Received: 21 December 2022 | Revised: 5 April 2023 | Accepted: 7 April 2023

DOI: 10.1111/dom.15090

ORIGINAL ARTICLE

WILEY

The improved health utility of once-weekly subcutaneous semaglutide 2.4 mg compared with placebo in the STEP 1-4 obesity trials

Jakob Bue Bjorner MD[1,2] | Sara Larsen MSc[3] | Christopher Lübker PhD[3] | Thomas Holst-Hansen PhD[3]

[1]QualityMetric Inc., LLC, Johnston, Rhode Island, USA

[2]Department of Public Health, University of Copenhagen, Copenhagen, Denmark

[3]Novo Nordisk A/S, Søborg, Denmark

Correspondence
Jakob Bue Bjorner, MD, 1301 Atwood Avenue, Suite 216E, Johnston, RI 02919, USA.
Email: jbjorner@qualitymetric.com

Funding information
Novo Nordisk A/S

Abstract

Aim: To assess health utility values in the Semaglutide Treatment Effect in People with obesity (STEP) 1-4 trials.

Materials and Methods: The STEP 1-4 phase 3a, 68-week, double-blind randomized controlled trials assessed the efficacy and safety of semaglutide 2.4 mg versus placebo in individuals with a body mass index (BMI) of 30 kg/m^2 or higher or a BMI of 27 kg/m^2 or higher and at least one comorbidity (STEP 1, 3 and 4), or a BMI of 27 kg/m^2 or higher and type 2 diabetes (STEP 2). Patients received lifestyle intervention plus intensive behavioural therapy in STEP 3. Health-related quality of life was assessed using the Short Form 36-item Health Survey version 2 (SF-36v2) at baseline and week 68. Scores were converted into Short Form Six-Dimension version 2 (SF-6Dv2) utility scores or mapped onto the European Quality of Life Five-Dimension Three-Level (EQ-5D-3L) utility index using UK health utility weights.

Results: At week 68, semaglutide 2.4 mg was associated with minor health utility score improvements from baseline (all trials), while scores for placebo typically decreased. SF-6Dv2 treatment differences by week 68 for semaglutide 2.4 mg versus placebo were significant in STEP 1 and 4 ($P \leq .001$), but not STEP 2 or 3. EQ-5D-3L treatment differences by week 68 for semaglutide 2.4 mg versus placebo were significant in STEP 1, 2 and 4 ($P < .001$ for all), but not STEP 3.

Conclusions: Semaglutide 2.4 mg was associated with improvement in health utility scores compared with placebo, reaching statistical significance in STEP 1, 2 and 4.

KEYWORDS
antiobesity drug, patient-reported outcomes, pharmaco-economics, randomized trial

1 | INTRODUCTION

Obesity is a complex chronic disease with interacting genetic, environmental and behavioural factors.[1] People living with obesity have an increased risk of complications, including type 2 diabetes (T2D), cardiovascular disease, dyslipidaemia, obstructive sleep apnoea, osteoarthritis, depression and impaired quality of life (QoL), among others.[1,2] Sustained weight loss has been shown to reduce the likelihood of obesity-related

complications, improve health-related quality of life (HRQoL) and reduce health care utilization and costs.[3-5] Once-weekly subcutaneous (s.c.) semaglutide 2.4 mg is a glucagon-like peptide-1 analogue approved for use in major geographies including the United States, the UK, the European Union and Canada. Semaglutide 2.4 mg is used as an adjunct to a reduced calorie diet and increased physical activity for chronic weight management in people with obesity (body mass index [BMI] ≥ 30 kg/m^2), or overweight (BMI ≥ 27 kg/m^2) with weight-related comorbidities (e.g. hypertension, T2D or dyslipidaemia).[6-10] The safety and efficacy of once-weekly s.c. semaglutide 2.4 mg was shown in the global phase 3a Semaglutide Treatment Effect in People with obesity (STEP) programme. In the STEP trials, people with obesity treated with semaglutide 2.4 mg achieved a weight loss of 15% to 17% sustained over 68 weeks in people with overweight and obesity without T2D.[11-13]

Patient-reported outcomes (PROs) were assessed throughout the trials using the Short Form 36-item Health Survey version 2 (SF-36v2). In health economic assessments, HRQoL is often summarized using a single 'health utility' value with scores anchored at 0 (equivalent to death) and 1 (perfect health).[14] The cost–benefit of treatment can be calculated by combining health utility values with a time horizon to form quality-adjusted life year (QALY) estimates.[14] These analyses can be used to inform policy based on cost-effectiveness by comparing the incremental costs of an intervention with the incremental gain in QALYs.

There are many generic instruments for estimating health utilities; among the most frequently used are the European Quality of Life Five-Dimension (EQ-5D) and the Short Form Six-Dimension (SF-6D) utility indices.[15,16] Here, we report the effect of treatment with once-weekly s.c. semaglutide 2.4 mg versus placebo on the SF-6D version 2 (SF-6Dv2) and EQ-5D-Three-Level (EQ-5D-3L) utility scores across the STEP 1-4 trials.[11-13,17]

2 | MATERIALS AND METHODS

2.1 | Trial designs and populations (STEP 1-4)

STEP 1-4 were 68-week, double-blind, randomized controlled trials assessing the efficacy and safety of semaglutide (2.4 mg s.c. once weekly) versus placebo for weight management in adults with obesity or overweight.[11-13,17] Key enrolment criteria, study objectives and endpoints are summarized in the subsequent paragraph and shown in Table S1. All trials complied with the International Conference on Harmonization Good Clinical Practice Guideline and the Declaration of Helsinki. All participants provided written informed consent.

Briefly, STEP 1, 3 and 4 enrolled adults (age ≥ 18 years) with obesity (BMI ≥ 30 kg/m^2) or with overweight (BMI ≥ 27 kg/m^2) and at least one weight-related co-morbidity.[11,13,17] STEP 2 enrolled adults with a BMI of 27 kg/m^2 or higher and T2D who had been treated with either diet and exercise alone or with a stable dose of up to three glucose-lowering agents within 90 days before screening.[12] Key exclusion criteria were self-reported weight change of 5 kg or more within 90 days before screening, or obesity treatment with surgery or a weight-loss device.[11-13,17] Comorbidities evaluated at the time of screening were dyslipidaemia, hypertension, coronary artery disease, cerebrovascular disease, obstructive sleep apnoea, impaired glucose metabolism, reproductive system disorders, liver disease, kidney disease, osteoarthritis, gout, asthma or chronic obstructive pulmonary disease (COPD).

2.1.1 | Treatment

In STEP 1, 1961 participants were randomized 2:1 to s.c. semaglutide 2.4 mg or placebo, in combination with lifestyle intervention.[11] In STEP 2, 1210 participants were randomized 1:1:1 to s.c. semaglutide 2.4 mg, s.c. semaglutide 1.0 mg or placebo, in combination with lifestyle intervention.[12] In STEP 3, 611 participants were randomized 2:1 to s.c. semaglutide 2.4 mg or placebo, as an adjunct to intensive behavioural therapy and initial low-calorie diet.[13] In STEP 4, 902 participants received open-label, once-weekly s.c. semaglutide during a 20-week run-in period that included a 16-week dose escalation period and 4 weeks on the target semaglutide dose of 2.4 mg. After the run-in period, 803 participants were randomized (2:1) to either continued semaglutide 2.4 mg or switched to placebo, for a further 48 weeks. Participants received semaglutide 2.4 mg or placebo in combination with lifestyle intervention for the trial duration (68 weeks).[17]

2.1.2 | Trial endpoints

STEP 1, 2 and 3 included the co-primary endpoints percentage change in body weight from baseline to week 68 and the achievement of a reduction in body weight of 5% or more at week 68. These endpoints were evaluated separately; however, superiority had to be shown for both to achieve the primary objective.[11-13] In STEP 4, the primary endpoint was the percentage change in body weight from week 20 to week 68.[17] 'Baseline' corresponded to week 0 in all trials, except STEP 4, in which randomization (baseline) followed a 20-week run-in period, during which all participants received semaglutide. The analysis of health utility scores was based on the trial product estimand, which modelled the average effect in all randomly assigned patients, assuming that patients had remained on treatment for the duration of the trials, and without initiation of obesity rescue medication.

Confirmatory secondary trial endpoints included the proportion of participants achieving a reduction in body weight of 10% or more, or 15% or more at week 68 (with the exception of STEP 4); and change from baseline to week 68 in waist circumference, systolic blood pressure and PRO assessments, including the SF-36v2 physical functioning score.

2.1.3 | Change in body weight at week 68

In the STEP 1-4 trials, estimated mean weight changes from randomization to week 68 with semaglutide 2.4 mg (vs. placebo) for the trial product estimand were −16.9% (−2.4%), −10.6% (−3.1%), −17.6% (−5.0%) and −8.8% (6.5%), respectively.[11-13,17]

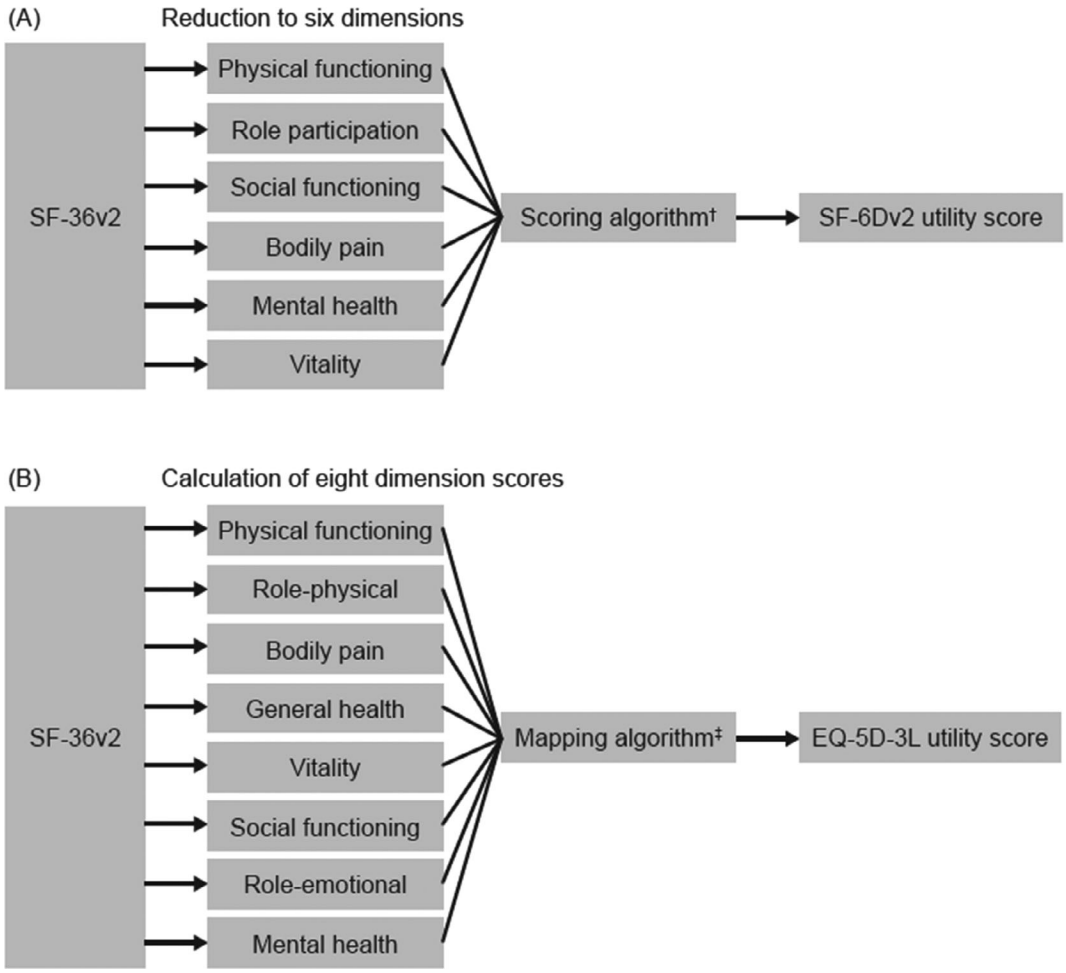

FIGURE 1 Mapping SF-36v2 data onto the A, SF-6Dv2, and B, EQ-5D-3L utility indices†‡. †Based on the methods of Brazier et al., 2020[16] and Mulhern et al., 2020.[19] ‡Based on the methods of Rowen et al., 2009.[20] EQ-5D-3L, European Quality of Life Five-Dimension Three-Level; SF-36v2, 36-item Short Form Health Survey version 2; SF-6Dv2, Short Form Six-Dimension version 2

2.2 | HRQoL assessments in STEP 1-4

HRQoL was assessed in the STEP 1-4 trials using the SF-36v2 comprising 36 questions (items) across eight domains (physical functioning, role limitations because of physical health problems, bodily pain, general health perception, vitality, social functioning, role limitations because of emotional problems, and mental health).[14,18] Changes from baseline to week 68 in the individual domains were analysed as prespecified secondary endpoints in the STEP 1-4 trials.[11-13,17]

Across the STEP 1-4 trials, there were greater improvements in SF-36v2 physical functioning scores with semaglutide 2.4 mg than with placebo at week 68: STEP 1: 2.6 versus 0.5 points, STEP 2: 2.7 versus 0.7 points, STEP 3: 2.4 versus 1.5 points and STEP 4: 1.0 versus −1.0 points.[11-13,17]

2.3 | Scoring of the SF-6Dv2 and EQ-5D-3L utility scores

In the current analysis, SF-36v2 scores were analysed post hoc and used to derive the SF-6Dv2 and EQ-5D-3L health utility scores using previously validated methods.[19,20] The item responses from STEP 1-4 were directly used to assign the SF-6Dv2 utility scores[16,19] and SF-36v2 scale scores were mapped onto the EQ-5D-3L utility index using an established mapping algorithm (Figure 1).[20] For national decision-making, it is recommended to use country-specific utility weights (value sets) that reflect the health preferences of individual countries.[21] This analysis used utility weights for the UK.

2.3.1 | SF-6Dv2 utility scores

The SF-6D classification system is one of the generic measures of health most widely used to calculate QALYs.[16] It uses responses from 11 SF-36 items to define health on six dimensions (physical functioning, role limitations, pain, vitality, social functioning and mental health). Each dimension is assessed using four to six severity levels that are combined to produce an overall health utility score.[16]

The SF-6Dv2 tool, a revised version of the SF-6D, was developed to increase the score range, and to adjust and clarify the assessment of some dimensions of HRQoL.[16] The SF-6Dv2 includes six dimensions (physical functioning, role participation, social

TABLE 1 Frequency of comorbidities[a] at baseline in the STEP 1-4 trials[11-13,17]

	STEP 1		STEP 2		STEP 3		STEP 4	
	Semaglutide 2.4 mg ($n = 1306$)	Placebo ($n = 655$)	Semaglutide 2.4 mg ($n = 404$)	Placebo ($n = 403$)	Semaglutide 2.4 mg ($n = 407$)	Placebo ($n = 204$)	Semaglutide 2.4 mg ($n = 535$)	Placebo ($n = 268$)
Number of comorbidities	Proportion of patients having specified number of comorbidities, *n* (%)							
None	328 (25.1)	163 (24.9)	[b]	[b]	99 (24.3)	49 (24.0)	144 (26.9)	70 (26.1)
1	337 (25.8)	187 (28.5)	47 (11.6)	33 (8.2)	93 (22.9)	53 (26.0)	160 (29.9)	78 (29.1)
2	298 (22.8)	135 (20.6)	77 (19.1)	72 (17.9)	96 (23.6)	43 (21.1)	103 (19.3)	68 (25.4)
3	183 (14.0)	96 (14.7)	98 (24.3)	112 (27.8)	62 (15.2)	38 (18.6)	77 (14.4)	34 (12.7)
4	96 (7.4)	43 (6.6)	89 (22.0)	105 (26.1)	31 (7.6)	14 (6.9)	38 (7.1)	15 (5.6)
5 or more	64 (4.9)	31 (4.7)	93 (23.0)	81 (20.1)	26 (6.4)	7 (3.4)	13 (2.4)	3 (1.1)
Type of comorbidity	Proportion of patients with specific type of comorbidity, *n* (%)							
Dyslipidaemia	499 (38.2)	226 (34.5)	265 (65.6)	284 (70.5)	145 (35.6)	67 (32.8)	189 (35.3)	99 (36.9)
Hypertension	472 (36.1)	234 (35.7)	276 (68.3)	287 (71.2)	145 (35.6)	67 (32.8)	199 (37.2)	99 (36.9)
Knee osteoarthritis	173 (13.2)	102 (15.6)	73 (18.1)	67 (16.6)	76 (18.7)	31 (15.2)	72 (13.5)	27 (10.1)
Obstructive sleep apnoea	159 (12.2)	71 (10.8)	68 (16.8)	54 (13.4)	58 (14.3)	19 (9.3)	61 (11.4)	33 (12.3)
Non-alcoholic fatty liver disease	101 (7.7)	62 (9.5)	85 (21.0)	94 (23.3)	23 (5.7)	12 (5.9)	37 (6.9)	18 (6.7)
Polycystic ovary syndrome[c]	62 (6.5)	34 (6.8)	7 (3.1)	10 (5.3)	17 (5.4)	10 (5.6)	15 (3.5)	10 (4.9)
Asthma or COPD	147 (11.3)	80 (12.2)	36 (8.9)	32 (7.9)	67 (16.5)	25 (12.3)	57 (10.7)	35 (13.1)
Coronary artery disease	32 (2.5)	17 (2.6)	26 (6.4)	33 (8.2)	6 (1.5)	4 (2.0)	4 (0.7)	3 (1.1)

Note: Data are from the full analysis set (STEP 2 and 3) or from randomized patients in the safety analysis set (STEP 1 and 4).
Abbreviations: COPD, chronic obstructive pulmonary disease; STEP, Semaglutide Treatment Effect in People with obesity; T2D, type 2 diabetes.
[a]Comorbidities evaluated included: dyslipidaemia, hypertension, coronary artery disease, cerebrovascular disease, obstructive sleep apnoea, impaired glucose metabolism, reproductive system disorders, liver disease, kidney disease, osteoarthritis, gout, asthma or COPD.
[b]All patients had at least one comorbidity (T2D) as an inclusion criterion.
[c]Proportion of female patients.

functioning, bodily pain, mental health, and vitality), and is scored from 10 items in the SF-36v2 with five or six severity levels for each dimension.[16]

Using UK utility weights, the SF-6Dv2 has a range from −0.574 to 1, anchored at 0 for death and 1 for perfect health. Of the 18 750 possible health states defined by the SF-6Dv2, 15.0% are below 0, indicating that they are valued as worse than death (e.g. a state of being in very severe pain and having depression/anxiety all the time).[19]

2.3.2 | EQ-5D utility scores

The EQ-5D assesses HRQoL in five dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression). Each dimension is assessed using either three response levels (EQ-5D-3L) or five response levels (EQ-5D-5L) relating to perceived problems (none, some, moderate, severe and extreme).[14] Mapping algorithms have also been established that derive EQ-5D-3L health utility scores from responses to other health surveys, including the SF-36 survey.[20] Rowen et al. used a random effects generalized least squares model, including SF-36 dimensions, squared terms and dimension interactions applied to inpatient data from the prospective 'Health Outcomes Data Repository', which includes responses from more than 23 000 inpatients to SF-36 and EQ-5D questions.[20] This model was used to derive the EQ-5D-3L utility scores from the SF-36v2 scores in the STEP trials. The EQ-5D-3L tool produces health utility scores anchored at 0 for death and 1 for perfect health.

TABLE 2 Indexation of estimated mean health utility scores in the STEP 1-4[a] trials

		SF-6Dv2 utility scores		EQ-5D-3L utility scores	
		Semaglutide 2.4 mg	Placebo	Semaglutide 2.4 mg	Placebo
STEP 1	Mean at week 0, by treatment arm	0.838	0.831	0.895	0.892
	Overall	0.836		0.894	
	Estimated mean at week 68	0.843	0.814	0.908	0.879
	Estimated change from baseline (week 0) to week 68	0.008	−0.021	0.013	−0.015
	Estimated treatment difference (95% CI)	0.029 (0.012, 0.046) $P = .001$		0.029 (0.018, 0.039) $P < .0001$	
STEP 2	Mean at week 0, by treatment arm	0.829	0.838	0.874	0.880
	Overall	0.835		0.881	
	Estimated mean at week 68	0.847	0.827	0.900	0.877
	Estimated change from baseline (week 0) to week 68	0.013	−0.008	0.018	−0.005
	Estimated treatment difference (95% CI)	0.021 (−0.002, 0.044) $P = .0729$		0.023 (0.010, 0.037) $P = .0009$	
STEP 3	Mean at week 0, by treatment arm	0.854	0.846	0.906	0.902
	Overall	0.851		0.905	
	Estimated mean at week 68	0.864	0.851	0.923	0.907
	Estimated change from baseline (week 0) to week 68	0.012	−0.001	0.018	0.002
	Estimated treatment difference (95% CI)	0.013 (−0.014, 0.041) $P = .3373$		0.015 (−0.001, 0.032) $P = .0607$	
STEP 4	Mean at week 0, by treatment arm	0.860	N/A	0.911	N/A
	Overall	0.860		0.911	
	Mean at week 20, by treatment arm	0.870	0.866	0.925	0.927
	Overall	0.869		0.926	
	Estimated mean at week 68	0.869	0.926	0.933	0.899
	Estimated change from baseline (week 20) to week 68	0.009	−0.036	0.007	−0.027
	Estimated treatment difference (95% CI)	0.045 (0.025, 0.066) $P < .0001$		0.034 (0.024, 0.044) $P < .0001$	

Abbreviations: CI, confidence interval; EQ-5D-3L, European Quality of Life Five-Dimension Three-Level; N/A, not applicable; SF-6Dv2, Short Form Six-Dimension version 2; STEP, Semaglutide Treatment Effect in People with obesity.

[a]Data are mean scores at the start of treatment (week 0, in-trial) and estimated mean at end of treatment (week 68), and estimated mean change from baseline from the on-treatment period, analysed using a mixed model for repeated measurements, including all responses before first discontinuation of treatment or initiation of other glucose-lowering medication or antiobesity therapies. Number included in the analysis set for semaglutide/placebo: STEP 1, 1306/655; STEP 2, 404/403; STEP 3, 407/204; STEP 4, 535/268.

2.4 | Statistical analysis

Health utility values were analysed for each trial using a mixed model for repeated measurements. This model handles missing data implicitly under the missing at random assumption. All responses before first discontinuation of treatment or initiation of other weight management drugs or bariatric surgery were included according to the trial product estimand.[11-13,17]

The mixed model included randomized treatment as a factor and baseline health utility score as covariate, all nested within visit. For STEP 2, the stratification groups (oral antidiabetic treatment status and HbA1c category at screening) and the interaction between stratification groups were included in the statistical model as additional factors. Results from statistical analysis were accompanied by two-sided 95% confidence intervals (CIs) and corresponding P values (significance defined as $P \leq .05$ in this analysis). The presented analysis was not adjusted for multiplicity. The method for analysing health utility scores is identical to the prespecified analysis for the individual SF-36 domain scores according to the trial product estimand.

3 | RESULTS

3.1 | Overview of STEP 1-4 baseline characteristics

The baseline demographic and clinical characteristics of enrolled patients were generally comparable across treatment groups in the STEP 1-4 trials, as shown in Table S2. The majority of patients were female (47.1%-88.2%), white (58.7%-84.3%) and their mean age ranged from 46 to 55 years. Across the individual trials at baseline (week 0), body weight ranged from 99.9 to 107.2 kg, and BMI and waist circumference ranged from 34.1-35.3 to 38.4 kg/m^2 and from 104.7-111.8 to 115.5 cm, respectively. The trials included participants with T2D (100% of the STEP 2 population) and without T2D, with

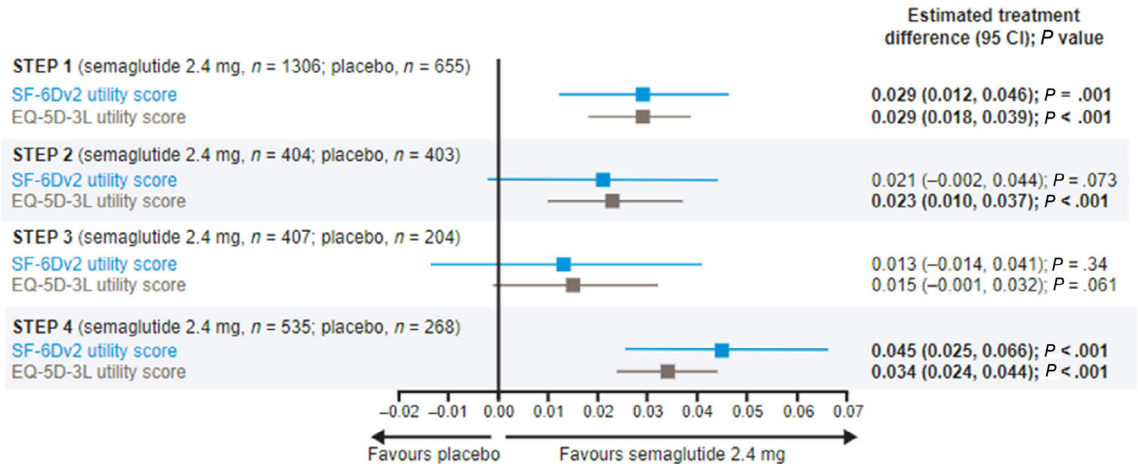

FIGURE 2 Estimated treatment differences in SF-6Dv2 and EQ-5D-3L utility scores†. †Statistically significant improvements versus placebo are indicated in bold. CI, confidence interval; EQ-5D-3L, European Quality of Life Five-Dimension Three-Level; SF-6Dv2, Short Form Six-Dimension version 2; STEP, Semaglutide Treatment Effect in People with obesity

either normoglycaemia or prediabetes at baseline (ranging from 47.1% to 85.0% and from 15% to 52.9%, respectively, in STEP 1, 3 and 4).[11-13,17] In addition, participants in STEP 2 also received additional antidiabetic medication (up to three oral antidiabetic drugs) compared with the other trials.

3.2 | Frequency of comorbidities in STEP 1-4

There were differences in the overall number and type of comorbidities identified across the trials. In STEP 1, 3 and 4, approximately 75% of participants enrolled had at least one co-morbidity at baseline, as shown in Table 1. In STEP 2, where T2D was an inclusion criterion, the majority of participants had two or more additional comorbidities at baseline (Table 1). In this trial, greater proportions of patients had concomitant dyslipidaemia, hypertension, non-alcoholic fatty liver disease or coronary artery disease, approximately 2-fold higher or more, than those in STEP 1, 3 and 4. The proportions of patients with other comorbidities were similar across the trials, although slightly smaller proportions in STEP 2 had asthma or COPD than in STEP 1, 3 and 4.

3.3 | SF-6Dv2 and EQ-5D-3L utility scores in STEP 1-4

3.3.1 | Baseline utility scores

At baseline, in the STEP 1-3 trials, mean overall SF-6Dv2 and EQ-5D-3L utility scores ranged from 0.835 to 0.851 and from 0.881 to 0.905, respectively. The corresponding values for STEP 4 at week 0 before the run-in period were 0.860 and 0.911, respectively. After the 20-week run-in of semaglutide treatment for all participants, the mean baseline overall SF-6Dv2 and EQ-5D-3L utility scores were 0.869 and 0.926 (Table 2). Baseline health status was perceived as being good across STEP 1-4 despite the presence of co-morbidities based on the comparison of mean SF-6Dv2 and EQ-5D scores to published norms.[20,22]

3.3.2 | Change in utility scores at 68 weeks

At week 68, improvements in SF-6Dv2 and EQ-5D-3L were observed with semaglutide 2.4 mg across all trials, whereas scores typically decreased with placebo (Table 2).

All changes in SF-6Dv2 utility scores from baseline favoured semaglutide. Treatment with semaglutide 2.4 mg significantly improved SF-6Dv2 utility scores by week 68 versus placebo in STEP 1 (estimated treatment difference [ETD]: 0.029; 95% CI: 0.012, 0.046) and STEP 4 (ETD: 0.045; 95% CI: 0.025, 0.066), but did not reach significance in STEP 2 (ETD: 0.021; 95% CI: −0.002, 0.044) or STEP 3 (ETD: 0.013; 95% CI: −0.014, 0.041) (Figure 2).

All changes in EQ-5D-3L utility scores from baseline favoured semaglutide. EQ-5D-3L utility score improvements at week 68 with semaglutide 2.4 mg versus placebo were significant in STEP 1 (ETD: 0.029; 95% CI: 0.018, 0.039), STEP 2 (ETD: 0.023; 95% CI: 0.010, 0.037) and STEP 4 (ETD: 0.034; 95% CI: 0.024, 0.044), but did not reach significance in STEP 3 (ETD: 0.015; 95% CI: −0.001, 0.032; Figure 2).

4 | DISCUSSION

Clinical studies have shown the weight-loss benefits of treatment with once-weekly s.c. semaglutide 2.4 mg in patients with overweight or obesity, both with and without T2D.[11-13,17,23-25]

The current analysis showed that treatment with semaglutide 2.4 mg is also associated with improvements in HRQoL. Significant improvements with semaglutide 2.4 mg versus placebo were observed

in STEP 1 and 4 for SF-6Dv2, and in STEP 1, 2 and 4 for EQ-5D-3L health utility scores. Furthermore, SF-6Dv2 and EQ-5D results were consistent across the trials, although the variance was greater in SF-6Dv2 than in EQ-5D.

The lack of statistical significance for the effect of semaglutide versus placebo on SF-6Dv2 in STEP 2 may be explained by the trial population, which only included patients with diagnosed T2D (average disease duration of 8.2 years and patients using up to three antidiabetic medications).[12] By contrast, patients in the other STEP trials were either normoglycaemic or were in the early stages of diabetes (prediabetes).[11-13,17] This may have an important role in different results observed between the trials. Across STEP 1-4, participants in STEP 2 treated with semaglutide 2.4 mg had the lowest relative body weight change from week 0 to week 68.[11-13,17] Differences between trials in populations and background therapies may have influenced the weight losses observed in each trial. Patients in STEP 1, 3 and 4 had a lower comorbidity burden (the majority had one or more), compared with those in STEP 2 with a higher burden (the majority had three or more including T2D). Additionally, the type of comorbid illness differed between the trials. Greater proportions of patients in STEP 2 had dyslipidaemia, hypertension, non-alcoholic fatty liver disease or coronary artery disease, compared with those in the other trials. The highest presence of co-morbidities at baseline in STEP 2 was also consistent with the lowest mean health utility score (0.835 and 0.881 for SF6Dv2 and EQ-5D-3L, respectively) observed across STEP 1-4. Nevertheless, improvements in HRQoL have previously been observed with once-weekly s.c. semaglutide 0.5 mg or 1.0 mg versus placebo in patients with T2D at high cardiovascular risk, with effects mediated partly by changes in HbA1c and body weight,[26] as well as with once-daily oral semaglutide versus dulaglutide in patients with T2D in Japan.[27]

Difference in additional therapy may also contribute to the changes in SF-36v2 and EQ-5D-3L scores observed across the trials. The lack of statistical significance in STEP 3 for treatment differences with respect to either health utility score may reflect the intensive behavioural therapy, which also included an initial low-calorie diet, offered to both the placebo and semaglutide groups in this trial.[13] The behavioural therapy contributed to greater weight loss in the placebo group in STEP 3 compared with the other trials, such that no (or negligible) decline occurred in HRQoL in the placebo group.

By contrast, the larger treatment differences in mean health utility score observed in STEP 4 may reflect the two-stage design of this trial, in which one-third of participants who reached the semaglutide maintenance dose of 2.4 mg at week 20 were randomized to placebo for the remainder of the trial. By week 68, the placebo group had gained weight relative to week 20, which may explain the comparatively large difference between semaglutide and placebo groups with respect to change in HRQoL.[17]

It is important to note that across the studies, baseline health utility scores were generally high, with limited room for improvement. Treatment differences were typically driven by an increase in mean health utility score in the semaglutide 2.4 mg group and a decrease in mean health utility score in the placebo group.

A limitation of the current study is that the EQ-5D scores were derived by a mapping from the SF-36v2, and not from the original EQ-5D survey. The mapping algorithm may overpredict the utility score for people in more severe health states, which can also limit the room for improvement.[20] In addition, HRQoL was not collected or reported after the randomized treatment phase was complete in the STEP 1-4 trials. Differences in health status may be evident following trial end given that weight regain can occur after pharmacological treatment is stopped. In the STEP 1 extension study, participants regained two-thirds of their prior weight loss with semaglutide in the first year,[28] while substantial weight regain was reported in other studies with orlistat and loracserin.[29,30] This effect was partially explored with the STEP 4 randomized withdrawal trial design, where a larger decrease in utility score was evident in the placebo group after randomization compared with the other STEP trials.

While some of the greatest increases in health utility in the semaglutide arm were observed in STEP 3, the difference between semaglutide and placebo effects on health utility was non-significant in this trial, because of the lack of deterioration in the control group, perhaps caused by the intensive behavioural intervention. Self-perceived improvements may become more evident as physical capabilities improve; when physical exercise is an integrated part of treatment, as with intensive behavioural therapy, this effect may be further enhanced.[31] This interplay between medical and behavioural treatment warrants further research.

Caution should always be used when comparing clinical programmes of different therapeutic interventions, even in patient populations with similar demographics, disease status and co-morbidities. However, health utility analysis may enable HRQoL improvements with a given treatment to be translated into a health utility score, which may indicate or confirm clinical improvements. In a pooled meta-analysis of six randomized clinical trials (RCTs) to evaluate health utility using SF-36 and mapped EQ-5D scores, insulin degludec was associated with a modest but significant improvement in health utility compared with insulin glargine in 4001 patients with diabetes.[32] Further health utility analyses are warranted in patients with overweight or obesity.

The current analysis used prespecified HRQoL outcome data from the phase 3 STEP 1-4 RCTs, in selected patient populations within a controlled trial setting. It is, therefore, probable that these health utility scores would differ from those based on data collected prospectively from real-world clinical practice. While RCTs remain the gold standard for determining clinical efficacy, real-world evidence (RWE) enables assessment within a wider context, focusing on effectiveness and treatment-related costs of lifestyle modification and pharmacotherapy in broader patient populations with varying baseline characteristics.[33,34] A previous RWE study of electronic health records from patients with T2D undergoing bariatric surgery reported similar postsurgical reductions in HbA1c and BMI to clinical trials.[35] As such, RWE-confirmed health utility scores would probably complement and lend further weight to findings obtained within the clinical trial environment, and this research may generate interesting findings.

In conclusion, once-weekly s.c. semaglutide 2.4 mg was associated with improvement in health utility scores compared with placebo, reaching statistical significance for STEP 1 and 4 with SF-6Dv2, and for STEP 1, 2 and 4 with EQ-5D scores. Baseline utility scores were generally high, and treatment differences were typically driven by an increase in mean utility score in the semaglutide 2.4 mg group and a decrease in mean utility score in the placebo group. These observations, in addition to the weight loss benefits reported in the primary analyses for the individual STEP 1-4 trials, show how health-related utility scores may indicate improvements in patients with overweight or obesity in an RCT setting.

AUTHOR CONTRIBUTIONS

All authors developed the study concept and design, interpreted data and critically revised and completed the manuscript. TH-H performed statistical analysis of the manuscript.

ACKNOWLEDGEMENTS

The STEP 1-4 trials included in this analysis were sponsored by Novo Nordisk A/S and are registered with ClinicalTrials.gov (NCT03548935 [STEP 1], NCT03552757 [STEP 2], NCT03611582 [STEP 3] and NCT03548987 [STEP 4]). The authors thank the trial participants, the investigators and trial site staff who conducted the trials. The authors would like to acknowledge the contribution of Henrik Hjorth Meinke to the analysis of the data and the development of the abstract. Medical writing support was provided by Suzanne Axford-Bryars, a contract writer working on behalf of Apollo, and Judah Lynch of Apollo, OPEN Health Communications, funded by Novo Nordisk A/S, in accordance with Good Publication Practice 3 (GPP) guidelines (www.ismpp.org/gpp-2022).

CONFLICT OF INTEREST

JBB is an employee of QualityMetric Inc. who distributes the SF-36v2 and SF-6Dv2. SL, CL and TH-H are all employees of Novo Nordisk A/S, Søborg, Denmark.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/dom.15090.

DATA AVAILABILITY STATEMENT

Data will be shared with bona fide researchers who submit a research proposal approved by the independent review board. Individual participant data will be shared in data sets in a de-identified and anonymized format. Data will be made available after research completion and approval of the product and product use in the EU and the USA. Information about data access request proposals can be found at novonordisk-trials.com.

ORCID

Jakob Bue Bjorner https://orcid.org/0000-0001-7033-8224
Sara Larsen https://orcid.org/0000-0002-3887-4948
Christopher Lübker https://orcid.org/0000-0003-3220-801X
Thomas Holst-Hansen https://orcid.org/0000-0002-5314-7802

REFERENCES

1. Garvey WT, Mechanick JI, Brett EM, et al. American Association of Clinical Endocrinologists and American College of endocrinology comprehensive clinical practice guidelines for medical care of patients with obesity. *Endocr Pract*. 2016;22(Suppl 3):1-203.
2. Franks P, Hanmer J, Fryback DG. Relative disutilities of 47 risk factors and conditions assessed with seven preference-based health status measures in a national U.S. sample: toward consistency in cost-effectiveness analyses. *Med Care*. 2006;44(5):478-485.
3. Kolotkin RL, Gabriel Smolarz B, Meincke HH, Fujioka K. Improvements in health-related quality of life over 3 years with liraglutide 3.0 mg compared with placebo in participants with overweight or obesity. *Clin Obes*. 2018;8(1):1-10.
4. Kolotkin RL, Andersen JR. A systematic review of reviews: exploring the relationship between obesity, weight loss and health-related quality of life. *Clin Obes*. 2017;7(5):273-289.
5. Krishna SG, Rawal V, Durkin C, et al. Weight loss surgery reduces healthcare resource utilization and all-cause inpatient mortality in morbid obesity: a propensity-matched analysis. *Obes Surg*. 2018;28(10):3213-3220.
6. FDA. WEGOVY® prescribing information. 2022 https://www.accessdata.fda.gov/drugsatfda_docs/label/2021/215256s000lbl.pdf. (accessed October 2022)
7. European Medicines Agency. WEGOVY® summary of product characteristics. 2022 https://www.ema.europa.eu/en/medicines/human/EPAR/wegovy (accessed June 6, 2022)
8. MHRA. Wegovy® UK summary of product characteristics. Available at: https://mhraproducts4853.blob.core.windows.net/docs/105969ec8e62c5b6ad38141075d0854eea90b676 2022.
9. Canada NN. Health Canada approves Wegovy™ for the treatment of adults with obesity. https://www.novonordisk.ca/content/dam/nncorp/ca/en/press-releases/2021/Health-Canada-approves-Wegovy%E2%84%A2-for-the-treatment-of-adults-with-obesity.pdf 2022.
10. EMA. Wegovy summary of product characteristics. https://www.ema.europa.eu/en/documents/product-information/wegovy-epar-product-information_en.pdf 2022.
11. Wilding JPH, Batterham RL, Calanna S, et al. Once-weekly semaglutide in adults with overweight or obesity. *N Engl J Med*. 2021;384(11):989-1002.
12. Davies M, Faerch L, Jeppesen OK, et al. Semaglutide 2.4 mg once a week in adults with overweight or obesity, and type 2 diabetes (STEP 2): a randomised, double-blind, double-dummy, placebo-controlled, phase 3 trial. *Lancet*. 2021;397(10278):971-984.
13. Wadden TA, Bailey TS, Billings LK, et al. Effect of subcutaneous semaglutide vs placebo as an adjunct to intensive behavioral therapy on body weight in adults with overweight or obesity: the STEP 3 randomized clinical trial. *JAMA*. 2021;325(14):1403-1413.
14. Kaplan RM, Hays RD. Health-related quality of life measurement in public health. *Annu Rev Public Health*. 2022;43:355-373.
15. EuroQol Group. EuroQol-a new facility for the measurement of health-related quality of life. *Health Policy*. 1990;16(3):199-208.
16. Brazier JE, Mulhern BJ, Bjorner JB, et al. Developing a new version of the SF-6D health state classification system from the SF-36v2: SF-6Dv2. *Med Care*. 2020;58(6):557-565.
17. Rubino D, Abrahamsson N, Davies M, et al. 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*. 2021;325(14):1414-1425.
18. 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.
19. Mulhern BJ, Bansback N, Norman R, Brazier J. Valuing the SF-6Dv2 classification system in the United Kingdom using a discrete-choice experiment with duration. *Med Care*. 2020;58(6):566-573.
20. Rowen D, Brazier J, Roberts J. Mapping SF-36 onto the EQ-5D index: how reliable is the relationship? *Health Qual Life Outcomes*. 2009;7(1):27.
21. EuroQol Group. *EQ-5D value sets*. 2021. https://euroqol.org/publications/key-euroqol-references/value-sets/.

22. Poder TG, Carrier N. Quebec health-related quality of life population norms in adults using the SF-6Dv2: decomposition by sociodemographic data and health problems. *Med Care*. 2022;60(7):545-554.
23. Garvey WT, Batterham RL, Bhatta M, et al. Two-year effect of semaglutide 2.4 mg once-weekly in adults with overweight or obesity (STEP 5): a randomized, double-blind, placebo-controlled, phase 3 trial. *Nat Med*. 2022;28:2083-2091.
24. Kadowaki T, Isendahl J, Khalid U, et al. Semaglutide once a week in adults with overweight or obesity, with or without type 2 diabetes in an east Asian population (STEP 6): a randomised, double-blind, double-dummy, placebo-controlled, phase 3a trial. *Lancet Diabetes Endocrinol*. 2022;10(3):193-206.
25. Rubino DM, Greenway FL, Khalid U, et al. Effect of weekly subcutaneous semaglutide vs daily liraglutide on body weight in adults with overweight or obesity without diabetes: the STEP 8 randomized clinical trial. *JAMA*. 2022;327(2):138-150.
26. Jódar E, Michelsen M, Polonsky W, et al. Semaglutide improves health-related quality of life versus placebo when added to standard of care in patients with type 2 diabetes at high cardiovascular risk (SUSTAIN 6). *Diabetes Obes Metab*. 2020;22(8):1339-1347.
27. Ishii H, Hansen BB, Langer J, Horio H. Effect of orally administered semaglutide versus dulaglutide on diabetes-related quality of life in Japanese patients with type 2 diabetes: the PIONEER 10 randomized, active-controlled trial. *Diabetes Ther*. 2021;12(2):613-623.
28. Wilding JPH, Batterham RL, Davies M, et al. Weight regain and cardiometabolic effects after withdrawal of semaglutide: the STEP 1 trial extension. *Diabetes Obes Metab*. 2022;24(8):1553-1564.
29. Sjöström L, Rissanen A, Andersen T, et al. Randomised placebo-controlled trial of orlistat for weight loss and prevention of weight regain in obese patients. European Multicentre Orlistat Study Group. *Lancet*. 1998;352(9123):167-172.
30. Smith SR, Weissman NJ, Anderson CM, et al. Multicenter, placebo-controlled trial of lorcaserin for weight management. *N Engl J Med*. 2010;363(3):245-256.
31. MacDonald CS, Nielsen SM, Bjørner J, et al. One-year intensive lifestyle intervention and improvements in health-related quality of life and mental health in persons with type 2 diabetes: a secondary analysis of the U-TURN randomized controlled trial. *BMJ Open Diabetes Res Care*. 2021;9(1):e001840.
32. Freemantle N, Evans M, Christensen T, Wolden ML, Bjorner JB. A comparison of health-related quality of life (health utility) between insulin degludec and insulin glargine: a meta-analysis of phase 3 trials. *Diabetes Obes Metab*. 2013;15(6):564-571.
33. Kim HS, Lee S, Kim JH. Real-world evidence versus randomized controlled trial: clinical research based on electronic medical records. *J Korean Med Sci*. 2018;33(34):e213.
34. Schneeweiss S, Patorno E. Conducting real-world evidence studies on the clinical outcomes of diabetes treatments. *Endocr Rev*. 2021;42(5):658-690.
35. Rassen JA, Murk W, Schneeweiss S. Real-world evidence of bariatric surgery and cardiovascular benefits using electronic health records data: a lesson in bias. *Diabetes Obes Metab*. 2021;23(7):1453-1462.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Bjorner JB, Larsen S, Lübker C, Holst-Hansen T. The improved health utility of once-weekly subcutaneous semaglutide 2.4 mg compared with placebo in the STEP 1-4 obesity trials. *Diabetes Obes Metab*. 2023;25(8):2142-2150. doi:10.1111/dom.15090