insulin	4 (5)	1 (1)	0.21
Anti-hypertensive agent			
Angiotensin-converting enzyme inhibitor	1 (1)	0 (0)	0.50
Angiotensin II receptor blocker	5 (7)	2 (3)	0.28
Calcium channel blocker	2 (3)	2 (3)	1.00
Diuretics	2 (3)	1 (1)	0.62
Alpha-blocker	0 (0)	0 (0)	0.54
Beta-blocker	1 (1)	0 (0)	0.50
Anti-dyslipidaemia agent			
Statin	4 (5)	6 (8)	0.75
Fibrate	4 (5)	0 (0)	0.06
Week 48			
Anti-diabetic agent			
Biguanide	25 (38)	28 (38)	1.00
Alpha-glucosidase inhibitor	10 (15)	4 (5)	0.09
Sulfonylurea	6 (9)	7 (9)	1.00
Insulin	3 (5)	1 (1)	0.34
Anti-hypertensive agent			
Angiotensin-converting enzyme inhibitor	1 (2)	0 (0)	0.47
Angiotensin II receptor blocker	3 (5)	2 (3)	0.66

Calcium channel blocker	1 (2)	2 (3)	1.00
Diuretics	2 (3)	1 (1)	0.60
Alpha-blocker	0 (0)	0 (0)	0.54
Beta-blocker	0 (0)	0 (0)	0.54
Anti-dyslipidaemia agent			
Statin	4 (6)	6 (8)	0.75
Fibrate	2 (3)	1 (1)	0.60

Data are presented as N (%).

P values are calculated with Fisher's exact test.

Supplementary table 2. Physical Activity and Dietary Intake.

Characteristics	Placebo (n=76)	Dapagliflozin (n=78)	P value
Physical activity (METs*hour per week)			
Baseline	7.7 (0-23.1)	9.9 (0-21.3)	0.18
Week 24	9.9 (0-22.2)	9.2 (0-20.0)	0.88
Week 48	8.0 (0-19.6)	6.6 (0-19.8)	0.63
Calorie intake (kcal/day)			
Baseline	2556.6 (903.6)	2464.3 (863.5)	0.52
Week 24	2214.7 (810.9) **	2178.7 (788.6) *	0.79
Week 48	2371.4 (879.3)	2214.2 (883.8)	0.30
Carbohydrate intake (%)			
Baseline	54.0 (9.8)	54.6 (12.0)	0.75
Week 24	53.9 (10.9)	55.6 (10.3)	0.33
Week 48	55.2 (10.8)	55.3 (13.3)	0.97
Protein intake (%)			
Baseline	15.4 (3.4)	15.7 (4.6)	0.65
Week 24	15.4 (3.6)	15.4 (3.5)	0.96
Week 48	15.4 (3.8)	15.4 (4.4)	0.92
Fat intake (%)			
Baseline	31.2 (8.5)	30.3 (9.2)	0.53
Week 24	31.3 (9.5)	29.6 (8.4)	0.25
Week 48	30.0 (8.7)	29.7 (9.9)	0.86

Data are mean (SD) or median (IQR); MET denotes metabolic equivalents. P values are estimated from general linear model.

* P < 0.05 vs. baseline.

** P < 0.01 vs. baseline.

Supplementary table 3. Sensitivity Analysis for Primary and Confirmatory Secondary End Points at Week 48 in Multiple Imputation Analysis and Per Protocol Analysis.

End points	Placebo (n=76)	Dapagliflozin (n=78)	RR (95% CI) †	P value
Multiple Imputation Analysis*				
Primary end point				
MASH improvement without worsening of fibrosis	31 (41)	47 (60)	1.50 (1.05 to 2.14)	0.03
Confirmatory secondary end points				
MASH resolution without worsening of fibrosis	7 (9)	21 (27)	2.97 (1.33 to 6.61)	0.008
Fibrosis improvement without worsening of MASH	19 (25)	40 (51)	2.10 (1.29 to 3.41)	0.003
Per Protocol Analysis				
Primary end point				
MASH improvement without worsening of fibrosis	23 (39)	41 (62)	1.59 (1.10 to 2.30)	0.01
Confirmatory secondary end points				
MASH resolution without worsening of fibrosis	6 (10)	18 (27)	2.67 (1.13 to 6.31)	0.02
Fibrosis improvement without worsening of MASH	15 (25)	35 (53)	2.07 (1.27 to 3.38)	0.002

* The multiple -imputation is based on the missing at random assumption that a participant with a missing end-of-study biopsy would have similar histological response as participant with an end-of-study biopsy and comparable baseline characteristics. Variables used for the imputation include treatment group, age, sex, baseline diabetes status, BMI, liver fat and stiffness assessed by FibroScan®, and histological characteristics (NAS, steatosis, ballooning, inflammation, and fibrosis score). One hundred imputation were generated and the results were pooled using Rubin's rule.

† The risk ratios (RRs) and P values were calculated using the Cochran-Mantel-Haenszel (CMH) test, controlling for baseline diabetes status.

Supplementary table 4. Sensitivity Analysis for Primary End Point with Different Definitions Used in Previous Studies.

Definitions of MASH improvement *	Placebo (n=76)	Dapagliflozin (n=78)	RR (95% CI) †	P value
1. decrease of at least 2 points in NAS, or NAS of less than or equal to 3 points, without worsening of fibrosis ‡	23 (30)	41 (53)	1.73 (1.16 to 2.58)	0.006
2. decrease of at least 2 points in NAS or NAS of less than or equal to 3 points, and decrease of at least 1 point in ballooning, without worsening of fibrosis	17 (22)	34 (44)	1.94 (1.19 to 3.16)	0.006
3. decrease of at least 2 points in NAS or NAS of less than or equal to 3 points, and decrease of at least 1 point in ballooning or lobular inflammation, without worsening of fibrosis	21 (28)	38 (49)	1.76 (1.15 to 2.70)	0.008
4. decrease of at least 2 points in NAS, and decrease of at least 1 point in ballooning, without worsening of fibrosis	16 (21)	33 (42)	2.01 (1.21 to 3.33)	0.005
5. decrease of at least 2 points in NAS, without worsening of fibrosis	22 (29)	39 (50)	1.72 (1.14 to 2.61)	0.008
6. decrease of at least 2 points in NAS, and decrease in at least 2 components, without worsening of fibrosis	20 (26)	35 (45)	1.70 (1.09 to 2.67)	0.02

Data are shown as N (%).

* Shown are the definitions of MASH improvement used in previous studies. Patients with missing biopsies were considered as non-responder.

† The risk ratios (RRs) and P values were calculated using the Cochran-Mantel-Haenszel (CMH) test, with stratified for baseline diabetes status.

‡ Used in the current DEAN trial.

Supplementary table 5. Sensitivity Analysis for Primary and Confirmatory Secondary End Points at Week 48 Adjusted for Sites.

End points	OR (95% CI)	P value	Adjusted OR (95% CI)	Adjusted P value
Primary end point				
MASH improvement without worsening of fibrosis	2.52 (1.30 to 4.87)	0.006	2.70 (1.37 to 5.31)	0.004
Confirmatory secondary end points				
MASH resolution without worsening of fibrosis	3.43 (1.29 to 9.16)	0.01	3.35 (1.24 to 9.05)	0.02
Fibrosis improvement without worsening of MASH	3.26 (1.59 to 6.68)	0.001	4.00 (1.84 to 8.68)	<0.001

ORs and 95% CIs were estimated using logistic models, with presence of diabetes as stratification factor.

Adjusted ORs and 95% CIs were estimated using logistic models, with presence of diabetes as stratification factor and sites as covariate.

Supplementary table 6. Mediation Analysis of Weight Change Percentage at Week 48.

End points	OR (95% CI) *	Adjusted OR (95% CI)†	Adjusted P value
Primary end point			
MASH improvement without worsening of fibrosis	2.52 (1.30 to 4.86)	1.56 (0.74 to 3.31)	0.25
Confirmatory secondary end points			
MASH resolution without worsening of fibrosis	3.43 (1.29 to 9.16)	1.84 (0.62 to 5.44)	0.27
Fibrosis improvement without worsening of MASH	3.26 (1.59 to 6.68)	2.18 (1.01 to 4.72)	0.048

*The odds ratios (ORs), 95% confidence interval (CI), and P values were calculated using logistic regression models, stratified for baseline diabetes status.

†The adjusted ORs, 95% CI, and p values were calculated using logistic regression models, stratified for baseline diabetes status, with baseline body weight and weight change percentage as covariates.

Supplementary table 7. Adverse Events.

Adverse events	Placebo (n=76)	Dapagliflozin (n=78)	P value
Any adverse events	49 (64)	44 (56)	0.33
Serious adverse event ^a	1 (1)	0 (0)	0.49
Adverse event leading to trial discontinuation	2 (3)	1 (1)	0.62
Adverse event from any system organ class ^b			
COVID19	7 (9)	8 (10)	1.00
Insomnia	3 (4)	5 (6)	0.72
Gout	2 (3)	5 (6)	0.44
Nasopharyngitis	5 (7)	2 (3)	0.27
Hypoglycaemia	2 (3)	4 (5)	0.68
Upper abdominal pain	2 (3)	4 (5)	0.68
Urinary tract infection	2 (3)	2 (3)	1.00
Gastrointestinal polyp	2 (3)	2 (3)	1.00
Dizziness	3 (4)	0 (0)	0.12
Fatigue	3 (4)	0 (0)	0.12
Gastroenteritis	2 (3)	1 (1)	0.62
Genital pruritus	2 (3)	1 (1)	0.62
Headache	2 (3)	1 (1)	0.62
Rash	2 (3)	1 (1)	0.62
Urinary tract stone	2 (3)	1 (1)	0.62
Bone fracture	0 (0)	2 (3)	0.50
Alopecia	1 (1)	1 (1)	1.00
Circumcision	1 (1)	1 (1)	1.00
Dermatitis	1 (1)	1 (1)	1.00
Face oedema	1 (1)	1 (1)	1.00

Data are presented as N (%).

^a One serious adverse event in diabetic ketoacidosis was reported.

^b Shown are the adverse events with at least two cases.