Supplementary Appendix

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

Supplement to: Wilding JPH, Batterham RL, Calanna S, et al. Once-weekly semaglutide in adults with overweight or obesity. N Engl J Med 2021;384:989-1002. DOI: 10.1056/NEJMoa2032183

SUPPLEMENTARY APPENDIX

CONTENTS

LIST OF INVESTIGATORS IN THE SEMAGLUTIDE TREATMENT EFFECT IN PEOPLE WITH OBESITY (STEP) 1 TRIAL

Argentina: Marianela Aguirre Ackermann (Corrientes), Cecilia Luquez (Córdoba), Marcos Mayer (Santa Rosa), Carla Musso (CABA), Susana Salzberg (Ciudad de Buenos Aires)

Belgium: Ides Colin (Boussu), Ann Mertens (Leuven), André Scheen (Liège), Jean-Paul Thissen (Bruxelles), Luc Van Gaal (Edegem)

Bulgaria: Zhivka Asyova (Sofia), Anna-Maria Borissova (Sofia), Nickolay Botushanov (Plovdiv), Ivona Daskalova (Sofia), Zdravko Kamenov (Sofia)

Canada: Adam Blackman (Toronto), Martin D'Amours (Quebec), Isabelle Labonte (Quebec), Stephanie Li (Edmonton), Derek Lowe (Surrey), Sean Wharton (Hamilton), Sanaz Zarinehbaf-Asadi (North York)

Denmark: Bjørn Richelsen (Aarhus N)

Finland: Kirsi Pietiläinen (Helsinki), Markku Savolainen (Oulu)

France: Sebastien Czernichow (Paris), Emmanuel Disse (Pierre Benite), Kamel Mohammedi (Pessac), Arnaud Monier (Le Coudray), Christine Poitou-Bernert (Paris), Pierre Serusclat (Venissieux), Jean-Francois Thuan (Narbonne)

Germany: Christel Contzen (Frankfurt), Moritz Mauro Erlinger (Stuttgart), Michael Esser (Essen), Thomas Linn (Giessen), Jörg Lüdemann (Falkensee), Karsten Milek (Hohenmölsen), Nicoletta Nalazek (Leipzig), Andrea Rinke (Bochum), Joachim Sauter (Wangen), Thomas Schürholz (Essen), Alexander Segner (St. Ingbert-Oberwürzbach), Liana Vismane (Berlin), Ulrich Wendisch (Hamburg)

India: Syamasis Bandyopadhyay (Kolkata), Dipti Chand (Nagpur), Piyush Desai (Surat), Vaishali Deshmukh (Pune), Yashdeep Gupta (New Delhi), P K Jabbar (Thiruvananthapuram), Dinesh Jain (Ludhiana), Neelaveni K (Hyderabad), Shriraam Mahadevan (Chennai), Rajesh Rajput (Rohtak), Sudhakar Reddy (Secunderabad), Kongara Srikanth (Guntur), A Unnikrishnan (Pune)

Japan: Satoshi Inoue (Suita-shi, Osaka), Arihiro Kiyosue (Tokyo), Osamu Matsuoka (Tokyo), Hiraku Ono (Chiba-shi, Chiba), Masamichi Yamada (Tokyo)

Mexico: Diego Espinoza Peralta (Hermosillo), Silvia Jimenez-Ramos (Guadalajara), Carlos Medina Pech (Merida)

Poland: Pawel Bogdanski (Poznan), Malgorzata Jozefowska (Lodz), Agata Leksycka (Gdynia), Jaroslaw Ogonowski (Szczecin)

Russian Federation: Diana Alpenidze (Saint-Petersburg), Olga Ershova (Yaroslavl), Marina Kharakhulakh (Tomsk), Vadim Klimontov (Novosibirsk), Ludmila Ruyatkina (Novosibirsk), Marina Sergeeva-Kondrachenko (Penza), Ekaterina Troshina (Moscow), Elena Zhdanova (Voronezh)

Taiwan: Kuo-Chin Huang (Taipei)

United Kingdom: Rachel Batterham (London), Matt Capehorn (Rotherham), Rhodri King (Taunton), Michael Lean (Glasgow), Barbara McGowan (London), Khin Swe Myint (Norwich), Adrian Park (Cambridge), Harpal Randeva (Coventry), Georgina Russell (Bristol), John Wilding (Liverpool)

United States: Hanid Audish (Spring Valley), Darlene Bartilucci (Jacksonville), Harold Bays (Louisville), Ronald Brazg (Renton), Robert Broker (Simpsonville), Kevin Cannon (Wilmington), Tira Chaicha-Brom (Austin), Matthew Davis (Rochester), H. Jackson Downey (Jacksonville), Stephen Fehnel (West Reading), Almena Free (Anniston), Amina Haggag (Anaheim), Mitzie Hewitt (Buckley), Priscilla Hollander (Dallas), Misal Khan (Panama City), Karen Laufer (Plantation), Robert McNeill (Salisbury), John Nardandrea, Jr (Ocala), Lisa Neff (Chicago), Kevin Niswender (Nashville), Patrick O'Neil (Charleston), John Pullman (Butte), Marina Raikhel (Lomita), Scott Redrick (Crystal River), John Reed III (Roswell), Michele Reynolds (Dallas), Luis Rivera-Colon (San Juan), Julio Rosenstock (Dallas), Erich Schramm (Ponte Vedra), John Scott (Richmond), Stephanie Shaw (Round Rock), Vijay Shivaswamy (Omaha), Timothy Smith (St. Peters), Joseph Soufer (Waterbury), Stephen Straubing (Chiefland), Danny Sugimoto (Chicago), Phillip Toth (Indianapolis), Ralph Wade (Bountiful), Holly Wyatt (Aurora), lan Wynne (Topeka).

SUPPLEMENTAL METHODS

INCLUSION AND EXCLUSION CRITERIA

Inclusion criteria

Subjects are eligible to be included in the trial only if all of the following criteria apply:

- Informed consent obtained before any trial-related activities. Trial-related activities are any procedures that are carried out as part of the trial, including activities to determine suitability for the trial.
- Male or female, age ≥18 years at the time of signing informed consent.
- Body mass index (BMI) ≥30.0 kg/m^2 or ≥27.0 kg/m^2 with the presence of at least one of the following weight-related comorbidities (treated or untreated): hypertension, dyslipidemia, obstructive sleep apnea, or cardiovascular disease.
- History of at least one self-reported unsuccessful dietary effort to lose body weight.

Exclusion criteria

Subjects are excluded from the trial if any of the following criteria apply:

Glycemia-related:

- Glycated hemoglobin (HbA_{1c}) ≥48 mmol/mol (6.5%) as measured by the central laboratory at screening.
- History of type 1 or type 2 diabetes mellitus.
- Treatment with glucose-lowering agent(s) within 90 days before screening.
- Treatment with a glucagon-like peptide-1 receptor agonist within 180 days before screening.

Obesity-related:

- A self-reported change in body weight >5 kg (11 lbs) 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 are allowed: (1) liposuction and/or abdominoplasty, if performed >1 year before screening; (2) lap banding, if the band has been removed >1 year before screening; (3) intragastric balloon, if the balloon has been removed >1 year before screening; or (4) duodenal-jejunal bypass sleeve, if the sleeve has 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 (e.g. schizophrenia, bipolar disorder).
- A Patient Health Questionnaire-9 score of ≥15 at screening.
- A 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:

- Use of non-herbal Chinese medicine or other non-herbal local medicine with unknown/unspecified content within 90 days before screening.
- 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 estimated glomerular filtration rate value of <15 mL/min/1.73 m^2 as defined by KDIGO 2012[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 are allowed.
- Any of the following: myocardial infarction, stroke, hospitalization for unstable angina, or transient ischemic attack within the past 60 days prior to screening.
- Subject presently 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 is defined as signed informed consent.
- Participation in another clinical trial within 90 days before screening.
- Other subject(s) from the same household participating in any semaglutide trial.
- Female who is pregnant, breast-feeding, or intends to become pregnant, or is 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 subject's safety or compliance with the protocol.

PATIENT-REPORTED OUTCOMES

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

SF-36 is a generic patient-reported outcome (PRO) instrument measuring health-related quality of life and general health status across disease areas. It consists of 36 questions (items) across eight domains (physical functioning, role limitations due to physical health problems, bodily 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) and mental component summary (MCS), created by aggregating the eight domains according to the scoring algorithm.[2] SF-36 scores are norm-based scores, i.e. transformed to a scale where the 2009 US general population has a mean of 50 and a standard deviation of 10. The lowest to highest scores are 19.03 to 57.60 for the physical functioning domain, 6.11 to 79.67 for the PCS, and –3.83 to 78.75 for the MCS, reported as norm-based scores. An increase in score represents an improvement in health status.

Impact of Weight on Quality of Life-Lite Clinical Trials Version (IWQOL-Lite-CT)

The IWQOL-Lite-CT is a 20-item PRO instrument used to assess weight-related physical and psychosocial functioning in three composite scores (physical, physical function, and psychosocial) and a total score.[3] The range of possible scores for the IWQOL-Lite-CT is 0–100. Larger values on composite scores as well as total scores of the IWQOL-Lite-CT indicate better patient functioning.

ENDPOINTS

Co-primary endpoints

In order of hierarchical testing procedure:

- Change from baseline to week 68 in body weight (%).
- Subjects who after 68 weeks achieve (yes/no) body weight reduction ≥5% from baseline.

Confirmatory secondary endpoints

In order of hierarchical testing procedure:

- Subjects who after 68 weeks achieved (yes/no):
    - Body weight reduction ≥10% from baseline.
    - Body weight reduction ≥15% from baseline.
- Change from baseline to week 68 in:
    - Waist circumference (cm).
    - Systolic blood pressure (mmHg).
    - SF-36 physical functioning score.
    - IWQOL-Lite-CT physical function score.

Supportive secondary endpoints

Efficacy endpoints

- Change from baseline to week 68 in:
    - Body weight (kg) and BMI (kg/m^2).
    - HbA_{1C} (%, 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, triglycerides.
    - C-reactive protein (mg/L).
    - Plasminogen activator inhibitor-1 activity (AU/mL).

  - Soluble leptin receptor (ng/mL) and leptin (ng/mL).
  - SF-36 scores: role-physical, bodily pain, general health, vitality, social functioning, role-emotional, mental health, PCS, MCS.
  - IWQOL-Lite-CT: physical score, psychosocial score, and total score.
- Body composition (assessed by dual energy X-ray absorptiometry [DEXA]) in a subset of participants:
  - Total fat mass (%, kg).
  - Total lean body mass (%, kg).
  - Regional visceral fat mass (%, kg).
- Body weight (%, kg) in the DEXA subset of participants.
- Subjects who after 68 weeks achieved (yes/no):
  - Body weight reduction ≥20% from baseline.
  - Responder definition value for SF-36 physical functioning score and IWQOL-Lite-CT physical function score.

Safety endpoints

- Number of treatment-emergent adverse events from baseline to week 75.
- Number of serious adverse events from baseline to week 75.
- Change from baseline to week 68 in:
  - Pulse (bpm).
  - Amylase (U/L).
  - Lipase (U/L).
  - Calcitonin (ng/L).

Exploratory endpoints

- Change from baseline to week 68 in:
  - Glycemic category (normo-glycemia, prediabetes, type 2 diabetes).
  - Antihypertensive medication (decrease, no change, increase).

- Lipid-lowering medication (decrease, no change, increase).
- The Stanford Presenteeism Scale, total score.
- Fatty liver index score category (<30, ≥30 and <60, ≥60).
- International Consultation on Incontinence Questionnaire-Urinary Incontinence-Short Form, sum score (assessed in female subjects).
- Subjects who from randomization to week 68 discontinued randomized trial product (yes/no).
- Time to permanent discontinuation of randomized trial product (weeks).
- Number of days/week with at least one entry in the food diary from baseline to week 68.
- Number of minutes/week of physical activity from baseline to week 68.

LABORATORY TESTING

The following central laboratory was used for testing in STEP 1:

- ICON Laboratory Services Inc. (all standard laboratory assessments for efficacy and safety, and central storage of biological samples for future analyses of genetic and circulating biomarkers)

The following laboratory assessments were performed:

Laboratory assessment	**Parameter**	**Assay/system**
Glucose metabolism	Fasting plasma glucose	Hexokinase/glucose-6-phosphate dehydrogenase assay (Abbott ARCHITECT System)
	HbA_{1c}	Calculation, ion exchange HPLC (BIO-RAD Variant II Hemoglobin test), boronate affinity chromatography and HPLC (Trinity Biotech Premier Hb9210)
	Fasting serum insulin	CMI (Abbott ARCHITECT System)
Lipids	Cholesterol	Enzymatic assay (Abbott ARCHITECT System)
	HDL-C	Accelerator selective detergent assay (Abbott ARCHITECT System)
	LDL-C	Liquid selective detergent assay (Abbott ARCHITECT System)
	Triglycerides	Glycerol phosphate oxidase assay (Abbott ARCHITECT System)
	VLDL-C	Calculated (Abbott ARCHITECT System)
	Free fatty acids	Enzymatic assay (Abbott ARCHITECT System)
Biomarkers	Plasminogen activator inhibitor-1 activity	Synthetic chromogenic substrate method (STAGO STA Analyzer)
	CRP	Turbidimetric/immunoturbidimetric assay (Abbott ARCHITECT System)
Hematology	Basophils	All assessed by automated cytochemistry/microscopy (Siemens Healthcare Diagnostics)
	Eosinophils	
	Erythrocytes	
	Hematocrit	
	Hemoglobin	
	Leukocytes	
	Lymphocytes	
	Monocytes	
	Neutrophils	
	Thrombocytes	
Biochemistry	ALT	NADH (without P-5'-P) assay (Abbott ARCHITECT System)
	Albumin	Bromcresol green assay (Abbott ARCHITECT System)

	Alkaline phosphatase	Para-nitrophenyl phosphate assay (Abbott ARCHITECT System)
	Amylase	CNPG3 substrate assay (Abbott ARCHITECT System)
	AST	NADH (without P-5'-P) assay (Abbott ARCHITECT System)
	Calcitonin	Chemiluminescent immunometric assay (Siemens Healthcare Diagnostics)
	Creatine kinase	N-acetyl-L-cysteine assay (Abbott ARCHITECT System)
	Creatinine	Enzymatic assay (Abbott ARCHITECT System)
	GGT	L-Gamma-glutamyl-3-carboxy-4-nitroanilide substrate assay (Abbott ARCHITECT System)
	Lipase	Quinone dye assay (Abbott ARCHITECT System)
	Potassium	Ion-selective electrode diluted (indirect) assay (Abbott ARCHITECT System)
	Sodium	Ion-selective electrode diluted (indirect) assay (Abbott ARCHITECT System)
	Thyroid-stimulating hormone	CMI (Abbott ARCHITECT System, Trinity Biotech Premier Hb9210)
	Total bilirubin	Diazonium salt assay (Abbott ARCHITECT System)
	Urea	Urease assay (Abbott ARCHITECT System)
Other tests	eGFR	Calculated using the CKD-EPI creatinine equation as defined by KDIGO 2012[1] (Abbott ARCHITECT System)

ALT, alanine aminotransferase; AST, aspartate aminotransferase; CKD-EPI, Chronic Kidney Disease Epidemiology Collaboration; CMI, chemiluminescent microparticle immunoassay; CNPG3, 2-chloro-4-nitrophenyl-a-D-maltotrioside; CRP, C-reactive protein; eGFR, estimated glomerular filtration rate; GGT, gamma-glutamyl transferase; HbA_{1c}, glycated hemoglobin; HDL-C, high-density lipoprotein cholesterol; HPLC, high-performance liquid chromatography; KDIGO, Kidney Disease Improving Global Outcomes; LDL-C, low-density lipoprotein cholesterol; NADH, nicotinamide adenine dinucleotide; P-5'-P, pyridoxal-5'-phosphate; VLDL-C, very low-density lipoprotein cholesterol

STATISTICAL ANALYSIS

Two estimands were employed to assess treatment efficacy from different perspectives, and accounted for intercurrent events and missing data differently, as described previously.[4] All analyses in the statistical hierarchy were based on the primary treatment policy estimand, which quantified the average treatment effect in all randomized participants regardless of adherence to treatment or starting rescue interventions (anti-obesity medications or bariatric surgery). Continuous endpoints were analyzed using analysis of covariance, with randomized treatment as a factor and baseline endpoint value as a covariate. Categorical endpoints were analyzed using logistic regression, with the same factor and covariate. Missing data were imputed using a multiple imputation approach.[5] The secondary trial product estimand quantified the average treatment effect in all randomized participants assuming they remained on randomized treatment for the entire study duration and without rescue interventions. For the trial product estimand, continuous endpoints were assessed using a mixed model for repeated measurements (MMRM); categorical endpoints were assessed using logistic regression with treatment as the only factor (for missing data, categorization was based on values predicted from an MMRM).

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	Statistical model	Imputation approach	Missing results at week 68, n (%)
Primary endpoints							
Primary	% weight change	1	Continuous	Treatment policy*	ANCOVA	RD-MI	Placebo: 78 (11.9) Semaglutide: 94 (7.2)
				Trial product†	MMRM	-	Placebo: 212 (32.4) Semaglutide: 356 (27.3)
Primary	5% responders	2	Binary	Treatment policy*	LR	RD-MI	Placebo: 78 (11.9) Semaglutide: 94 (7.2)
				Trial product†	LR	MMRM	Placebo: 212 (32.4) Semaglutide: 356 (27.3)
Confirmatory secondary endpoints							
Primary	10% responders	3	Binary	Treatment policy*	LR	RD-MI	Placebo: 78 (11.9) Semaglutide: 94 (7.2)
				Trial product†	LR	MMRM	Placebo: 212 (32.4) Semaglutide: 401 (30.7)
Primary	15% responders	4	Binary	Treatment policy*	LR	RD-MI	Placebo: 78 (11.9) Semaglutide: 94 (7.2)
				Trial product†	LR	MMRM	Placebo: 212 (32.4) Semaglutide: 356 (27.3)
Primary	Waist circumference change (cm)	5	Continuous	Treatment policy*	ANCOVA	RD-MI	Placebo: 80 (12.2) Semaglutide: 96 (7.4)
				Trial product†	MMRM	-	Placebo: 212 (32.4)

							Semaglutide: 356 (27.3)
Secondary	Systolic blood pressure change (mmHg)	6	Continuous	Treatment policy*	ANCOVA	RD-MI	Placebo: 81 (12.4) Semaglutide: 96 (7.4)
				Trial product†	MMRM	-	Placebo: 212 (32.4) Semaglutide: 356 (27.3)
Secondary	SF-36 physical functioning score change	7	Continuous	Treatment policy*	ANCOVA	RD-MI	Placebo: 86 (13.1) Semaglutide: 103 (7.9)
				Trial product†	MMRM	-	Placebo: 216 (33.0) Semaglutide: 364 (27.9)
Secondary	IWQOL-Lite-CT physical function score change	8	Continuous	Treatment policy*	ANCOVA	RD-MI	Placebo: 86 (13.1) Semaglutide: 105 (8.0)
				Trial product†	MMRM	-	Placebo: 216 (33.0) Semaglutide: 366 (28.0)

*Designated as the primary estimand.

†Designated as the secondary estimand.

ANCOVA, analysis of covariance; FAS, full analysis set; IWQOL-Lite-CT, Impact of Weight on Quality of Life-Lite Clinical Trials Version; LR, logistic regression; MMRM, mixed model for repeated measurements; RD-MI, multiple imputation using retrieved subjects; SF-36, Short Form36v2® Health Survey, Acute Version.

Test order refers to the order of the endpoint in the statistical test hierarchy. All analyses were performed using the full analysis set.

See Section 2.3 in the Statistical Analysis Plan for a description of imputation methods.

SUPPLEMENTARY TABLES

Table S1. Baseline demographics and clinical characteristics in the DEXA subpopulation

	DEXA subpopulation			
	Semaglutide 2.4 mg once weekly (N=95)	**Placebo once weekly (N=45)**	**Total (N=140)**	**Overall study population (N=1961)***
Age – years	50 ± 12	52 ± 13	51 ± 12	46 ± 13
Female sex – n (%)	72 (75.8)	34 (75.6)	106 (75.7)	1453 (74.1)
Race – n (%)				
White	75 (78.9)	41 (91.1)	116 (82.9)	1472 (75.1)
Black or African American	18 (18.9)	3 (6.7)	21 (15.0)	111 (5.7)
Asian	1 (1.1)	1 (2.2)	2 (1.4)	261 (13.3)
Other†	1 (1.1)	0	1 (0.7)	117 (6.0)
Hispanic or Latino ethnic group – n (%)	2 (2.1)	6 (13.3)	8 (5.7)	236 (12.0)
Body weight – kg	98.3 ± 15.9	98.7 ± 12.1	98.4 ± 14.7	105.3 ± 21.9
BMI				
Mean – kg/m^2	34.8 ± 3.6	35.0 ± 3.6	34.8 ± 3.6	37.9 ± 6.7
<30 kg/m^2 – n (%)	7 (7.4)	4 (8.9)	11 (7.9)	117 (6.0)
≥30 – <35 kg/m^2 – n (%)	41 (43.2)	17 (37.8)	58 (41.4)	643 (32.8)
≥35 – <40 kg/m^2 – n (%)	43 (45.3)	23 (51.1)	66 (47.1)	614 (31.3)
≥40 kg/m^2 – n (%)	4 (4.2)	1 (2.2)	5 (3.6)	587 (29.9)
Waist circumference – cm	109.4 ± 10.6	111.0 ± 10.1	109.9 ± 10.4	114.7 ± 14.6
Glycated hemoglobin – %	5.7 ± 0.4	5.7 ± 0.3	5.7 ± 0.3	5.7 ± 0.3
Body composition (DEXA)				
Total fat mass‡				
Kg	42.1 ± 10.1	43.3 ± 9.2	42.5 ± 9.8	–
%	43.4 ± 7.5	44.6 ± 8.1	43.8 ± 7.7	–
Regional visceral fat mass§				
Kg	1.3 ± 0.6	1.5 ± 0.7	1.3 ± 0.6	–
%	33.8 ± 9.9	36.3 ± 12.3	34.6 ± 10.7	–
Total lean body mass‡				
Kg	52.4 ± 11.6	51.5 ± 10.8	52.1 ± 11.3	–
%	53.9 ± 7.4	52.7 ± 7.7	53.5 ± 7.5	–

All data presented as mean ± standard deviation, unless indicated otherwise.

*Overall study population data included for comparative purposes. Additional baseline demographic and clinical characteristics for the overall study population are described in Table 1 in the main manuscript.

†Including American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Other and Not Applicable, the latter of which is how race was recorded in France.

‡Percentages calculated for total fat mass and lean body mass as the respective values divided by total body mass.

§Visceral fat mass was calculated in the L4 region (males or females), android region (males only), or in the gynoid region (females only), depending on the methodology of the scanner available at participating study sites. Percentage visceral fat mass values are relative to the area assessed, not total body mass.

BMI, body mass index; DEXA, dual energy X-ray absorptiometry.

Table S2. Co-primary, Confirmatory and Selected Supportive Secondary Endpoints for the Trial Product Estimand*†

	Semaglutide 2.4 mg once weekly (N=1306)	Placebo once weekly (N=655)	Treatment comparison for semaglutide vs. placebo [95% CI]
Co-primary endpoint assessed in the overall population			
Body weight change from baseline to week 68 – %	–16.86	–2.44	ETD: –14.42 [–15.29; –13.55]
Body weight reduction ≥5% – proportion of participants (%) at week 68	92.4	33.1	OR: 37.0 [28.0; 49.0]
Confirmatory secondary endpoints assessed in the overall population			
Body weight reduction ≥10% – proportion of participants (%) at week 68	74.8	11.8	OR: 30.0 [22.5; 40.0]
Body weight reduction ≥15% – proportion of participants (%) at week 68	54.8	5.0	OR: 31.8 [21.0; 48.3]
Waist circumference change from baseline to week 68 – cm	–15.22	–4.48	ETD: –10.75 [–11.61; –9.88]
Systolic blood pressure change from baseline to week 68 – mmHg	–7.08	–1.14	ETD: –5.93 [–7.19; –4.68]
SF-36 physical functioning score change from baseline to week 68	2.56	0.50	ETD: 2.06 [1.43; 2.70]
IWQOL-Lite-CT physical function score change from baseline to week 68	16.08	6.51	ETD: 9.57 [7.71; 11.44]
Selected supportive secondary endpoints assessed in the overall population			
Body weight reduction ≥20% – proportion of participants (%) at week 68	34.8	2.0	OR: 42.2 [20.8; 85.6]
Body weight change from baseline to week 68 – kg	–17.4	–2.7	ETD: –14.7 [–15.6; –13.7]
BMI change from baseline to week 68 – kg/m^2	–6.27	–0.95	ETD: –5.33 [–5.65; –5.00]
Glycated hemoglobin change from baseline to week 68 – percentage-points	–0.50	–0.16	ETD: –0.34 [–0.37; –0.31]
Fasting plasma glucose change from baseline to week 68 – mg/dL	–9.90	–1.00	ETD: –8.90 [–9.84; –7.96]
Diastolic blood pressure change from baseline to week 68 – mmHg	–2.99	–0.59	ETD: –2.40 [–3.28; –1.52]
Lipids ratio to baseline at week 68‡			
Total cholesterol	0.96	1.00	ETR: 0.96 [0.94; 0.97]
HDL cholesterol	1.05	1.02	ETR: 1.03 [1.02; 1.05]
LDL cholesterol	0.96	1.01	ETR: 0.95 [0.93; 0.98]
VLDL cholesterol	0.76	0.92	ETR: 0.82 [0.79; 0.85]
Free fatty acids	0.81	0.92	ETR: 0.88 [0.82; 0.94]
Triglycerides	0.76	0.92	ETR: 0.82 [0.79; 0.86]
CRP ratio to baseline at week 68‡	0.42	0.80	ETR: 0.52 [0.47; 0.57]

Supportive secondary endpoints assessed in the DEXA subpopulation

	N=95	N=45	
Body composition change from baseline to week 68 (DEXA)			
Total fat mass			
Kg change	–10.40	–1.17	ETD: –9.23 [–12.72; –5.74]
Percentage-points change in total fat mass proportion[§]	–4.19	–0.19	ETD: –4.00 [–6.27; –1.73]
Regional visceral fat mass[¶]			
Kg change	–0.47	–0.03	ETD: –0.45 [–0.60; –0.30]
Percentage-points change in regional visceral fat mass proportion[‖]	–2.65	0.58	ETD: –3.23 [–5.35; –1.10]
Total lean body mass			
Kg change	–6.92	–1.48	ETD: –5.44 [–7.07; –3.81]
Percentage-points change in total lean body mass proportion[§]	3.61	0.11	ETD: 3.50 [1.35; 5.64]

*The trial product estimand assesses treatment effect if trial product was taken as intended (i.e. if all participants adhered to treatment and did not receive rescue intervention). Treatment policy estimand data are reported in Table 2. Denominators for the percentages of participants observed to have body-weight reduction of ≥5%, ≥10%, ≥15%, and ≥20% at week 68 are the numbers of participants for whom data were available from the week 68 visit — 1059 participants in the semaglutide group and 499 participants in the placebo group.

†All analyses in the statistical hierarchy were based on the primary treatment policy estimand and P values are therefore not reported.

‡Data presented as ratio to baseline and estimated treatment ratio (ratios to baseline and corresponding baseline values were log-transformed prior to analysis).

§Percentage-point changes in total fat mass and lean body mass proportions, which are calculated as the respective values divided by total body mass.

¶Visceral fat mass was calculated in the L4 region (males or females), android region (males only), or in the gynoid region (females only), depending on the methodology of the scanner available at participating study sites.

‖Percentage-point changes in visceral fat mass proportions, which are the visceral fat mass relative to the area assessed, not total body mass.

Endpoints were analyzed using a mixed model for repeated measurements.

BMI, body mass index; CI, confidence interval; CRP, C-reactive protein; DEXA, dual energy X-ray absorptiometry; ETD, estimated treatment difference; ETR, estimated treatment ratio; HDL, high-density lipoprotein; IWQOL-Lite-CT, Impact of Weight on Quality of Life-Lite Clinical Trials Version; LDL, low-density lipoprotein; OR, odds ratio; SF-36, Short Form36v2® Health Survey, Acute Version; VLDL, very-low-density lipoprotein.

Table S3. Selected Supportive Secondary and Exploratory Endpoints for the Treatment Policy Estimand*

	Semaglutide 2.4 mg once weekly	Placebo once weekly	Treatment comparison for semaglutide vs. placebo [95% CI]
	Overall population		
	N=1306	N=655	
Fasting serum insulin – pmol/L, ratio to baseline at week 68	0.74	0.93	ETR: 0.79 [0.74, 0.83]
Alanine aminotransferase – ratio to baseline at week 68†	0.76	0.94	ETR: 0.81 [0.77, 0.86]
Aspartate aminotransferase – ratio to baseline at week 68†	0.89	0.99	ETR: 0.90 [0.88, 0.93]
Antihypertensive medication – %			
Decreased	14	5	
No change	53	62	
Increased	12	22	
Stopped	20	11	
Lipid lowering medication – %			
Decreased	4	5	
No change	69	63	
Increased	10	20	
Stopped	17	12	

*The treatment policy estimand assesses treatment effect regardless of treatment discontinuation or rescue intervention; see Supplementary Appendix, Table S4 for corresponding data for the trial product estimand (assesses treatment effect assuming all participants adhered to treatment and did not receive rescue intervention). Supportive secondary endpoint analyses were not adjusted for multiplicity and P values are therefore not reported for these endpoints.

†Not a prespecified endpoint.

CI, confidence interval; ETR, estimated treatment ratio.

Table S4. Proportions of Participants Achieving a Clinically Meaningful Within-person Improvement in Score from Baseline to Week 68 (%), Treatment Policy Estimand*†‡

	Semaglutide 2.4 mg once weekly	Placebo once weekly	Treatment comparison for semaglutide vs. placebo [95% CI]
SF-36 physical functioning (≥3.7 points) ‡	40.0	27.0	OR: 2.08 [1.60, 2.70]
IWQOL-Lite-CT physical function (≥14.6 points) ‡	51.2	32.9	OR: 2.72 [2.14, 3.47]

SF-36 scores are norm-based scores (transformed to a scale where the 2009 US general population has a mean of 50 and a standard deviation of 10).

*The treatment policy estimand assesses treatment effect regardless of treatment discontinuation or rescue intervention; see Supplementary Appendix, Table S4 for corresponding data for the trial product estimand (assesses treatment effect assuming all participants adhered to treatment and did not receive rescue intervention). Supportive secondary endpoint analyses were not adjusted for multiplicity and P values are therefore not reported for these endpoints.

†Not a prespecified endpoint.

‡Threshold values for clinically meaningful within-patient improvements (responder thresholds) are anchor-based obesity-specific thresholds.

CI, confidence interval; IWQOL-Lite-CT, Impact of Weight on Quality of Life-Lite Clinical Trials Version; OR, odds ratio; SF-36, SF-36v2® Health Survey acute version.

Table S5. Supportive Secondary Endpoints Assessed in the DEXA Subpopulation for the Treatment Policy Estimand*

	Semaglutide 2.4 mg once weekly	Placebo once weekly	Treatment comparison for semaglutide vs. placebo [95% CI]
	N=95	N=45	
Body composition change from baseline to week 68 (DEXA)			
Total fat mass			
Kg change	–8.36	–1.37	ETD: –6.99 [–9.79; –4.19]
Percentage-points change in total fat mass proportion†	–3.48	–0.19	ETD: –3.29 [–4.94; –1.65]
Regional visceral fat mass‡			
Kg change	–0.36	–0.10	ETD: –0.27 [–0.39; –0.15]
Percentage-points change in regional visceral fat mass proportion§	–1.99	–0.01	ETD: –1.98 [–3.69; –0.27]
Total lean body mass			
Kg change	–5.26	–1.83	ETD: –3.43 [–4.74; –2.13]
Percentage-points change in total lean body mass proportion†	3.04	0.09	ETD: 2.94 [1.40; 4.49]

*The treatment policy estimand assesses treatment effect regardless of treatment discontinuation or rescue intervention; see Supplementary Appendix, Table S4 for corresponding data for the trial product estimand (assesses treatment effect assuming all participants adhered to treatment and did not receive rescue intervention). Supportive secondary endpoint analyses were not adjusted for multiplicity and P values are therefore not reported for these endpoints.

†Percentage-point changes in total fat mass and total lean body mass proportions, which are calculated as the respective values divided by total body mass.

‡Visceral fat mass was calculated in the L4 region (males or females), android region (males only), or in the gynoid region (females only), depending on the methodology of the scanner available at participating study sites.

§Percentage-point changes in visceral fat mass proportions, which are the visceral fat mass relative to the area assessed, not total body mass.

Continuous endpoints were 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.[5] Categorical endpoints were analyzed using logistic regression, with the same factor and covariate.

CI, confidence interval; DEXA, dual energy X-ray absorptiometry; ETD, estimated treatment difference.

Table S6. Supportive Secondary Safety Endpoints, On-treatment*

	Semaglutide 2.4 mg once weekly		Placebo once weekly	
	N	**Mean**	**N**	**Mean**
Pulse – bpm				
Baseline	1306	72 ± 10	655	72 ± 10
Week 68	1059	75 ± 9	499	71 ± 10
Change from baseline to week 68†	1306	3.52	655	–0.74
Estimated treatment difference (semaglutide vs. placebo) [95% CI]†		4.26 [3.38; 5.15]		
Amylase – U/L				
Baseline	1306	48 (35.7)	655	48 (35.5)
Week 68	1053	55 (37.3)	497	49 (35.9)
Ratio to baseline at week 68	1053	1.14 (21.6)	497	1.03 (21.4)
Lipase – U/L				
Baseline	1306	25 (53.8)	654	25 (52.6)
Week 68	1053	36 (59.4)	497	24 (51.4)
Ratio to baseline at week 68	1053	1.41 (49.3)	496	0.97 (37.3)
Calcitonin – ng/L				
Baseline	1306	1.4 (77.3)	655	1.3 (77.4)
Week 68	1050	1.4 (77.5)	497	1.3 (70.8)
Ratio to baseline at week 68	1050	0.99 (37.6)	497	0.95 (40.9)

Data are descriptive statistics presented as arithmetic mean ± standard deviation or geometric mean (coefficient of variation), unless indicated otherwise.

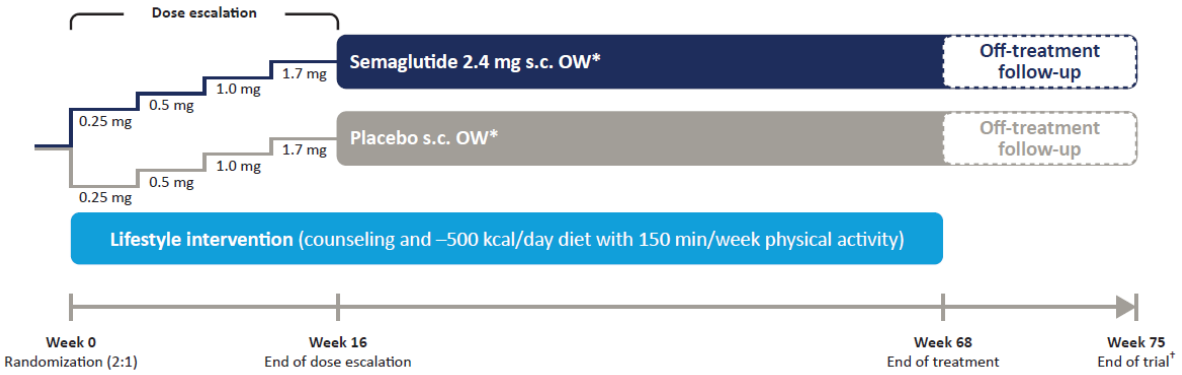
*During treatment with trial product (any dose of trial medication administered within the previous 2 weeks [i.e. any period of temporary treatment interruption with trial product was excluded]).

†Trial product estimand data (assesses treatment effect if trial product was taken as intended [i.e. if all participants adhered to treatment and did not receive rescue intervention]) analyzed using a mixed model for repeated measurements.

CI, confidence interval.

SUPPLEMENTARY FIGURES

Figure S1. Trial Design

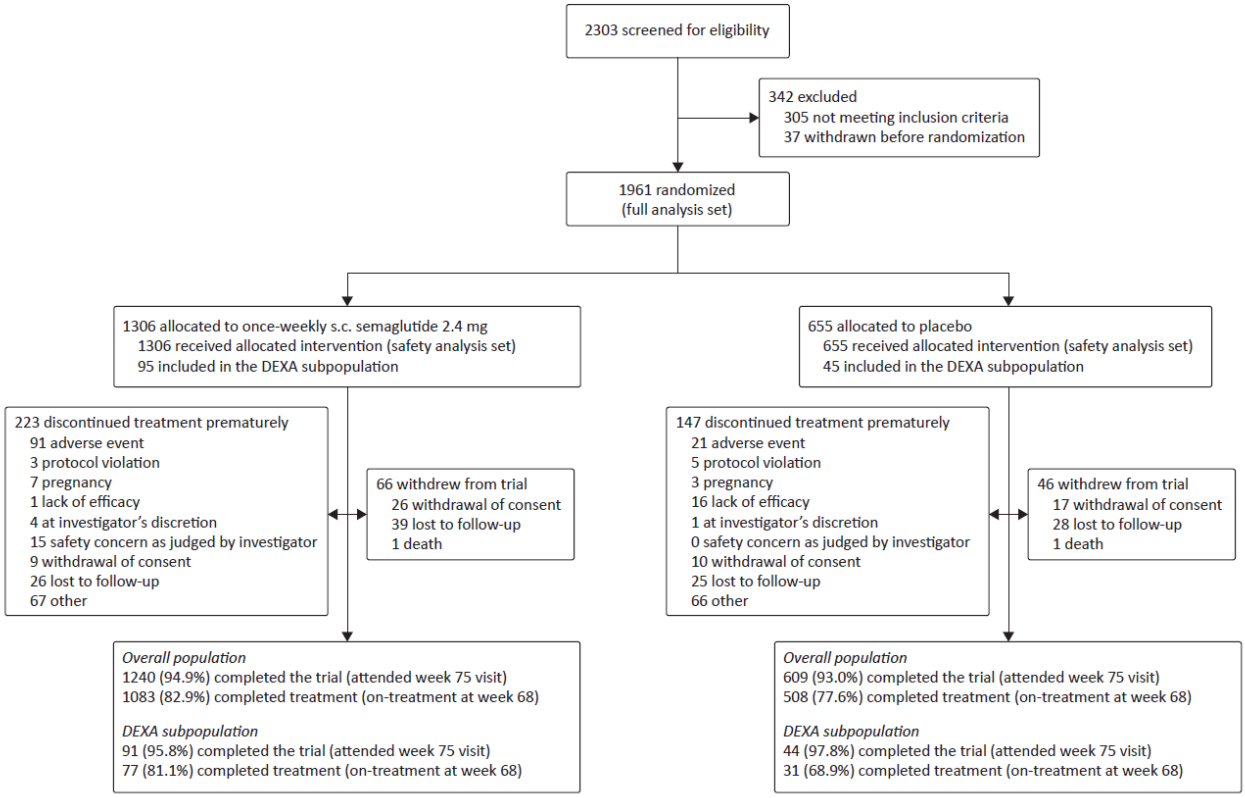


*As an adjunct to lifestyle intervention (–500 kcal/day diet with 150 min/week physical activity).

†End of trial for the main phase.

OW, once weekly; s.c., subcutaneous.

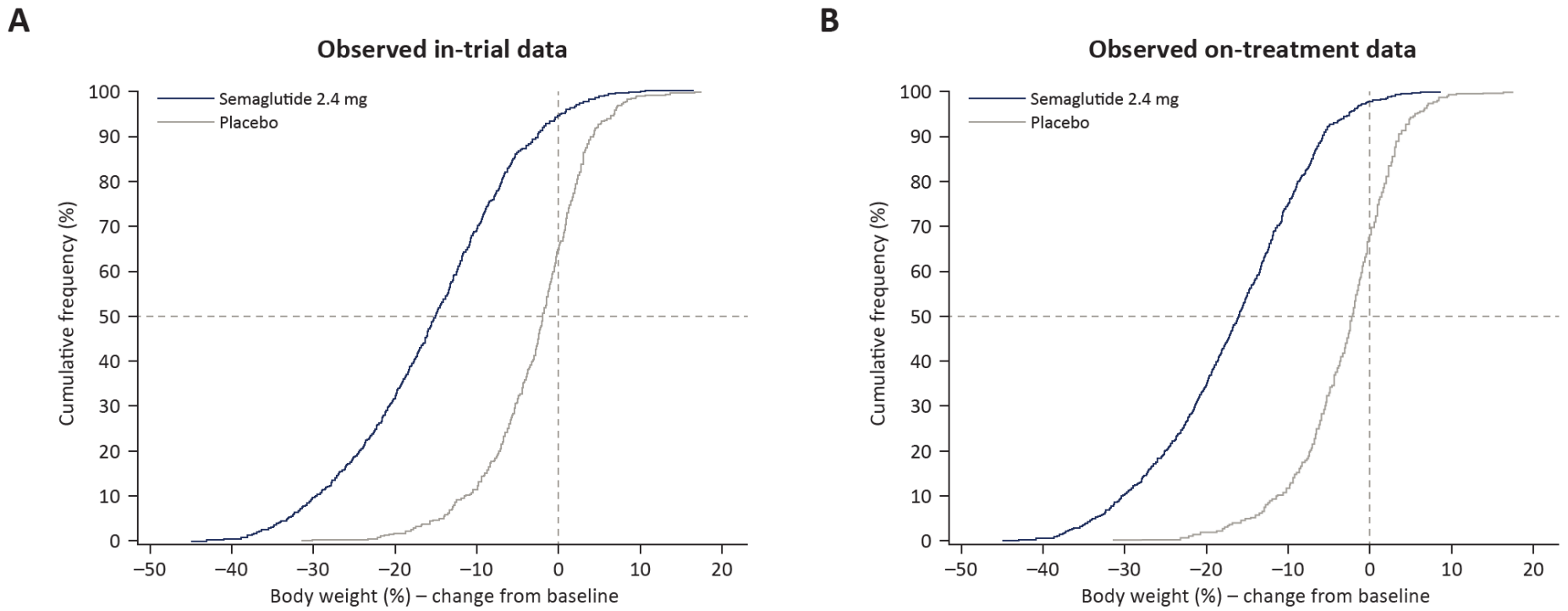
Figure S2. Participant Flow



Among treatment completers in the semaglutide group, 89.6% were receiving the 2.4 mg maintenance dose at week 68, 4.4% were receiving a dose between 1.7 mg and <2.4 mg, and 5.2% were receiving a semaglutide dose <1.7mg; the remainder did not have a dose reported at this timepoint. Among treatment completers in the placebo group, 98.0% completed treatment with the placebo equivalent of the semaglutide 2.4 mg dose; the remainder were on a lower dose or did not have a dose reported at this timepoint.

DEXA, dual energy X-ray absorptiometry; s.c., subcutaneous.

Figure S3. Cumulative Distribution Plot of Change from Baseline to Week 68 in Body Weight

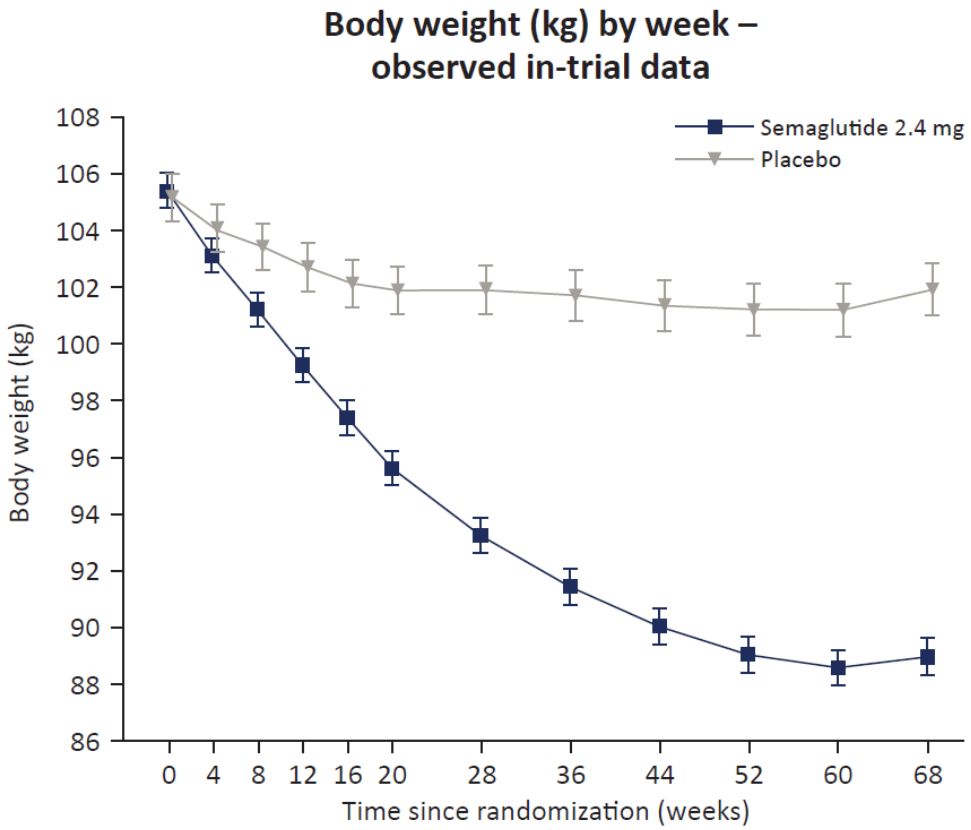


Cumulative distribution plot of observed percentage change from baseline over time in body weight for participants in the full analysis set during the in-trial observation period* (A) and on-treatment observation period† (B).

*From randomization to last contact with trial site, regardless of treatment discontinuation or rescue intervention.

†During treatment with trial product (any dose of trial medication administered within the previous 2 weeks [i.e. any period of temporary treatment interruption with trial product was excluded]).

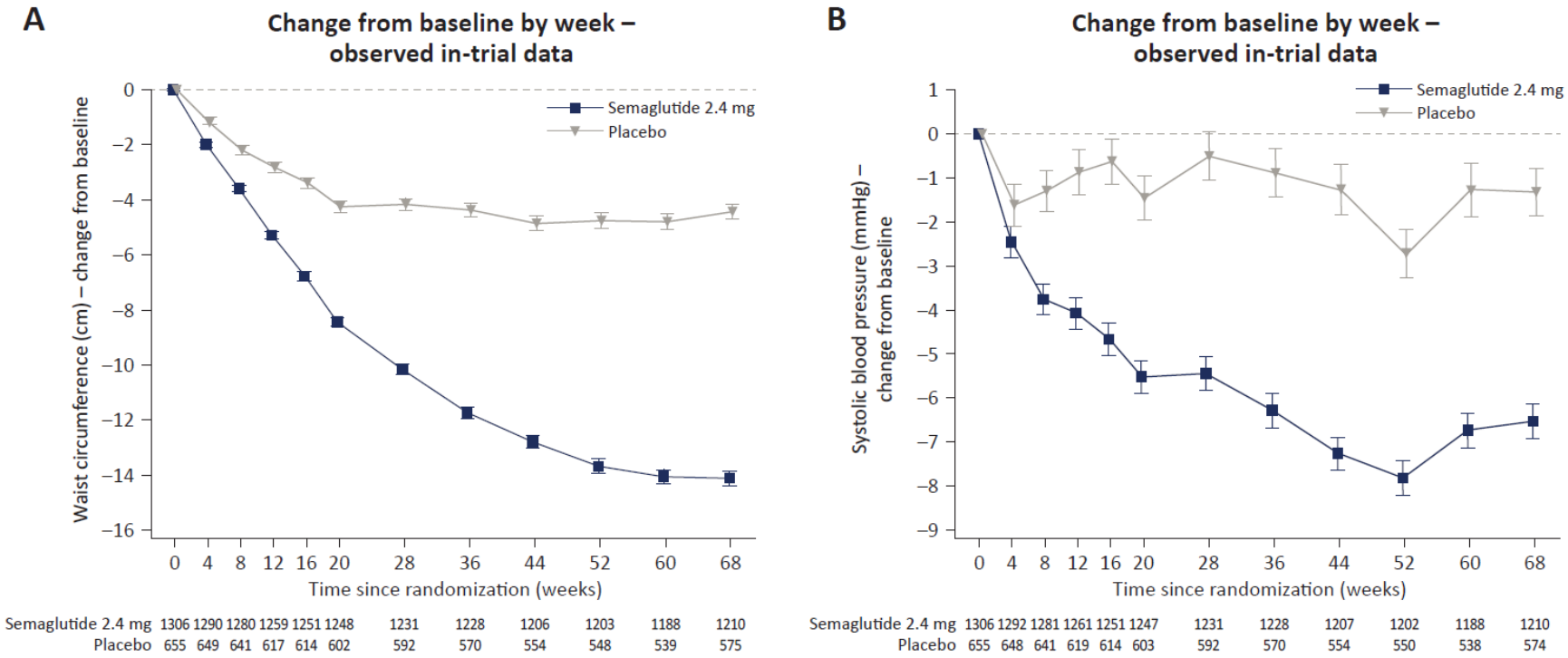
Figure S4. Body Weight in Kilograms by Week



Observed mean body weight (kg) over time for participants in the full analysis set during the in-trial observation period.* Error bars are ± standard error of the mean. N numbers represent the number of participants with available data contributing to the means at each visit.

*From randomization to last contact with trial site, regardless of treatment discontinuation or rescue intervention.

Figure S5. Semaglutide 2.4 mg Once Weekly Compared with Placebo on Selected Confirmatory Secondary Endpoints*

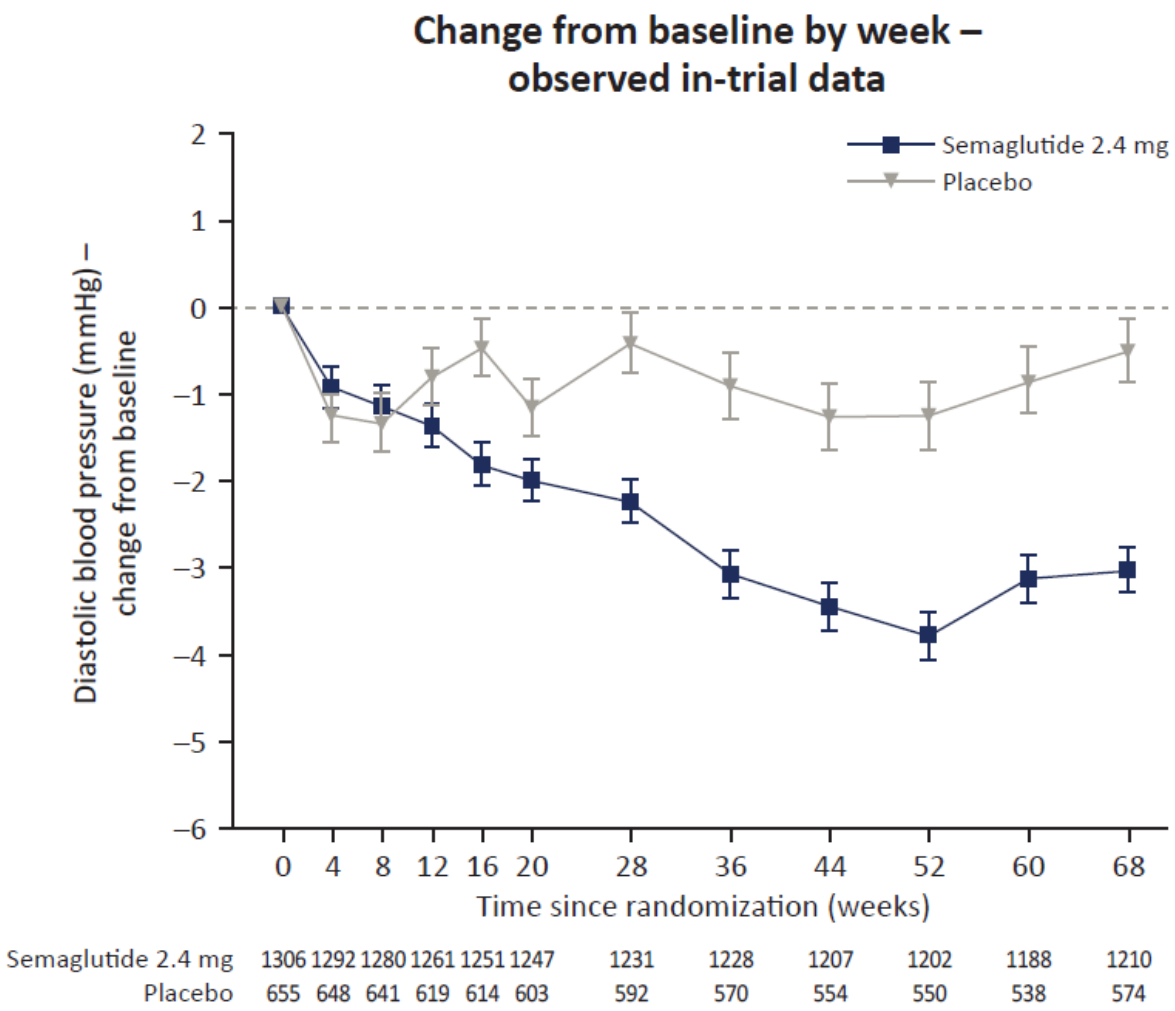


Observed mean change from baseline over time in waist circumference (A), and systolic blood pressure (B) for participants in the full analysis set during the in-trial observation period.† Error bars are ± standard error of the mean. N numbers represent the number of participants with available data contributing to the means at each visit.

*The secondary confirmatory endpoints of achievement of weight loss ≥10% and ≥15% are reported in the results text and in Figure 1; secondary confirmatory endpoints of SF-36 and IWQOL-Lite-CT are reported in Figure S7.

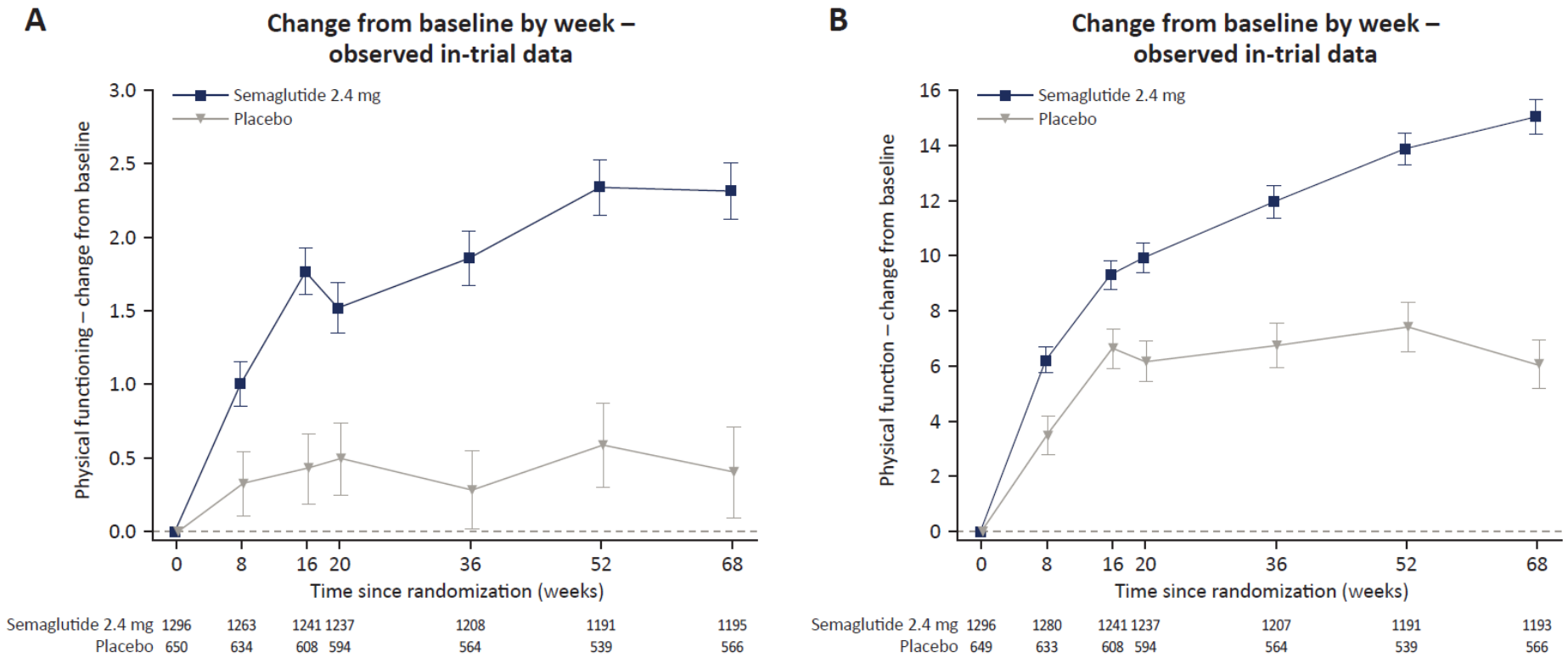
†From randomization to last contact with trial site, regardless of treatment discontinuation or rescue intervention.

Figure S6. Change from Baseline by Week in Diastolic Blood Pressure

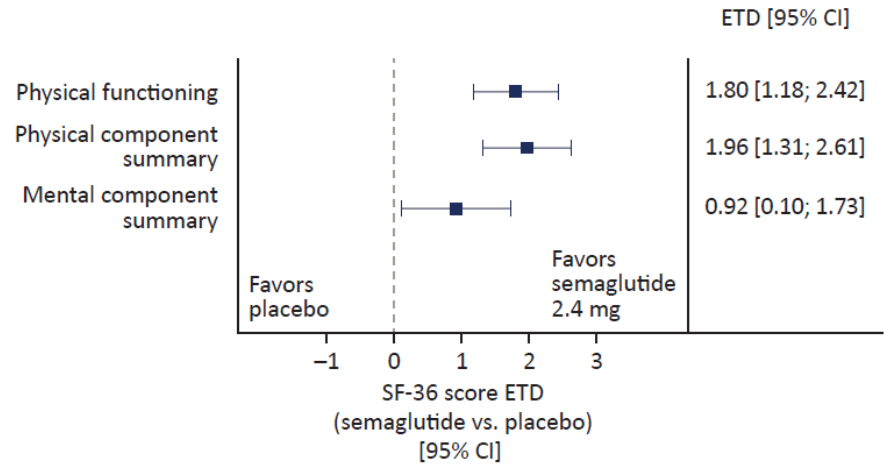


Data presented as observed mean change from baseline over time in diastolic blood pressure in the full analysis set during the in-trial observation period (from randomization to last contact with trial site, regardless of treatment discontinuation or rescue intervention). Error bars are ± standard error of the mean. N numbers represent the number of participants with available data contributing to the means at each visit.

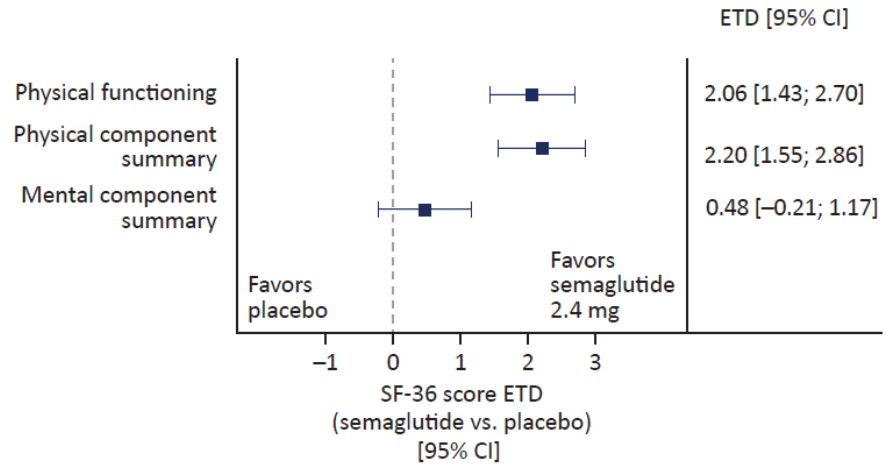
Figure S7. Semaglutide 2.4 mg Once Weekly Compared with Placebo on Patient-reported Outcomes for the SF-36 and IWQOL-Lite-CT



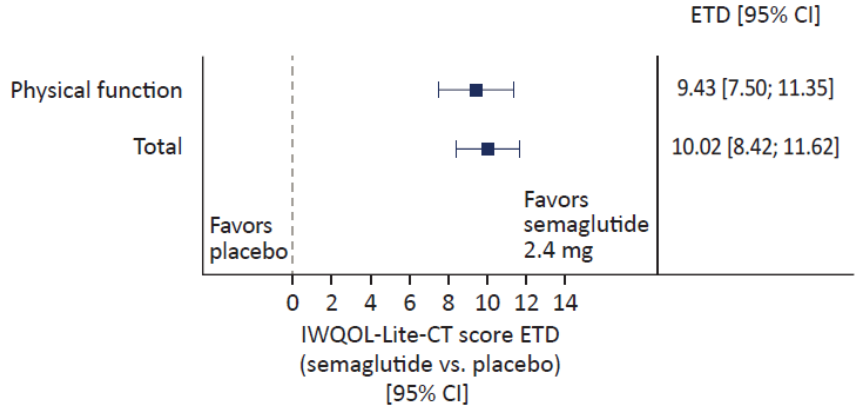
C **SF-36 change from baseline to week 68 – treatment policy estimand**



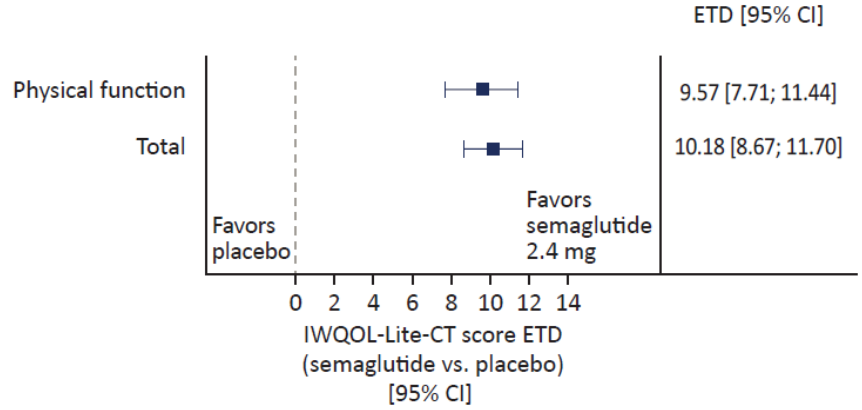
D **SF-36 change from baseline to week 68 – trial product estimand**



E **IWQOL-Lite-CT change from baseline to week 68 – treatment policy estimand**



F **IWQOL-Lite-CT change from baseline to week 68 – trial product estimand**



Observed mean change from baseline over time in SF-36 physical functioning score (A), and IWQOL-Lite-CT physical function score (B) for participants in the full analysis set during the in-trial observation period.† Error bars are ± standard error of the mean. N numbers represent the number of participants with available data contributing to the means at each visit.

†From randomization to last contact with trial site, regardless of treatment discontinuation or rescue intervention.

Panels (C–F) are data presented as estimated treatment differences for semaglutide vs. placebo (boxes) and associated 95% CIs (whiskers) for participants in the full analysis set based on the treatment policy estimand‡ for (C) and (E), and the trial product estimand$ for (D) and (F). SF-36 scores are norm-based scores, which were transformed to a scale where the 2009 US general population has a mean of 50 and a standard deviation of 10.

‡Assesses treatment effect regardless of treatment discontinuation or rescue intervention. Endpoints were 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.[5]

$Assesses treatment effect if trial product was taken as intended (i.e. if all participants adhered to treatment and did not receive rescue intervention). Endpoints were analyzed using a mixed model for repeated measurements.

CI, confidence interval; ETD, estimated treatment difference; IWQOL-Lite-CT, Impact of Weight on Quality of Life-Lite Clinical Trials Version; SF-36, Short Form36v2® Health Survey, Acute Version.

Figure S8. Prevalence and Duration of Gastrointestinal Events by Severity

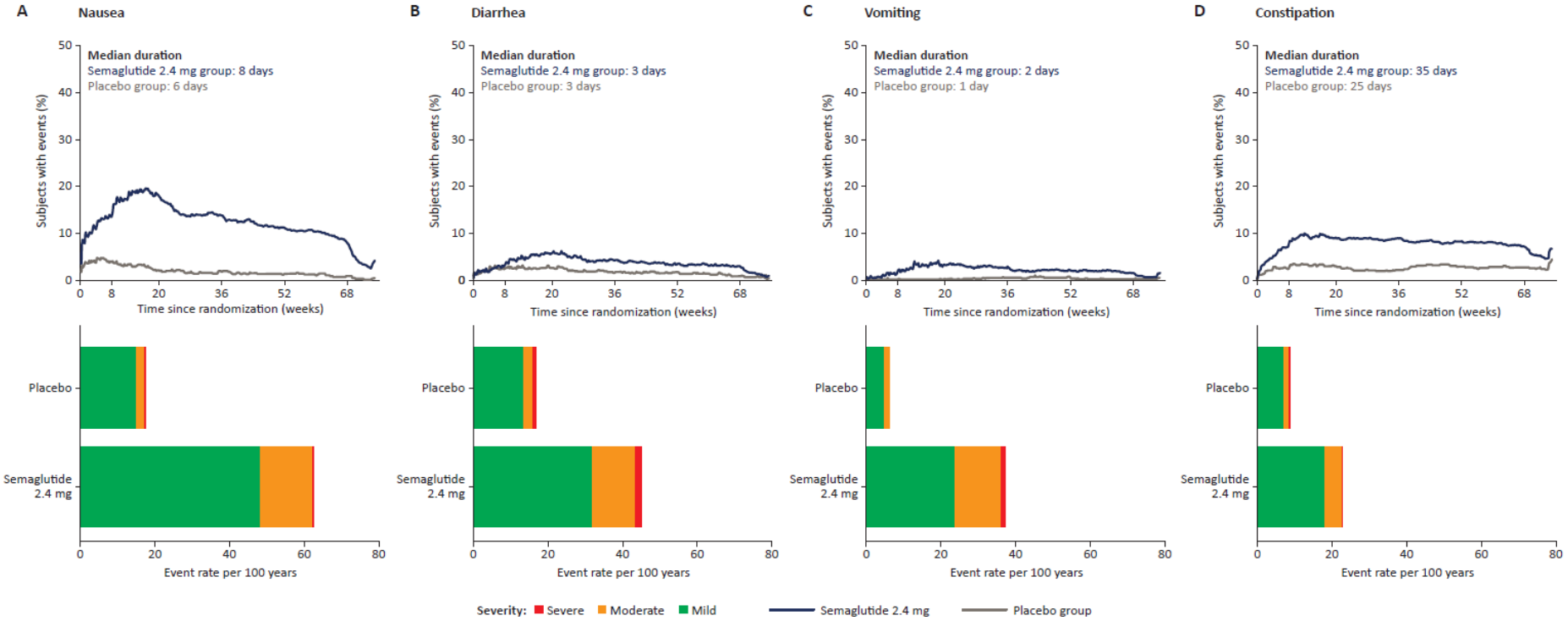
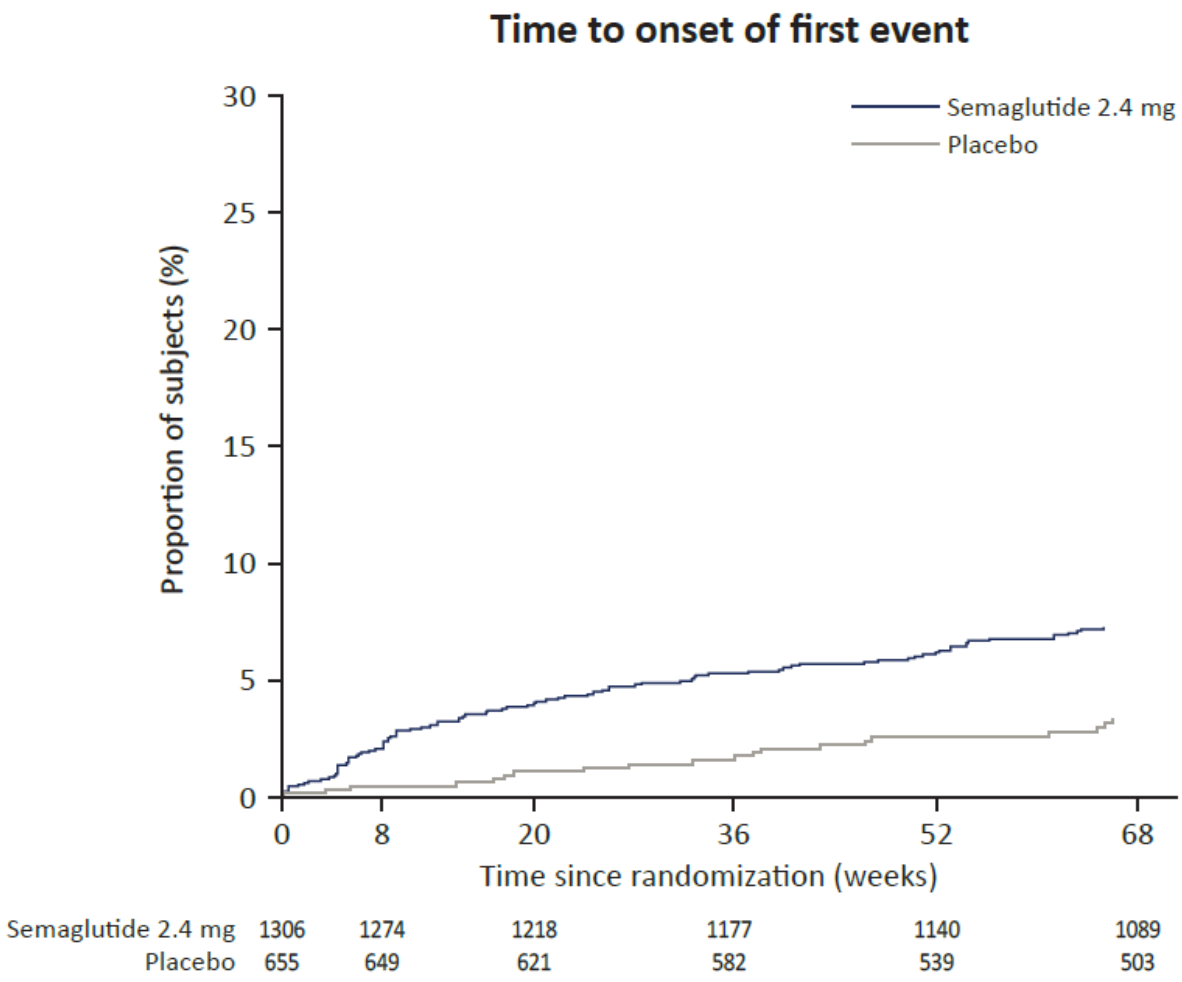


Figure presents the proportion of participants receiving semaglutide or placebo who reported nausea (A), diarrhea (B), vomiting (C), or constipation (D) events classed as mild, moderate, or severe, over the course of the treatment period and the median duration of the event. Data are on-treatment observation period data (during treatment with trial product [any dose of trial medication administered within the previous 49 days (i.e. any period of temporary treatment interruption with trial product was excluded)]). Adverse events were classified by severity as mild (easily tolerated, causing minimal

discomfort and not interfering with everyday activities), moderate (causes sufficient discomfort and interferes with normal everyday activities) or severe (prevents normal everyday activities).

Figure S9. Time to First Onset of Adverse Events Leading to Permanent Trial Product Discontinuation



Data are on-treatment observation period data (during treatment with trial product [any dose of trial medication administered within the previous 49 days (i.e. any period of temporary treatment interruption with trial product was excluded)]).

REFERENCES

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;3:1-150.

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):903-12.

3. Kolotkin RL, Williams VSL, Ervin CM, et al. Validation of a new measure of quality of life in obesity trials: Impact of Weight on Quality of Life-Lite Clinical Trials version. Clin Obes 2019;9:e12310.

4. Wharton S, Astrup A, Endahl L, et al. Estimating and interpreting treatment effects in clinical trials for weight management: implications of estimands, intercurrent events and missing data. Int J Obes 2020; In press.

5. 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;6:1050-61.