Supplemental Online Content

Rubino D, Abrahamsson N, Davies M, et al; STEP 4 Investigators. Effect of continued weekly subcutaneous semaglutide vs placebo on weight loss maintenance in adults with overweight or obesity: the STEP 4 randomized clinical trial. *JAMA*. Published online March 23, 2021. doi:10.1001/jama.2021.3224

This supplemental material has been provided by the authors to give readers additional information about their work.

eAppendix 1. Participant Enrollment and Exclusions by Study Site

Site	Enrolled into run-in, No.	Withdrew prior to randomization, No.
100	35	4
101	25	2
200	32	0
202	10	0
250	6	0
251	5	0
252	3	0
253	7	1
254	7	0
255	7	0
300	10	0
301	10	0
302	10	0
303	10	0
304	10	1
305	10	1
306	10	0
352	12	2
353	12	0
354	13	1
355	12	2
401	16	0
402	12	1
403	11	1
404	7	0
406	15	0
407	8	0
451	12	1
452	12	1
453	12	1
454	13	1
455	6	1
456	5	0
501	16	2
502	16	0
503	11	1

Site	Enrolled into run-in, No.	Withdrew prior to randomization, No.
504	8	1
505	5	0
506	14	1
552	15	1
553	13	1
554	13	0
555	13	2
556	12	0
601	15	2
602	14	2
603	12	4
604	13	3
605	15	1
606	16	3
607	14	4
608	12	2
609	15	4
610	14	2
611	14	2
612	13	3
613	15	0
614	12	3
615	14	1
616	13	3
617	12	4
618	8	1
619	16	5
620	14	3
621	7	2
622	13	2
623	11	5
624	16	2
625	17	4
627	14	0
628	15	1
629	17	1

eAppendix 2. Eligibility Criteria

Inclusion and Exclusion Criteria

Inclusion Criteria

Participants were eligible to be included in the trial only if all of the following criteria applied:

- Informed consent obtained before any trial-related activities. Trial-related activities are any procedures that were carried out as part of the trial, including activities to determine suitability for the trial
- Male or female, aged ≥18 years at the time of signing informed consent
- Body mass index ≥30 kg/m^2, or ≥27 kg/m^2 with the presence of ≥1 of the following weight-related comorbidities (treated or untreated): hypertension, dyslipidemia, obstructive sleep apnea, or cardiovascular disease
- History of ≥1 self-reported unsuccessful dietary effort to lose body weight

Exclusion Criteria

Participants were excluded from the trial if any of the following criteria applied:

Glycemia-Related:

- Glycated hemoglobin ≥6.5% (48 mmol/mol), as measured by the central laboratory at screening
- History of type 1 or type 2 diabetes
- Treatment with glucose-lowering agent(s) within 90 days before screening

Obesity-Related:

- A self-reported change in body weight >11 lbs (5 kg) within 90 days before screening irrespective of medical records
- Treatment with any medication for the indication of obesity within the past 90 days before screening
- Previous or planned (during the trial period) obesity treatment with surgery or a weight loss device. However, the following were allowed: (1) liposuction and/or abdominoplasty, if performed >1 year before screening; (2) lap banding, if the band had been removed >1 year before screening; (3) intragastric balloon, if the balloon had been removed >1 year before screening; or (4) duodenal-jejunal bypass sleeve, if the sleeve had been removed >1 year before screening
- Uncontrolled thyroid disease, defined as thyroid stimulating hormone >6.0 mIU/L or <0.4 mIU/L as measured by the central laboratory at screening

Mental Health:

- History of major depressive disorder within 2 years before screening
- Diagnosis of other severe psychiatric disorder (eg, schizophrenia, bipolar disorder)
- A Patient Health Questionnaire-9 score of ≥15 at screening
- Lifetime history of a suicidal attempt
- Suicidal behavior within 30 days before screening

- Suicidal ideation corresponding to type 4 or 5 on the Columbia-Suicide Severity Rating Scale within the past 30 days before screening

General Safety:

- Presence of acute pancreatitis within the past 180 days prior to the day of screening.
- History or presence of chronic pancreatitis
- Calcitonin ≥100 ng/L as measured by the central laboratory at screening
- Personal or first-degree relative(s) history of multiple endocrine neoplasia type 2 or medullary thyroid carcinoma
- Renal impairment measured as an estimated glomerular filtration rate <15 mL/min/1.73 m^2, as defined by Kidney Disease: Improving Global Outcomes 2012 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease,[1] by the central laboratory at screening
- History of malignant neoplasms within the past 5 years prior to screening. Basal and squamous cell skin cancer and any carcinoma in-situ were allowed
- Any of the following: myocardial infarction, stroke, hospitalization for unstable angina, or transient ischemic attack within the past 60 days prior to screening
- Participant classified as being in New York Heart Association Class IV
- Surgery scheduled for the duration of the trial, except for minor surgical procedures, in the opinion of the investigator
- Known or suspected abuse of alcohol or recreational drugs
- Known or suspected hypersensitivity to trial product(s) or related products
- Previous participation in this trial; participation was defined as signed informed consent
- Participation in another clinical trial within 90 days before screening
- Other participant(s) from the same household participating in any semaglutide trial
- Female who was pregnant, breast-feeding or intended to become pregnant or was of child-bearing potential and not using a highly effective contraceptive method
- Any disorder, unwillingness or inability, not covered by any of the other exclusion criteria, which in the investigator's opinion, might jeopardize the participant's safety or compliance with the protocol

Randomization Criteria

- Attended the randomization visit (Week 20) ***and*** had escalated to target dose after 16 weeks ± 3 days since Week 0 ***and*** were at target dose at the randomization visit (Week 20)

eAppendix 3. Patient-Reported Outcome Assessments

Short Form36v2® Health Survey, Acute Version (SF-36)

The SF-36 is a generic patient-reported outcome (PRO) instrument that measures health-related quality of life and general health status across disease conditions. It consists of 36 questions (items) across 8 domains (i.e., physical functioning, role limitations due to physical health problems, body pain, general health, vitality, social functioning, role limitations due to emotional problems, and mental health). The SF-36 also provides two aggregated scores, the physical component summary (PCS) score and mental component summary (MCS) score, created by aggregating the eight domains according to the scoring algorithm.[2] SF-36 scores are norm-based scores (also referred to as T-score metrics), i.e., transformed to a scale where the 2009 US general population has a mean of 50 and an SD of 10. Scores are calculated using PRO-CoRE version 1.5, a scoring software provided by Optum. The range of lowest to highest scores for the physical functioning domain is 19.03 to 57.60, for the physical component summary it is 6.11 to 79.67, and for the mental component summary it is –3.83 to 78.75. An increase in score represents an improvement in health status.

eAppendix 4. Supportive Secondary Endpoints

Efficacy Endpoints

From Randomization (Week 20) to Week 68

- Change in:
  - Body weight (kg) and body mass index (kg/m^2)
  - Glycated hemoglobin (%, mmol/mol), fasting plasma glucose (mg/dL), and fasting serum insulin (mIU/L)
  - Diastolic blood pressure (mmHg)
  - Lipids (mg/dL): total cholesterol, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, very low-density lipoprotein cholesterol, free fatty acids, and triglycerides
  - SF-36 scores: role-physical, bodily pain, general health, vitality, social functioning, role-emotional, mental health, PCS, MCS
- Subjects who after 68 weeks achieved (yes/no):
  - Responder definition value for SF-36 physical functioning score
- Subjects who after 68 weeks gained weight (yes/no)

From Week 0 to Week 68

- Percentage change in body weight
- Participants who achieved (yes/no) a body weight reduction <0%, ≥5%, ≥10%, ≥15%, or ≥20%

Safety-Related Endpoints

From Week 0 to Randomization (Week 20)

- Number of treatment-emergent adverse events (AEs)
- Number of serious AEs
- Change in pulse (bpm), amylase (U/L), lipase (U/L), and calcitonin (ng/L)

From Randomization (Week 20) to Week 75

- Numbers of treatment-emergent AEs and serious AEs

From Randomization (Week 20) to Week 68

- Change in pulse (bpm), amylase (U/L), lipase (U/L), and calcitonin (ng/L)

Exploratory Endpoints

- Change from randomization (week 20) to week 68 in:
  - Glycemic category (normoglycemia, prediabetes, type 2 diabetes) based on American Diabetes Association 2017 classification[3]
  - Antihypertensive medication (decrease, no change, increase)
  - Lipid-lowering medication (decrease, no change, increase)
  - The Stanford Presenteeism Scale, total score
  - Weight-Related Sign and Symptom Measure total score
- Participants who from randomization (week 20) to week 68 permanently discontinued trial product (yes/no)
- Time to permanent discontinuation of trial product (weeks) from week 0 to randomization (week 20)
- Participants who from randomization (week 20) to week 68 permanently discontinued randomized trial product (yes/no)
- Time to permanent discontinuation of randomized trial product (weeks) from randomization (week 20) to week 68

eTable 1. Analysis and Imputation Methods to Address the Treatment Policy and Trial Product Estimands for the Primary and Confirmatory Secondary Endpoints in the Statistical Testing Hierarchy

Objective	Endpoint	Test order	Endpoint type	Estimand	Analysis set	Statistical model	Imputation approach
Primary endpoint							
From randomization (week 20) to week 68:							
Primary	% weight change	1	Continuous	Treatment policy[a]	FAS	ANCOVA	RD-MI
				Trial product[b]	FAS	MMRM	–
Confirmatory secondary endpoints							
From randomization (week 20) to week 68:							
Primary	Waist circumference change (cm)	2	Continuous	Treatment policy[a]	FAS	ANCOVA	RD-MI
				Trial product[b]	FAS	MMRM	–
Secondary	SBP change (mmHg)	3	Continuous	Treatment policy[a]	FAS	ANCOVA	RD-MI
				Trial product[b]	FAS	MMRM	–
Secondary	SF-36 physical functioning score change	4	Continuous	Treatment policy[a]	FAS	ANCOVA	RD-MI
				Trial product[b]	FAS	MMRM	–

Abbreviations: ANCOVA, analysis of covariance; FAS, full analysis set; MMRM, mixed model for repeated measurements; RD-MI, multiple imputation using retrieved participants; SBP, systolic blood pressure; SF-36, Short Form36v2® Health Survey, Acute Version.

Test order refers to the order of the endpoint in the statistical test hierarchy.

[a]Designated as the primary estimand.

[b]Also known as the hypothetical estimand and designated as the secondary estimand.

eTable 2. Clinical Characteristics at Randomization (Week 20) for the Total Population (Full Analysis Set)[a]

	Total (N = 803)
Body weight, kg, mean (SD)	96.1 (22.6)
BMI, kg/m^2	
Mean (SD)	34.4 (7.0)
<25, No. (%)	16 (2.0)
≥25–<30, No. (%)	222 (27.6)
≥30–<35, No. (%)	263 (32.8)
≥35–<40, No. (%)	168 (20.9)
≥40, No. (%)	134 (16.7)
Waist circumference, cm, mean (SD)	105.3 (16.2)
SBP, mmHg, mean (SD)	121 (13)
DBP, mmHg, mean (SD)	78 (9)
Hemoglobin A_{1c}, %, mean (SD)	5.4 (0.3)
FPG, mg/dL, mean (SD)	87.6 (7.7)
Fasting lipid profile, mg/dL, median (IQR)[b]	
Total cholesterol	177.2 (154.1–201.2)
HDL cholesterol	44.0 (37.5–51.7)
LDL cholesterol	110.8 (91.9–130.9)
vLDL cholesterol	18.5 (14.3–24.7)
Free fatty acids	12.5 (8.7–18.0) [n = 802]
Triglycerides	94.3 (73.0–125.5)
SF-36 physical functioning score, mean (SD)[b,c]	53.9 (5.5) [n = 802]
Pulse, bpm, mean (SD)[c]	76 (9)
eGFR, mL/min/1.73 m^2, median (IQR)[d,e]	94.9 (82.0–107.4)

Abbreviations: BMI, body mass index; DBP, diastolic blood pressure; eGFR, estimated glomerular filtration rate; FPG, fasting plasma glucose; HDL, high-density lipoprotein; IQR, interquartile range; LDL, low-density lipoprotein; SBP, systolic blood pressure; SD, standard deviation; SF-36, Short Form36v2® Health Survey, Acute Version; vLDL, very-low-density lipoprotein.

[a]Demographics data (age, sex, race, and ethnicity) and comorbidity data are not reported as they were the same as at week 0 (see Table 1).

[b]Participant numbers are provided where the number analyzed differed from the number in the full analysis set.

[c]SF-36 measures health-related quality of life and general health status. SF-36 scores are norm-based scores. Norm-based scores are scores transformed to a scale where the 2009 US general population has a mean of 50 and an SD of 10. Further information on the SF-36 is provided in eAppendix 4, Supplement 1.

[d]Data are for the safety analysis set.

[e]Assessed at screening (week –1).

eTable 3. Geometric Mean Values for Fasting Lipid Parameters and Estimated Glomerular Filtration Rate at Weeks 0 and 20 (Full Analysis Set)

	Week 0 (start of run-in period with semaglutide treatment) (N = 803)	Change during run-in period[a]	Week 20 (randomization)		
			Continued semaglutide 2.4 mg/wk (N = 535)	Switched to placebo (N = 268)	Total (N = 803)
Fasting lipid profile, mg/dL, geometric mean (CV)[b]					
Total cholesterol	193.1 (19.5) [n = 798]	0.9 (14.5)	175.9 (20.3)	175.1 (20.8)	175.6 (20.4)
HDL-C	49.6 (24.8) [n = 798]	0.9 (13.7)	44.5 (21.6)	43.6 (22.5)	44.2 (21.9)
LDL-C	114.8 (28.8) [n = 798]	0.9 (20.5)	108.7 (29.2)	109.1 (30.5)	108.8 (29.6)
vLDL-C	23.7 (47.4) [n = 798]	0.8 (33.7)	19.2 (42.1)	18.6 (43.4)	19.0 (42.5)
Free fatty acids	12.1 (60.9) [n = 789]	1.0 (73.7)	12.3 (57.9) [n = 534]	11.7 (62.0)	12.1 (59.3) [n = 802]
Triglycerides	121.5 (48.3) [n = 798]	0.8 (34.4)	98.1 (42.3)	95.3 (43.4)	97.1 (42.7)
eGFR, mL/min/1.73 m^2, geometric mean (CV)[c,d]	97.4 (19.1)	0.9 (10.5)	91.1 (22.1)	94.6 (18.3)	92.2 (21.0)

Abbreviations: CV, coefficient of variation (in percentage); eGFR, estimated glomerular filtration rate; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; vLDL-C, very-low-density lipoprotein cholesterol.

See Table 1 and eTable 2 for median (interquartile range) values for these parameters.

[a]Observed geometric mean (CV) ratio to week 0 for the individual participants in the total randomized population.

[b]Participant numbers are provided where the number analyzed differed from the number in the full analysis set.

[c]Data are for the safety analysis set.

[d]Assessed at screening (week –1).

eTable 4. Demographics and Clinical Characteristics at Start of Run-In (Week 0) for Enrolled Participants and Non-Randomized Participants

	Enrolled participants (N = 902)	Participants excluded from randomization (N = 99)
Age, years, mean (SD)	46 (12)	46 (14)
Female sex, No. (%)	717 (79.5)	83 (83.8)
Race, No. (%)[a]		
White	751 (83.3)	79 (79.8)
Black or African American	123 (13.6)	19 (19.20)
Asian	19 (2.1)	0
Other[b]	9 (1.0)	1 (1.0)
Hispanic or Latino ethnic group, No. (%)	70 (7.8)	7 (7.1)
Body weight, kg, mean (SD)	106.8 (23.1)	103.6 (25.8)
BMI, kg/m^2		
Mean (SD)	38.3 (7.0)	37.9 (7.4)
<25, No. (%)	0	0
≥25–<30, No. (%)	30 (3.3)	8 (8.1)
≥30–<35, No. (%)	308 (34.1)	34 (34.3)
≥35–<40, No. (%)	276 (30.6)	27 (27.3)
≥40, No. (%)	288 (31.9)	30 (30.3)
Waist circumference, cm, mean (SD)	115.1 (15.6)	113.3 (16.5)

Abbreviations: BMI, body mass index; SD, standard deviation.

[a]To meet regulatory requirements, race and ethnicity were recorded in this study and were determined by the participant according to fixed selection categories (with the option of answering 'other', 'not applicable' or 'unknown').

[b]Native Hawaiian or Other Pacific Island, or Other.

eTable 5. Changes in Efficacy Endpoints During the Randomized Period (Weeks 20–68; Trial Product Estimand[a]; Full Analysis Set)

	Estimated mean change (95% CI)			
	Continued semaglutide 2.4 mg/wk (N = 535)	Switched to placebo (N = 268)	Difference[b]	*P* value
Primary endpoint				
Body weight, % change	–8.8 (–9.5 to –8.1)	+6.5 (5.6 to 7.5)	–15.3 (–16.5 to –14.1)	<.001
Confirmatory secondary endpoints				
Waist circumference, cm	–7.3 (–7.9 to –6.6)	+3.0 (2.1 to 4.0)	–10.3 (–11.4 to –9.2)	<.001
SBP, mmHg	+0.1 (–1.0 to 1.2)	+4.9 (3.3 to 6.5)	–4.9 (–6.8 to –2.9)	<.001
SF-36 physical functioning score[c]	+1.0 (0.7 to 1.3)	–1.0 (–1.5 to –0.5)	+2.0 (1.4 to 2.6)	<.001
Supportive secondary endpoints				
Body weight, kg	–8.0 (–8.6 to –7.3)	+5.6 (4.7 to 6.5)	–13.6 (–14.7 to –12.5)	<.001
BMI, kg/m^2	–2.9 (–3.1 to –2.7)	+2.0 (1.7 to 2.4)	–5.0 (–5.3 to –4.6)	<.001
DBP, mmHg	+0.3 (–0.5 to 1.0)	+1.6 (0.5 to 2.7)	–1.3 (–2.6 to 0.0)	.052
Hemoglobin A_{1c}, %	–0.2 (–0.2 to –0.1)	+0.1 (0.1 to 0.1)	–0.3 (–0.3 to –0.2)	<.001
FPG, mg/dL	–1.4 (–2.2 to –0.7)	+7.1 (6.0 to 8.2)	–8.6 (–9.9 to –7.2)	<.001
Fasting serum insulin, % change[d]	–19 (–23 to –15)	–1 (–7 to 6)	–19 (–25 to –12)	<.001
Fasting lipid profile, % change[d]				
Total cholesterol	+5 (3 to 6)	+10 (9 to 12)	–5 (–7 to –3)	<.001
HDL cholesterol	+18 (16 to 19)	+19 (17 to 21)	–1 (–4 to 1)	.23
LDL cholesterol	+1 (–1 to 2)	+6 (4 to 9)	–5 (–8 to –2)	<.001
vLDL cholesterol	–6 (–9 to –3)	+9 (4 to 14)	–14 (–18 to –9)	<.001
Free fatty acids	–21 (–26 to –17)	–13 (–20 to –5)	–10 (–18 to 0)	.047

	Estimated mean change (95% CI)			
	Continued semaglutide 2.4 mg/wk (N = 535)	Switched to placebo (N = 268)	Difference[b]	*P* value
Triglycerides	–6 (–9 to –3)	+9 (4 to 14)	–14 (–18 to –9)	<.001
SF-36 scores[c]				
Physical component summary	+0.8 (0.4 to 1.2)	–0.7 (–1.3 to –0.1)	+1.5 (0.7 to 2.2)	<.001
Mental component summary	+0.1 (–0.4 to 0.7)	–2.4 (–3.2 to –1.5)	+2.5 (1.5 to 3.5)	<.001
Participants who gained weight, No. (%)[e]	61 (12.3)	190 (81.2)	0.0 (0.0 to 0.0)	<.001

Abbreviations: BMI, body mass index; CI, confidence interval; DBP, diastolic blood pressure; FPG, fasting plasma glucose; HDL, high-density lipoprotein; LDL, low-density lipoprotein; SBP, systolic blood pressure; SF-36, Short Form36v2® Health Survey, Acute Version; vLDL, very low-density lipoprotein.

[a]Assessed the treatment effect assuming participants remained on randomized treatment for the planned study duration without rescue intervention, analyzed using a mixed model for repeated measurements, with randomized treatment as a factor and baseline endpoint value as a covariate.[4] Analyses were not controlled for multiple comparisons. See Table 2 for corresponding data for the treatment policy estimand (assessed the treatment effect regardless of treatment discontinuation or rescue intervention).

[b]Absolute differences between estimated mean changes unless stated otherwise. The differences between mean percent changes in body weight, fasting serum insulin, and the fasting lipid profile and mean changes in hemoglobin A_{1c} are expressed in percentage points.

[c]SF-36 measures health-related quality of life and general health status. SF-36 scores are norm-based scores. Norm-based scores are scores transformed to a scale where the 2009 US general population has a mean of 50 and an SD of 10. Further information on the SF-36 is provided in eAppendix 4, Supplement 1.

[d]Data are presented as percent change from week 20 to 68, and estimated relative differences, calculated based on the estimated ratio to baseline and the estimated treatment ratio from the corresponding analysis of covariance. The formula for calculation was: (estimated ratio – 1)*100.

[e]Data are presented as observed proportions of participants who gained weight from week 20 to 68, and estimated odds ratios.

eTable 6. Change in Antihypertensive Medication During the Randomized Period (Weeks 20–68; Observed In-Trial Data; Full Analysis Set)

	Continued semaglutide 2.4 mg/wk (N = 149)	Switched to placebo (N = 67)
No change	97 (65.1)	48 (71.6)
Stopped/decreased	38 (25.5)	8 (11.9)
Increased	14 (9.4)	11 (16.4)

N represents the number of participants taking antihypertensive medication at week 20. Data are No. (%). Change in antihypertensive medication was based on the investigator's assessment of overall use.

eTable 7. Change in Lipid-Lowering Medication During the Randomized Period (Weeks 20–68; Observed In-Trial Data; Full Analysis Set)

	Continued semaglutide 2.4 mg/wk (N = 70)	Switched to placebo (N = 36)
No change	59 (84.3)	27 (75.0)
Stopped/decreased	8 (11.4)	4 (1.6)
Increased	3 (4.3)	5 (13.9)

N represents the number of participants taking lipid-lowering medication at week 20. Data are No. (%). Change in lipid lowering medication was based on the investigator's assessment of overall use.

eTable 8. Adverse Event and Tolerability Profile During the Run-In Period (Weeks 0–20; Safety Analysis Set)

	Semaglutide (N = 902)		
	No. (%)	Events	Events per 100 patient years
Any AE	760 (84.3)	3775	1105
Serious AEs	21 (2.3)	23	6.7
Discontinuation of trial product due to AEs[a]	48 (5.3)	–	–
Fatal events[b]	0	–	–
AEs reported in ≥5% of participants[c]			
Nausea	422 (46.8)	629	184.2
Diarrhea	212 (23.5)	309	90.5
Constipation	200 (22.2)	232	67.9
Vomiting	140 (15.5)	239	70.0
Dyspepsia	103 (11.4)	127	37.2
Decreased appetite	102 (11.3)	115	33.7
Headache	96 (10.6)	119	34.8
Nasopharyngitis	92 (10.2)	102	29.9
Eructation	71 (7.9)	88	25.8
Abdominal pain	68 (7.5)	84	24.6
Fatigue	67 (7.4)	69	20.2
Gastroesophageal reflux disease	58 (6.4)	60	17.6
Abdominal distension	50 (5.5)	53	15.5
Flatulence	50 (5.5)	73	21.4
Abdominal pain upper	49 (5.4)	64	18.7
Safety areas of interest (MedDRA)[d]			
Gastrointestinal disorders	644 (71.4)	2148	628.9
Gallbladder-related disorders	6 (0.7)	6	1.8
Hepatic disorders	4 (0.4)	5	1.5
Cardiovascular disorders[b]	30 (3.3)	37	10.7
Allergic reactions	29 (3.2)	34	10.0
Injection site reactions	25 (2.8)	33	9.7
Malignant neoplasms[b]	1 (0.1)	1	0.3
Psychiatric disorders	55 (6.1)	70	20.5
Acute renal failure	1 (0.1)	2	0.6
Hypoglycemia	1 (0.1)	1	0.3

Abbreviations: AE, adverse event; MedDRA, Medical Dictionary for Regulatory Activities.

Data are for treatment-emergent AEs during the on-treatment period (any dose of trial product administered within the prior 49 days), unless stated otherwise.

[a]Based on the number of participants from the full analysis set stating an AE as the reason for permanent trial product discontinuation on the end of treatment form.

[b]In-trial observation period data (time from randomization to last contact with trial site, irrespective of treatment discontinuation or rescue intervention).

[c]Most common AEs by preferred term reported in ≥5% of participant in either treatment group.

[d]Identified via MedDRA searches.

eTable 9. Other Reasons for Withdrawal From Trial Prior to Randomization (Run-in Failure Participants)

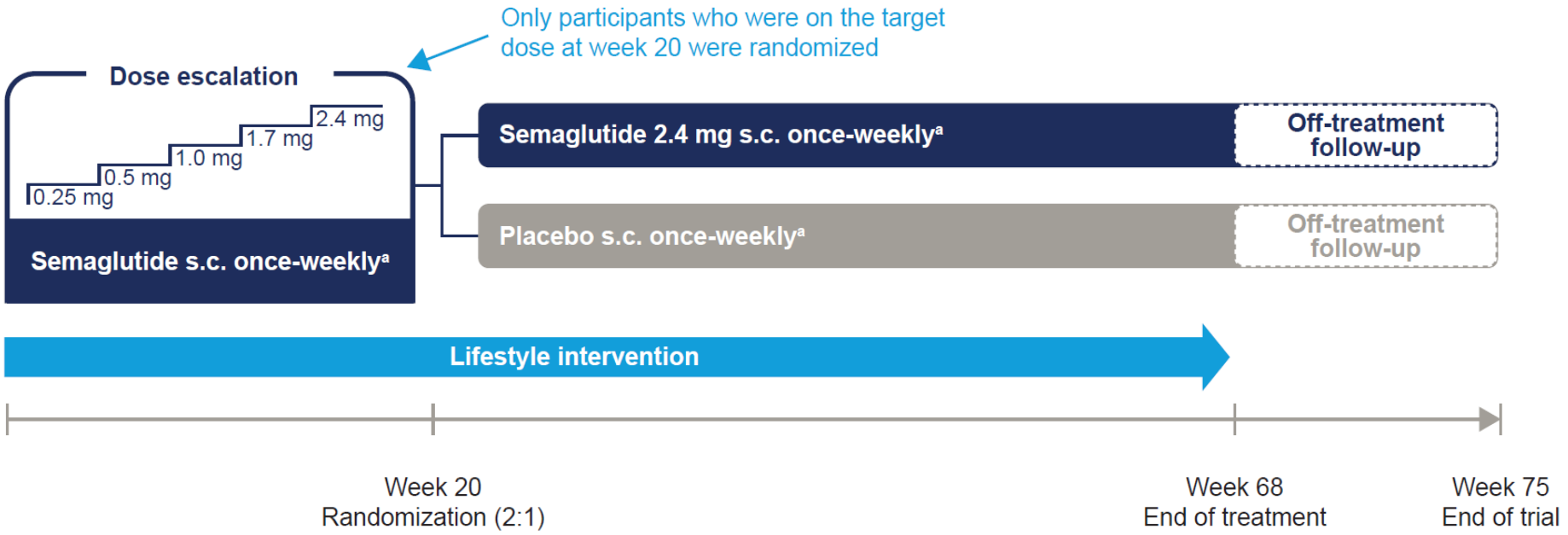
Participant (N = 9)	Primary reason
453003	Participant felt that study drug was not effective as expected
455005	Very busy, not able to commit to study visits schedule and diet
501009	Participant did not want to take investigational product because they believed it made them ill
501011	Participant decision; participant reported that they felt sick/ill after injecting
503010	Participant decision to stop investigational product
606015	Participant moved away
609010	Participant not compliant
611007	Participant moved away and was not willing to transfer to another study site
614007	Participant did not randomize due to medication tolerability issues

eTable 10. Other Reasons for Premature Discontinuation of Trial Product After Randomization (Randomized Participants)

Participant	Primary reason
Randomized to continued semaglutide 2.4 mg/wk (N = 12)	
100006	Participant chose to withdraw from the trial due to a difficult family situation
403005	Participant was not willing to take study medication anymore because they did not want to lose more weight
403013	Due to enough weight loss (in participant's opinion) not willing to participate anymore
601015	Per sponsor, if more than two consecutive doses missed, participant did not complete treatment
605009	Missing to administer the trial product
609014	Participant's work schedule
612005	Participant moved and only agreed to phone call at end of study
620013	Participant did not want to continue in the trial
621012	Participant maintains they were compliant but all duns returned since 03Oct2019 appear unused
624017	Participant did not want to take medication anymore
624018	Participant was concerned with significant weight loss and did not want to lose any more weight
629019	Participant refused
Randomized to placebo (N = 23)	
100003	Due to close relative being seriously ill
200005	Fear of injection
201015	Did not experience a positive effect
251001	Participant decided to continue the study through telephone visits, came to the site to return study investigational product
354009	Did not want to participate in the study anymore
401001	Participant was not willing to take study medication anymore
403002	Too busy to attend visit
403004	Not willing to participate anymore
403007	Not willing to participate anymore due to many control visits because of cochlea implant
407008	Participant was not willing to take study medication anymore, lack of motivation
452004	Participant decision
452014	Participant decision
501004	Participant missed several visit and did not take treatment during that time
501005	Participant missed doses
603005	Participant non-compliance per investigational product review; participant states they were compliant

Participant	Primary reason
604015	Discontinued due to inability to come in for visits due to family illness
610015	Personal life stressor; too much to do outside of study
613003	Participant decided after randomization to go vegan and they also no longer wanted to comply with diaries
618001	The participant cancelled their final appointment due to 'family emergency' and then became a lost to follow up
620008	Participant relocated out of state for work
620012	Participant decided not to resume investigational product to avoid recurrent gastritis
620015	Participant did not want to take drug due to life difficulties
625008	Participant choice

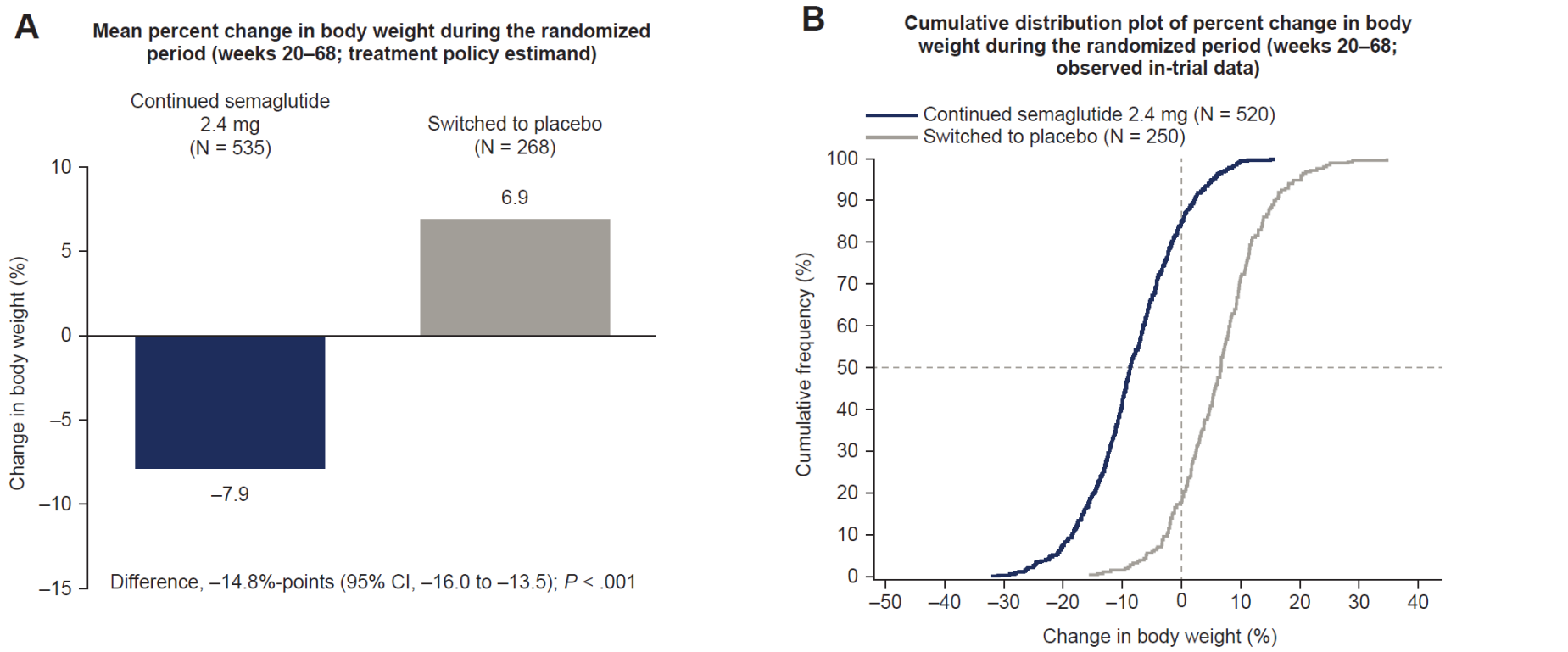
eFigure 1. STEP 4 Study Design



Abbreviations: s.c., subcutaneous.

[a]As an adjunct to lifestyle intervention (–500 kcal/day diet with 150 min/week physical activity).

eFigure 2. Effect of Semaglutide 2.4 mg Once Weekly Compared With Placebo on Body Weight During the Randomized Period (Weeks 20–68; Full Analysis Set)



Abbreviations: CI, confidence interval.

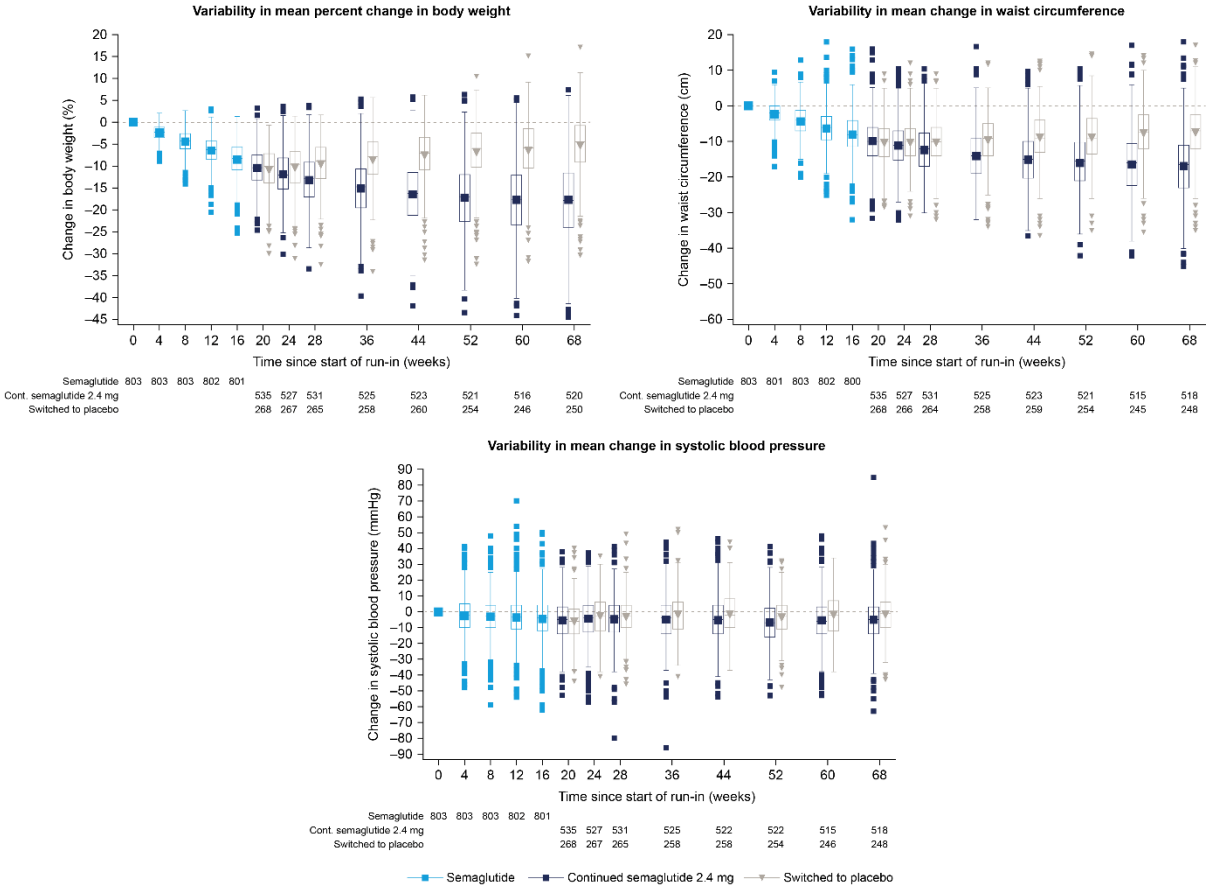
Panel A: Data presented are the estimated mean percent changes from weeks 20 to 68 in body weight, with the difference for continued semaglutide vs placebo and associated 95% CI, analyzed using the full analysis set based on the treatment policy estimand[a]

Panel B: Cumulative distribution plot of observed percent change in body weight from weeks 20 to 68 for participants in the full analysis set during the in-trial observation period[b]

[a]The treatment policy estimand assessed the treatment effect regardless of treatment discontinuation or rescue intervention. Analyzed using analysis of covariance, with randomized treatment as a factor and baseline endpoint value as a covariate, and a multiple imputation approach for missing data.[4]

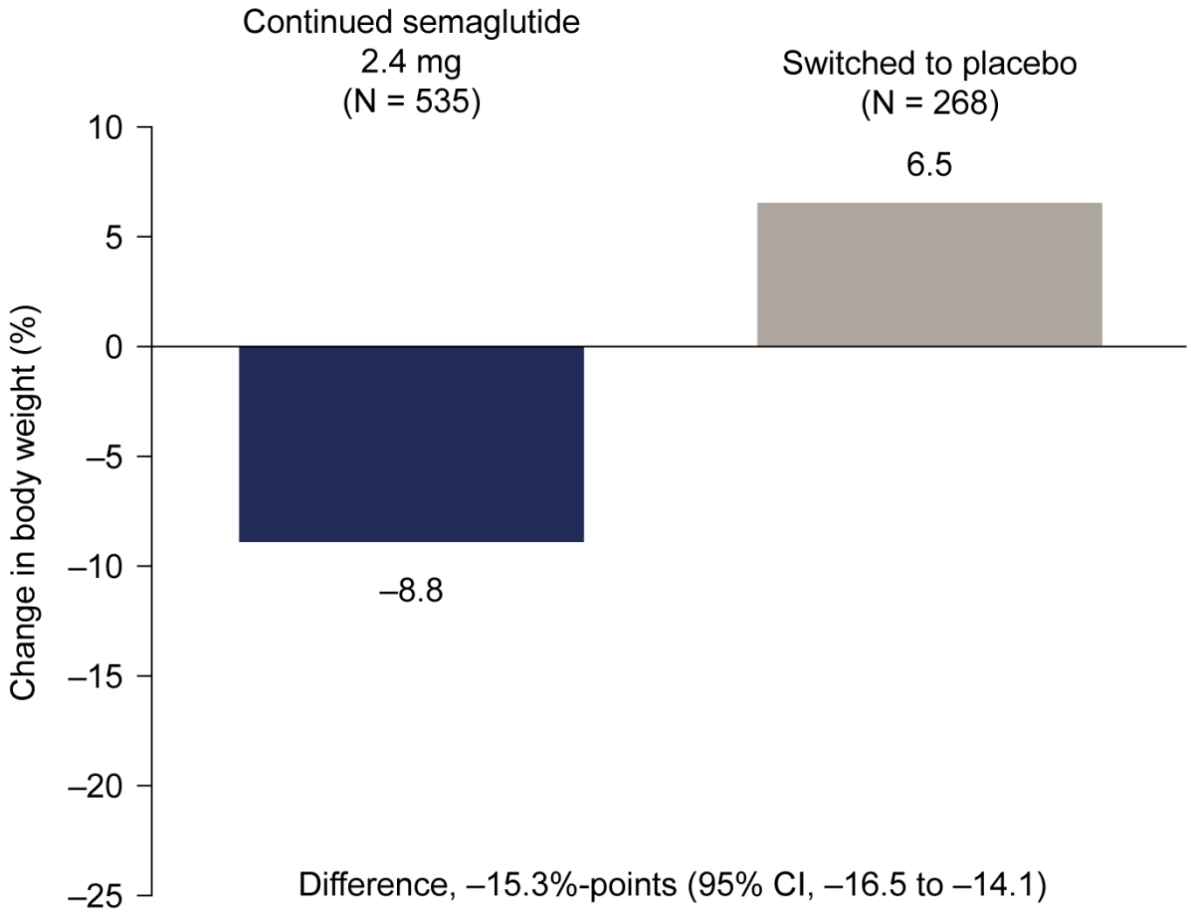
[b]The time from week 0 to date of last contact with trial site.

eFigure 3. Variability in Selected Efficacy Outcomes During the Entire Trial (Weeks 0–68; Observed In-Trial Data; Full Analysis Set)



Data presented are observed data for the full analysis set from the in-trial period. The middle lines represent median observed changes from week 0, symbols in the boxes represent mean observed percent change, box tops and bottoms represent interquartile range, whiskers extend to the most extreme observed values with 1.5 times the interquartile range of the nearer quartile, and symbols beyond these points represent observed values outside that range. More negative values indicate greater reductions.

eFigure 4. Percent Change in Body Weight During the Randomized Period (Weeks 20–68; Trial Product Estimand[a]; Full Analysis Set)

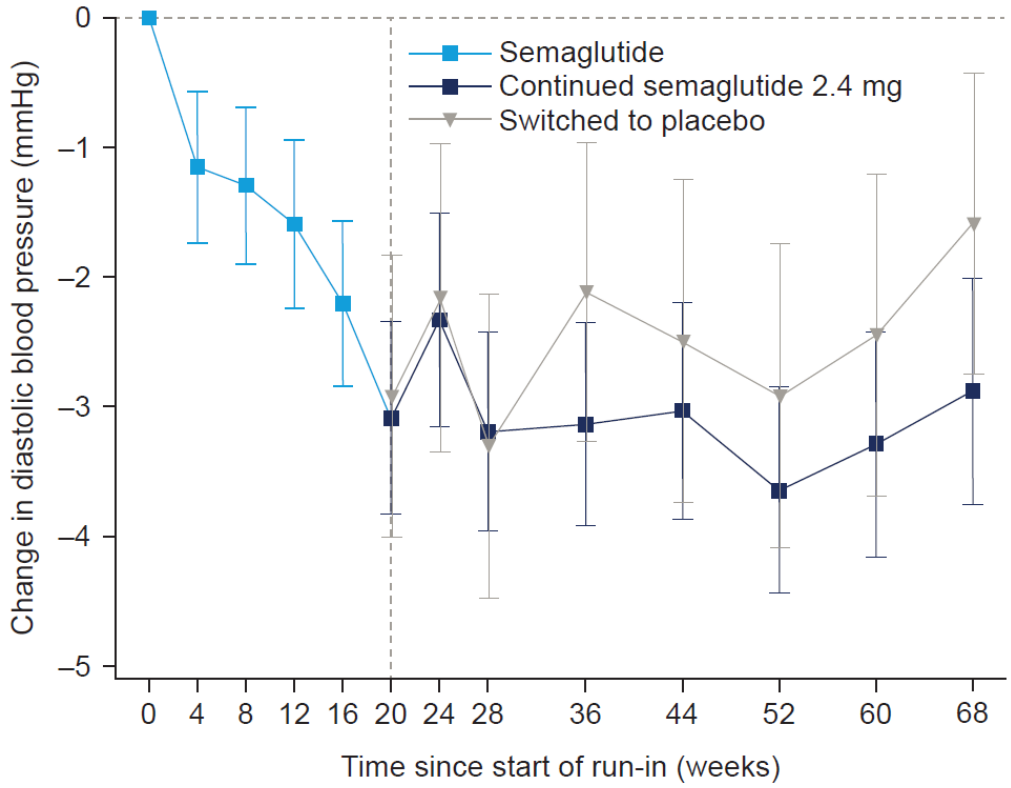


Abbreviations: CI, confidence interval.

Data presented are the estimated mean percent changes from weeks 20 to 68 in body weight for the full analysis set, with the absolute difference for continued semaglutide vs placebo and associated 95% CI, analyzed using the full analysis set based on the trial product estimand.[a]

[a]Assessed the treatment effect assuming participants remained on randomized treatment for the planned study duration without rescue intervention, analyzed using a mixed model for repeated measurements, with randomized treatment as a factor and baseline endpoint value as a covariate.[4]

eFigure 5. Mean Change in Diastolic Blood Pressure During the Entire Trial (Weeks 0–68; Observed In-Trial Data; Full Analysis Set)

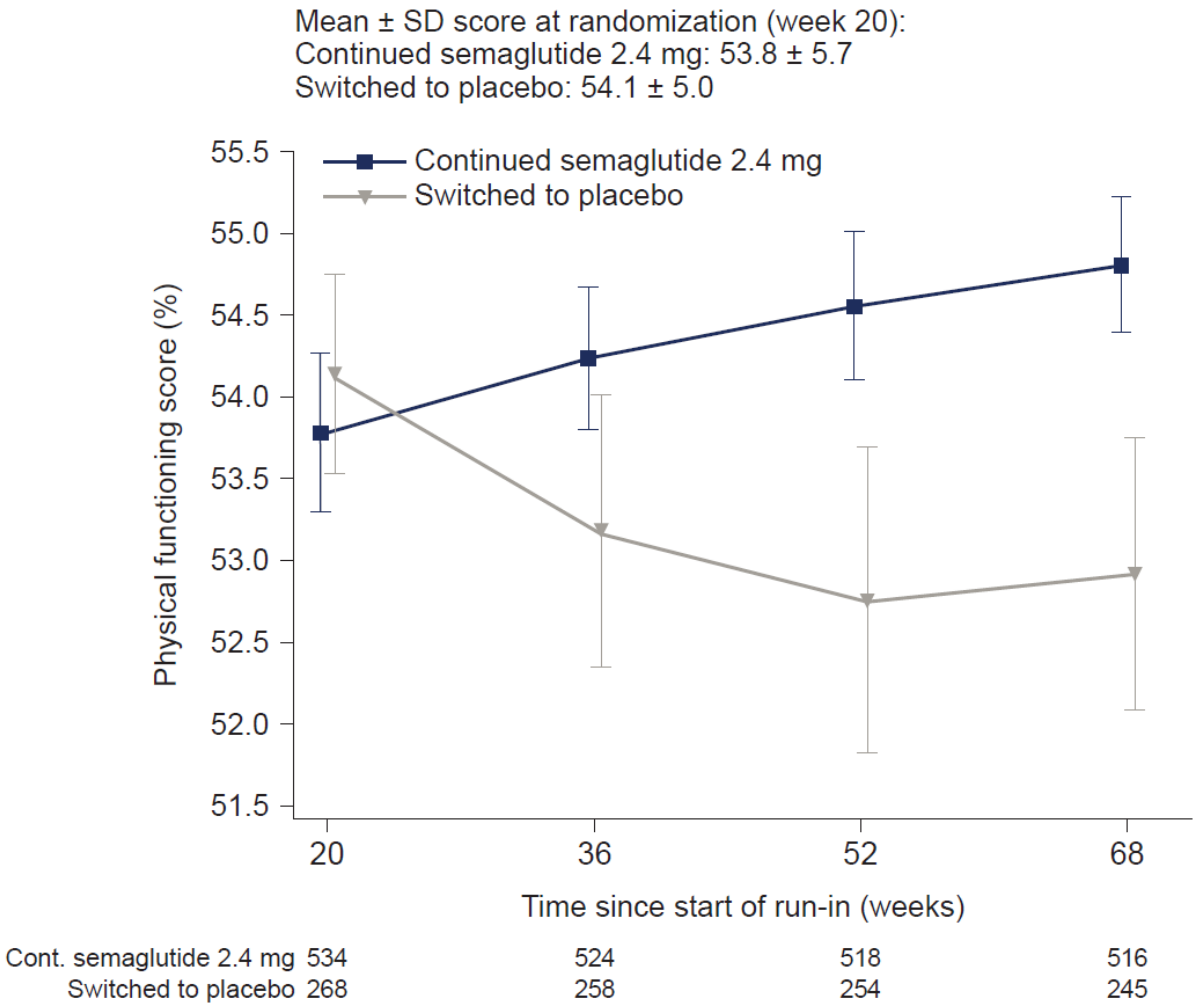


Semaglutide	803	803	803	802	801			
Cont. semaglutide 2.4 mg			535	527	531	525	522	522	515	518
Switched to placebo			268	267	265	258	258	254	246	248

Abbreviations: cont., continued.

Data presented are observed data for the full analysis set from the in-trial period (time from week 0 to date of last contact with trial site). Error bars represent 95% confidence intervals for the mean. Participant numbers shown denote participants contributing to the mean. The dashed vertical line at week 20 represents the randomization timepoint.

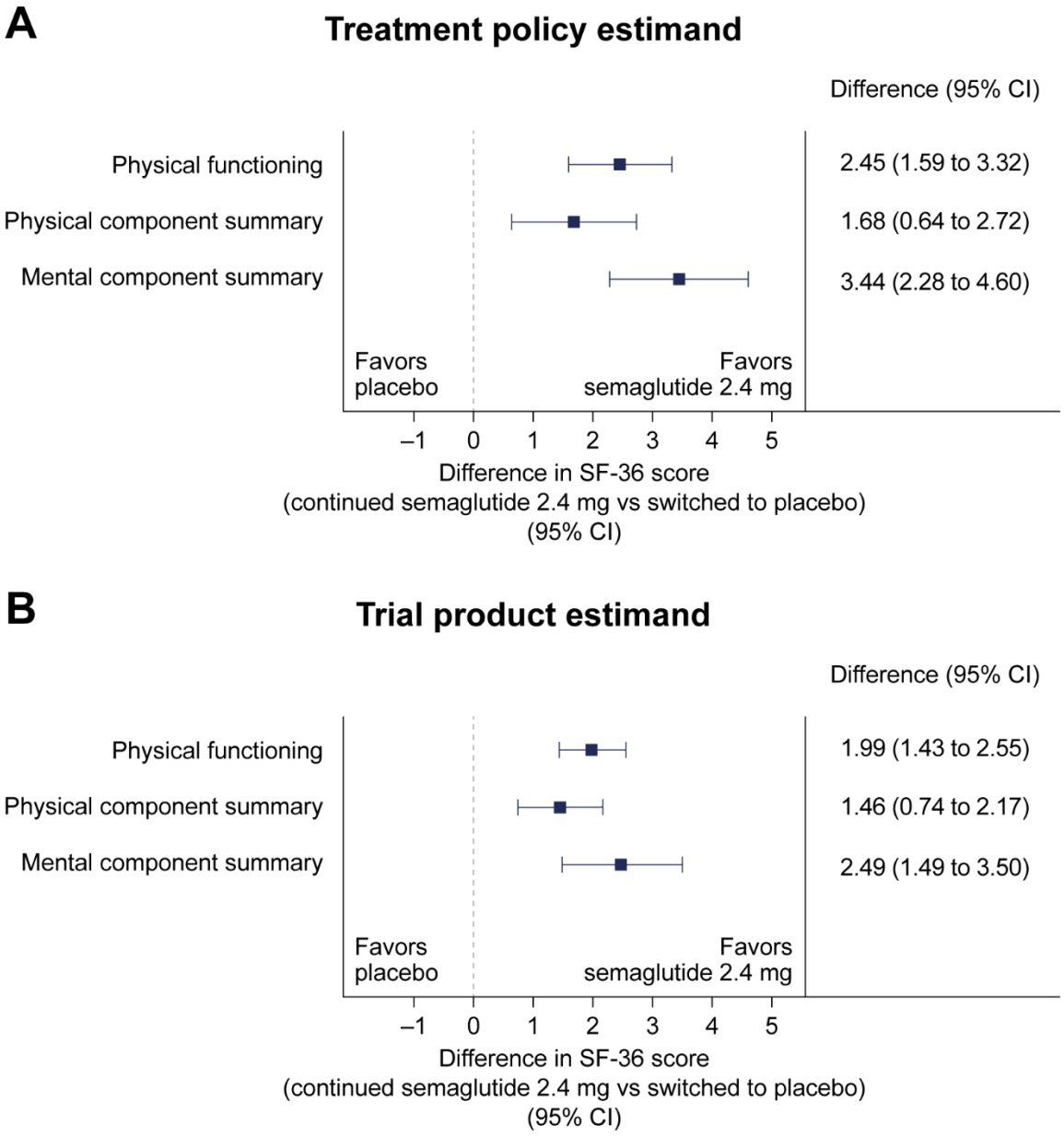
eFigure 6. Mean SF-36 Physical Functioning Score During the Randomized Period (Weeks 20–68; Observed In-Trial Data; Full Analysis Set)



Abbreviations: CI, confidence interval; cont., continued; SD, standard deviation; SF-36, Short Form36v2® Health Survey, Acute Version.

Data presented are observed data for the full analysis set from the in-trial period (time from week 20 to date of last contact with trial site). Error bars represent 95% confidence intervals for the mean. Participant numbers shown denote participants contributing to the mean. SF-36 measures health-related quality of life and general health status. SF-36 scores are norm-based scores. Norm-based scores are scores transformed to a scale where the 2009 US general population has a mean of 50 and an SD of 10. Further information on the SF-36 is provided in eAppendix 4, Supplement 1.

eFigure 7. Change in SF-36 Domain Scores During the Randomized Period (Weeks 20–68; Full Analysis Set)



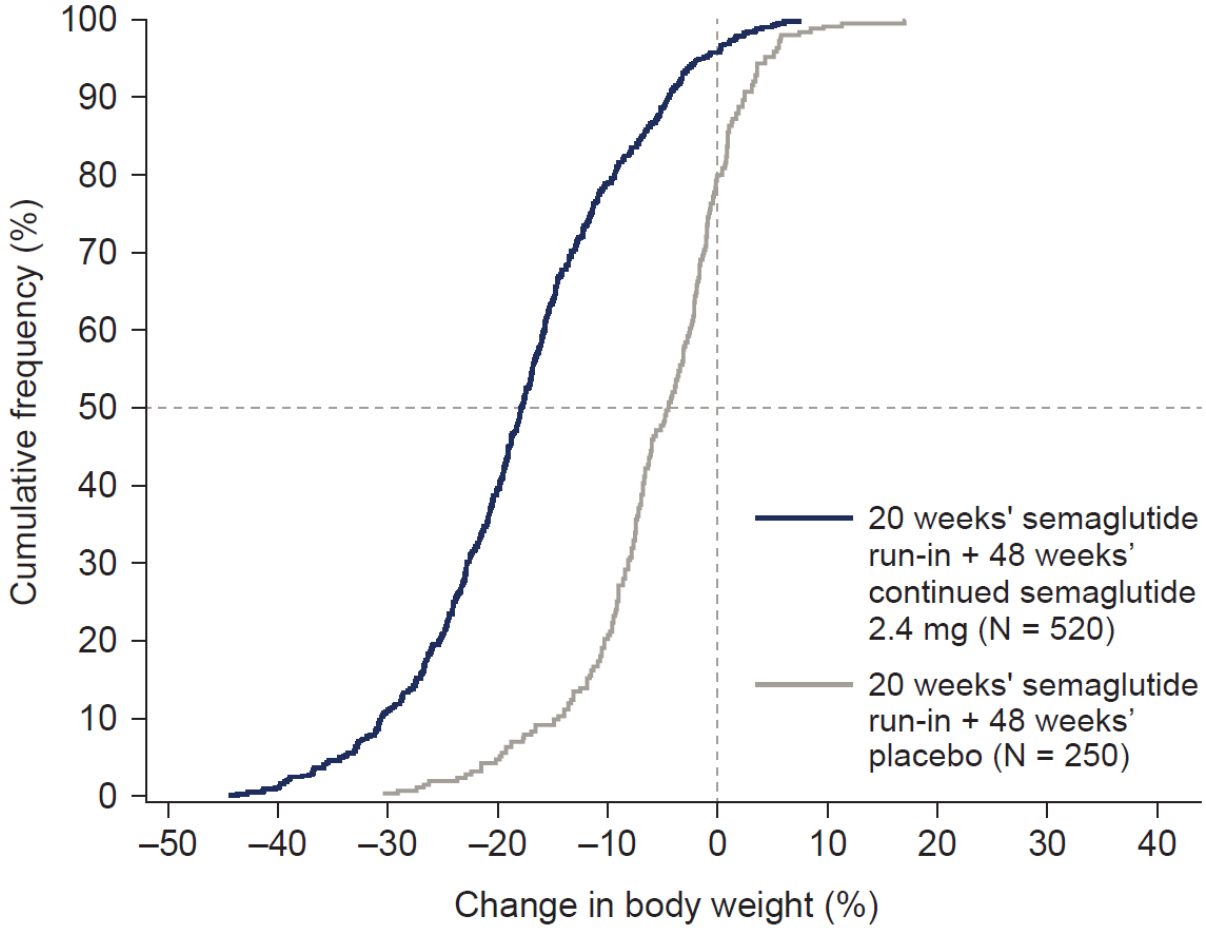
Abbreviations: CI, confidence interval; SF-36, Short Form36v2® Health Survey, Acute Version.

Panel A: Treatment policy estimand data – assessed the treatment effect in all randomized participants regardless of treatment discontinuation or rescue intervention. Analyzed using analysis of covariance, with randomized treatment as a factor and baseline endpoint value as a covariate, and a multiple imputation approach for missing data.[4]

Panel B: Trial product estimand – assessed the treatment effect in all randomized assuming they remained on randomized treatment for the planned study duration without rescue intervention. Analyzed using a mixed model for repeated measurements, with randomized treatment as a factor and baseline endpoint value as a covariate.[4]

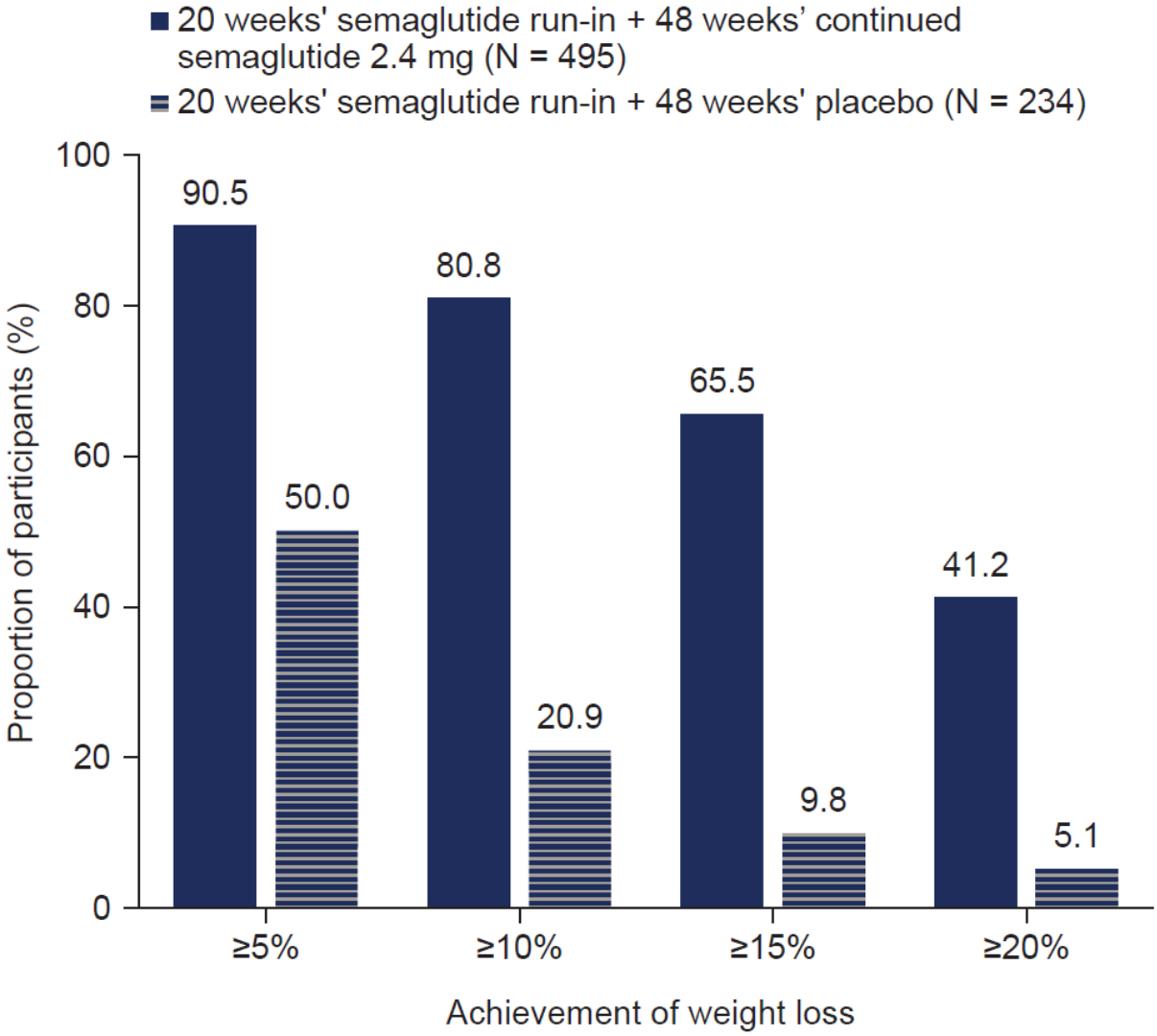
SF-36 measures health-related quality of life and general health status. SF-36 scores are norm-based scores. Norm-based scores are scores transformed to a scale where the 2009 US general population has a mean of 50 and a standard deviation of 10. Further information on the SF-36 is provided in eAppendix 4, Supplement 1.

eFigure 8. Cumulative Distribution Plot of Percent Change in Body Weight During the Entire Trial (Weeks 0–68; Observed In-Trial Data; Full Analysis Set)



Data presented are observed data for the full analysis set from the in-trial period (time from week 0 to date of last contact with trial site).

eFigure 9. Proportion of Participants Achieving Thresholds of Weight Loss During the Entire Trial (Weeks 0–68; Observed On-Treatment Data; Full Analysis Set)



Data presented are observed data among all randomized participants in the full analysis set with a week 68 assessment while on-treatment (defined as any dose of trial product within prior 14 days).

eReferences

1. Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2012 Clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney Int Suppl*. 2013;3(1):1-150.
2. Ware JE Jr, Gandek B. Overview of the SF-36 Health Survey and the International Quality of Life Assessment (IQOLA) Project. *J Clin Epidemiol*. 1998;51(11):903-912.
3. American Diabetes Association. 2. Classification and diagnosis of diabetes. *Diabetes Care*. 2017;40(Suppl):S11-S24.
4. Kushner RF, Calanna S, Davies M, et al. Semaglutide 2.4 mg for the treatment of obesity: key elements of the STEP Trials 1 to 5. *Obesity*. 2020;28(6):1050-1061.