

An Indirect Comparison of Changes in the Impact of Weight on Quality of Life Among Subjects with Type 2 Diabetes Treated with Antihyperglycemic Agents in Dual Therapy with Metformin

Shana B. Traina · April Slee · Suzy Van Sanden · Joris Diels

Received: September 22, 2017 / Published online: December 13, 2017
© The Author(s) 2017. This article is an open access publication

ABSTRACT

Introduction: It is important to capture the patient experience with a diabetes treatment in clinical trials; however, use of instruments to assess patient-reported outcomes (PROs) in diabetes trials is inconsistent and results may not be reported alongside primary efficacy data. In lieu of head-to-head data, indirect comparisons can be used to compare competing interventions. In this study, we used indirect comparison methods to assess differences in PRO score changes between canagliflozin and other antihyperglycemic agents as add-on to metformin.

Methods: Literature searches were performed to identify studies that reported the same PRO

Enhanced content To view enhanced content for this article go to <http://www.medengine.com/Redeem/17FCF06030E8C76D>.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13300-017-0343-7>) contains supplementary material, which is available to authorized users.

S. B. Traina (✉)
Global Market Access CVM, Janssen Global Services,
LLC, 920 Route 202 South, Raritan, NJ 08869, USA
e-mail: straina@its.jnj.com

A. Slee
Axio Research, Seattle, WA, USA

S. Van Sanden · J. Diels
Janssen Research & Development, Beerse, Belgium

instruments that were collected across four trials of canagliflozin in dual or triple therapy. Extensive searches identified only one study that was sufficiently similar in design and reported common PRO results using the Impact of Weight on Quality of Life-Lite (IWQoL-Lite): the DURATION-2 study of exenatide once-weekly (QW) versus sitagliptin and pioglitazone. This study was compared with the CANTATA-D study of canagliflozin versus sitagliptin. Bayesian indirect comparisons were performed to assess mean change in IWQoL-Lite total score. A fixed-effects model with noninformative priors was used to estimate between-treatment differences. Sensitivity analyses examined differences in trial populations.

Results: In the primary analysis, the probability that canagliflozin treatment results in greater improvement in IWQoL-Lite total score versus exenatide, sitagliptin, and pioglitazone was 60.0%, 89.9%, and 99.5%, respectively. When the CANTATA-D population was restricted using DURATION-2 inclusion/exclusion criteria, canagliflozin was also associated with a higher probability of having greater improvement in IWQoL-Lite than exenatide, sitagliptin, and pioglitazone.

Conclusions: These findings suggest that improvements in the impact of weight on health-related quality of life may be greater with canagliflozin than exenatide, sitagliptin, and pioglitazone. This analysis also demonstrates

the application of indirect comparison methodology to PRO data and provides examples of advantages and challenges associated with performing indirect comparisons of PRO data.

Keywords: Canagliflozin; Exenatide; Health-related quality of life; Indirect comparison; IWQoL-Lite; Sitagliptin; Type 2 diabetes mellitus; Pioglitazone

INTRODUCTION

Lowering blood glucose levels is the primary goal of type 2 diabetes mellitus (T2DM) management [1]. Trials evaluating new T2DM therapies are designed to measure changes in glycated hemoglobin (HbA1c) levels [2]; however, glycemic endpoints and other clinical measurements may not fully capture the patient experience with treatment. The potential inclusion of endpoints beyond glycemic efficacy in T2DM trials has received attention from the US Food and Drug Administration, which held a workshop to discuss outcomes that are relevant and important to people living with T2DM [3]. These additional measures may include patient-reported outcomes (PROs), which can be used to provide concrete and cost-effective ways to track relationships between physical changes, such as weight gain or loss, and changes in self-care behaviors and health-related quality of life [4]. Tracking such changes is particularly important in T2DM because self-care behaviors such as healthy eating, physical activity, daily monitoring activities, medication adherence, and risk reduction strategies are integral to disease management [5]. However, it can be challenging for people with T2DM to be motivated to adhere to healthy behaviors symptoms with benefits that are often not immediately tangible, but rather develop over the long term. Furthermore, negative side effects of some antihyperglycemic agents (AHAs) can be disincentives to medication adherence and disease management behaviors [6].

Weight-related PROs are particularly important in T2DM because more than 85% of people

with T2DM are overweight and more than 50% are obese [7]. Recent studies have shown that weight reduction in even modest amounts (e.g., 2–3 kg) can improve a person's weight-related quality of life and satisfaction with physical and emotional health [6]. In addition, weight loss of approximately 5–10% has been associated with a higher likelihood of achieving clinically meaningful changes in glycemic control [8]. While glycemic control is the fundamental goal in the treatment of T2DM, adherence to behaviors that promote glycemic control may be critically undermined by a lack of simultaneous weight management [9]. Weight management can be especially challenging for patients who use antihyperglycemic agents that are associated with weight gain (e.g., insulin, sulfonylureas, thiazolidinediones), as treatment-related weight gain can compromise weight-management efforts, glycemic control, and motivation to perform disease management behaviors. Newer treatments developed for T2DM have been shown to be weight neutral [e.g., dipeptidyl peptidase-4 (DPP-4) inhibitors] or to promote weight reduction [e.g., glucagon-like peptide-1 receptor agonists and sodium glucose co-transporter 2 (SGLT2) inhibitors] [1]. Beyond glycemic control and weight loss, some of these newer agents also offer a range of additional benefits, such as blood pressure reduction and a low risk of hypoglycemia [8, 10].

Canagliflozin is an SGLT2 inhibitor approved as an adjunct to diet and exercise to improve glycemic control among adults with T2DM. In clinical trials, canagliflozin improved HbA1c levels and reduced body weight in a broad range of subjects with T2DM as monotherapy and in combination therapy [11]. The canagliflozin clinical program included head-to-head randomized controlled trials of canagliflozin versus sitagliptin (dual therapy with metformin and triple therapy with metformin plus sulfonylurea) and glimepiride (dual therapy with metformin) [12–14]. In dual therapy, canagliflozin 100 mg demonstrated noninferiority and canagliflozin 300 mg demonstrated superiority in HbA1c lowering versus sitagliptin and glimepiride at 52 weeks [12, 14]; canagliflozin 300 mg demonstrated superiority in HbA1c lowering

versus sitagliptin in triple therapy at 52 weeks [13]. Both canagliflozin doses significantly lowered body weight compared with sitagliptin and glimepiride in the three head-to-head trials [12–14]. In addition to direct comparisons in clinical trials, network meta-analyses of clinical trial data have been performed to provide indirect comparisons of efficacy with canagliflozin versus other antihyperglycemic agents. These analyses have shown that HbA1c reductions and weight loss with canagliflozin as monotherapy and in dual and triple therapy were generally similar or greater versus other antihyperglycemic agents, including insulin, sulfonylureas, DPP-4 inhibitors, and other SGLT2 inhibitors [15–17]. Together with previous analyses demonstrating a relationship between weight-related quality of life and weight loss [4, 6], these findings imply that canagliflozin-associated weight loss may provide greater improvements in many aspects of health-related quality of life compared with other antihyperglycemic agents across lines of therapy. However, data from head-to-head comparative trials are not currently available to test this hypothesis, and indirect comparisons of PRO data with canagliflozin have not been performed.

Lack of direct comparisons are not unique to PROs, and the need for comparative effectiveness data has motivated international organizations to develop guidelines on appropriate methods for indirect comparisons of clinical outcomes [18–22]. Network meta-analyses have been used to evaluate differences between multiple treatment regimens with respect to clinical outcomes in various therapeutic areas [23, 24]. Although the application of indirect comparisons to PRO data is a logical extension, few indirect comparisons have been performed using these data. A likely contributor to the paucity of indirect comparisons is the wide variety of instruments being used to capture different health-related quality of life concepts in clinical trials [25–27]. For example, 20 different PRO instruments have been used in clinical trials of recently approved antihyperglycemic agents [28]. Several instruments can be used to measure the same or similar concepts; however, comparing results from different

instruments has the potential to introduce bias and heterogeneity [29]. Even when the same instrument is used in different trials, indirect comparison is difficult if the trials were of different length, if PRO assessments were conducted at different time points, if different comparators were used, or if the results were not presented separately by treatment groups. In addition, because PROs are usually not the primary endpoint in large-scale studies, PRO results may not be published, resulting in significant reporting bias [30]. Furthermore, unlike traditional biomarker endpoints, PROs can be sensitive to cultural differences, so indirect comparison may not provide fair assessments if studies were conducted in culturally different geographic regions.

The goal of this analysis was to apply indirect comparison methods to assess differences in changes in PROs between canagliflozin and other antihyperglycemic agents using published results from randomized controlled trials. In addition, this manuscript describes some key considerations when performing indirect comparisons using PRO data.

METHODS

Literature Search Criteria and Trial Identification

Extensive literature searches were conducted to identify relevant publications for comparison using MEDLINE, MEDLINE-In-Process, Embase, and Cochrane CENTRAL Register electronic databases. Search strategy details are described in Appendix 1. In addition, hand searches were performed to capture data from conferences, registries, and health technology assessment websites related to unpublished clinical trials. The ClinicalTrials.gov website was also searched to identify any trials with results that were either not yet published in a scientific journal or not indexed in the electronic databases that were searched.

Searches were conducted to identify T2DM studies that used any of the publicly licensable PRO measures used in phase 3 studies of canagliflozin. These measures included Impact of

Weight on Quality of Life-Lite (IWQoL-Lite), the Multidimensional Diabetes Questionnaire (MDQ), motivational items, Short Form-36 (SF-36), the EuroQoL-5 Dimension (EQ-5D), and the Diabetes Utility Index (DUI).

Detailed searches were conducted for English-language articles published from 2005 to May 3, 2013 using terms related to the publicly licensable PRO measures of interest and more general terms such as quality of life, survey, questionnaires, patient-reported outcome or PRO, patient input, self-reported outcome, instrument, assessment, patient perspective, focus group, and cognitive interview (see Appendix 1). Search terms were combined to capture the disease of interest (T2DM), the endpoint of interest (any of the PRO measures listed above), and the study type of interest (randomized controlled trials). Searches were conducted separately by AHA class (i.e., DPP-4 inhibitors, GLP-1 receptor agonists, SGLT2 inhibitors). Ongoing searches have been performed and reviewed on a monthly basis from May 2013 to present to identify additional articles meeting the study selection criteria. Relevant articles were independently identified by two researchers who manually reviewed article titles and abstracts, where available. Studies were ineligible if they were published in abstract form only, were review articles, reported data from open-label or other non-blinded studies (e.g., observational studies), had no matching PRO endpoints, or only described economic implications of treatments (e.g., cost-effectiveness). Studies were eligible for inclusion in the analysis if they were sufficiently similar to a phase 3 study of canagliflozin that collected PRO data. Specifically, studies eligible for inclusion were required to

- Report PRO results by treatment arm
- Have similar assessment time points (26 ± 4 or 52 ± 4 weeks)
- Have similar background therapy (dual therapy with metformin or triple therapy with metformin plus sulfonylurea)
- Have a placebo or an active comparator arm in common with a canagliflozin trial (sitagliptin or glimepiride)
- Describe the analysis methods in sufficient detail that they could be replicated using

data from a canagliflozin trial (e.g., model specification for least squares mean change with presentation of least squares mean changes by treatment)

A total of 311 articles were screened in the initial search. Of these, only one clinical trial had published PRO data that could be compared with data from a canagliflozin phase 3 clinical trial (Fig. 1). An illustration of the resulting network is shown in Fig. 2. No additional relevant studies were identified in subsequent searches.

Sources for Model Estimates

The DURATION-2 study evaluated exenatide once-weekly (QW) compared with sitagliptin or pioglitazone as add-on to metformin [31, 32]. As such, DURATION-2 was sufficiently similar to the CANTATA-D (DIA3006) phase 3 study [12] that evaluated canagliflozin compared with placebo and sitagliptin in subjects on a background of metformin. DURATION-2 reported changes in IWQoL-Lite total score, EQ-5D index score, Diabetes Treatment Satisfaction Questionnaire scale score, and the Psychological General Well-being index score from baseline to week 26. Data from a separate publication reporting results from a 26-week open-label extension of DURATION-2 were not included as no PRO data were reported. In the CANTATA-D study, IWQoL-Lite assessments were performed at baseline, week 26, and week 52 [33]. The CANTATA-D study did not employ the EQ-5D, Diabetes Treatment Satisfaction Questionnaire, or Psychological General Well-being instruments; therefore, indirect comparisons could only be made using IWQoL-Lite data for DURATION-2 versus CANTATA-D.

The IWQoL-Lite questionnaire is an instrument that has been broadly used to quantitatively assess the relationship between body weight and health-related quality of life [34]. The IWQoL-Lite total score is composed of five domains in which weight may impact day-to-day living; scores are used to derive a holistic assessment of the overall burden of weight on quality of life. IWQoL-Lite is a relevant clinical endpoint in T2DM because it provides a

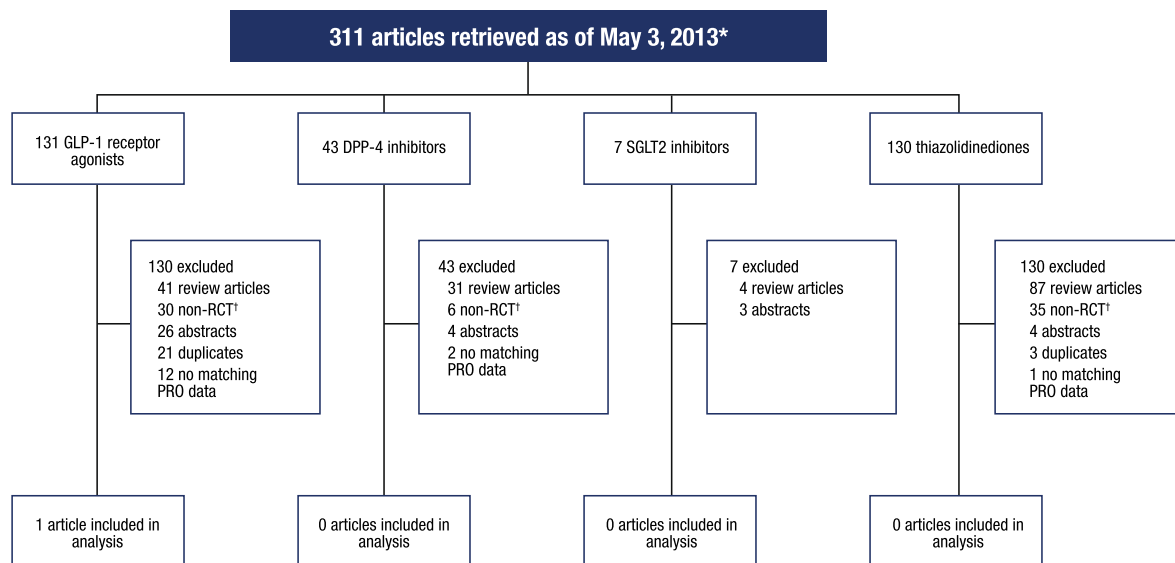


Fig. 1 Flow diagram. *GLP-1* glucagon-like peptide-1, *DPP-4* dipeptidyl peptidase-4, *SGLT2* sodium glucose co-transporter 2, *RCT* randomized controlled trial, *PRO* patient-reported outcome. *Monthly updated searches

through July 31, 2017 found no additional studies meeting study inclusion criteria. [†]Includes publications reporting study designs, open-label, observational, or non-blinded studies, and economic analyses

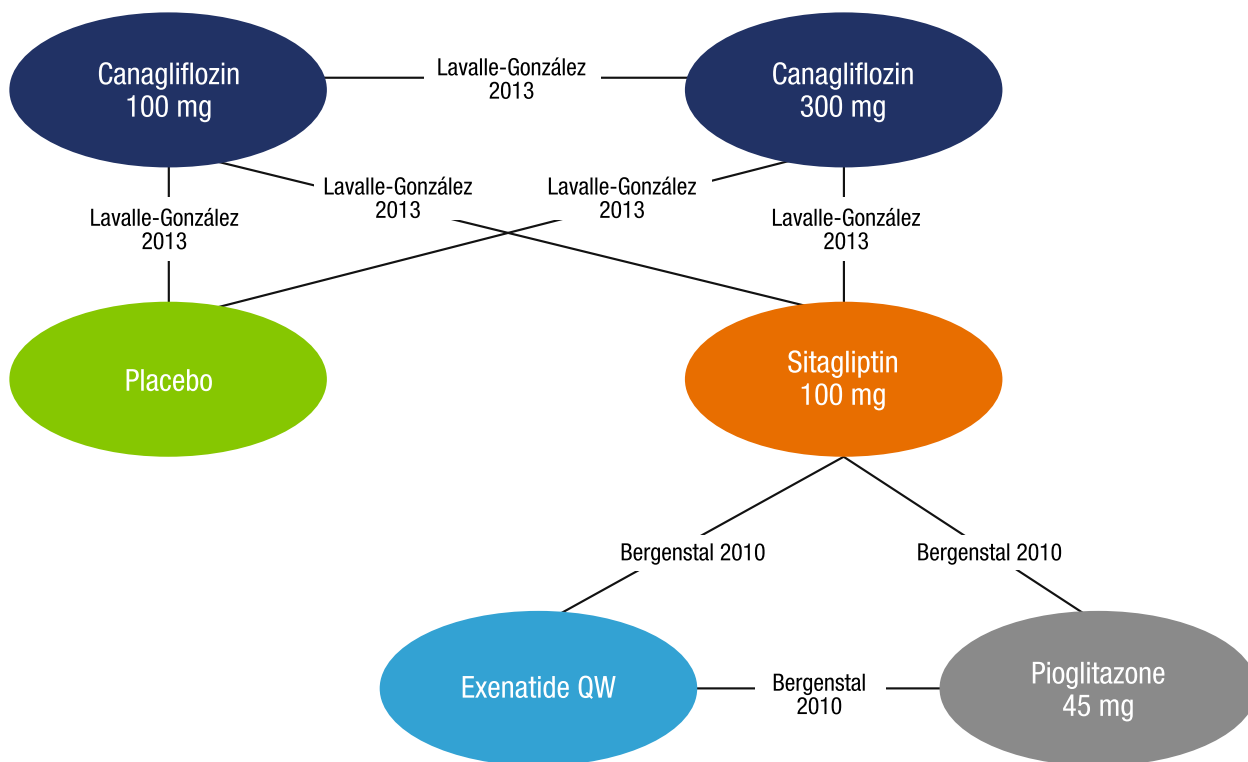


Fig. 2 Network of evidence. *QW* once weekly

measure of the effects of antihyperglycemic treatment beyond HbA1c.

Model Development

While the methodology for indirect comparisons of PRO data is generally similar to the methodology used for clinical variables, there are some notable differences. Most importantly, potential effect modifiers include characteristics specific to the PRO concept under investigation. In this case, individuals may perceive body weight in the context of cultural norms and may interpret PRO instrument items differently on the basis of a set of variables that may be irrelevant for the comparison of biomarkers. Previous studies have identified gender, race, and baseline body mass index (BMI) as characteristics that might have a material impact on results [35]. To address the potential impact of race on the results, the DURATION-2 race categories were mapped onto the CANTATA-D data. The inclusion and exclusion criteria of these two trials were compared, along with other published baseline characteristics and the countries in which the studies were conducted. Baseline characteristics were compared across studies using the *t* test or Chi square test.

In DURATION-2 ($N = 491$), inclusion and exclusion criteria were more restrictive than the criteria used in CANTATA-D ($N = 1020$), so sensitivity analyses were performed to account for differences in the trial populations. Specifically, analyses were performed using data from the total CANTATA-D study population (overall population) and using data from a subset of the CANTATA-D population ($n = 516$) that met all of the inclusion and exclusion criteria specified in DURATION-2 [31] (CANTATA-DD2 population).

Subject-level data from DURATION-2 were not available [31], so the published least squares mean changes from baseline to week 26 in IWQoL-Lite total scores were used. The DURATION-2 analysis of covariance (ANCOVA) models included treatment, country, randomization stratification factor (baseline HbA1c strata $< 9\%$ versus $\geq 9\%$), and baseline IWQoL-Lite score as fixed covariates and used the last observation carried forward (LOCF) to impute missing values

for subjects who withdrew prior to week 26. Subject-level data from CANTATA-D comparing canagliflozin to sitagliptin were available [33], so comparable least squares means were derived from ANCOVA models using LOCF imputation. The analysis of CANTATA-D data used fixed covariates of treatment, baseline IWQoL-Lite total score, and a binary term indicating whether the subject was from a DURATION-2 country (i.e., USA, India, or Mexico). Because published results from the DURATION-2 study did not report the proportions of subjects from each country in the study population, further refinement in country-level analysis was not possible.

Data for canagliflozin 100 and 300 mg were pooled as the impact on weight loss was fairly similar for both doses. Least squares means for the CANTATA-D data were evaluated using the baseline population distribution from DURATION-2 for race, mean IWQoL-Lite total score, and the modeled impact of having all subjects in one of the DURATION-2 countries. In order to evaluate the sensitivity of the results to the LOCF imputation assumption, a complete case analysis was also performed for the canagliflozin CANTATA-D trial, and these results are included in the current analysis for comparison. However, complete case analyses were not included in the DURATION-2 manuscript for comparison [31].

Indirect Comparison

The goal of this analysis was to compare the effect of canagliflozin to exenatide on the basis of the change in IWQoL-Lite total score. Since these treatments were not compared within the same trial, they could not be compared directly. However, an indirect comparison is made possible because both trials studied sitagliptin, which can serve as a common comparator (i.e., the difference between exenatide QW and sitagliptin in DURATION-2 can be compared to the difference between canagliflozin and sitagliptin in CANTATA-D). We utilized a Bayesian approach to enable the calculation of the probabilities of each intervention being the most effective. In other words, this approach quantifies the uncertainty in ranking treatments in

order to make statements such as “There is $X\%$ probability that treatment A results in a better outcome than treatment B”, which is more natural in a clinical decision-making context. The Bayesian approach is supported by several international guidelines for indirect comparisons and network meta-analysis [18–22].

Mean changes in IWQoL-Lite total score based on the DURATION-2 publication and calculated from the CANTATA-D data were used as model inputs. Separate analyses were performed using adjusted and unadjusted mean changes from CANTATA-D. When a sufficient number of studies are contributing information to an indirect comparison, a random effects model may be used to better account for study variability or differences in the clinical and methodological execution of the trials. Given that there were only two studies available for this analysis, fitting a random-effects model was not appropriate as there was not enough information to properly estimate the between-study variability. Thus, a fixed-effects model with noninformative priors was used to estimate the difference in IWQoL-Lite total score between canagliflozin and exenatide QW, sitagliptin, and pioglitazone. Specification of a prior distribution or uncertainty about treatment differences prior to data analysis is a requirement for Bayesian models. Noninformative priors were selected for this analysis, which is appropriate for situations in which we know very little about the treatment differences and want to model the data without making assumptions about the uncertainty.

Compliance with Ethics

This article does not contain any new studies with human participants performed by any of the authors.

RESULTS

Baseline Characteristics

Baseline characteristics of the DURATION-2 and CANTATA-D analysis populations are

summarized in Table 1. For DURATION-2, the indirect comparison analysis was based on data from subjects who had an IWQoL-Lite score at baseline and at least one follow-up time point [$n = 401/491$ (82%)]. Baseline characteristics for this subset of subjects were not presented in published results, so the overall trial characteristics were used (Table 1) [31]. For the CANTATA-D study, the analysis was also based on data from subjects with an IWQoL-Lite score at baseline and at least one change score prior to week 26 [$n = 1020/1101$ (93%)].

Important differences in race and gender were observed in the DURATION-2 and CANTATA-D populations. The DURATION-2 population had higher proportions of Asian and Black subjects and a lower proportion of White subjects compared with the CANTATA-D study, and the proportion of male subjects was slightly higher in DURATION-2. In addition, baseline HbA1c values were lower in CANTATA-D than in DURATION-2; this may have been due to the higher mean baseline metformin dose in CANTATA-D (2146–2190 mg/day in each group) compared with DURATION-2 (1480–1583 mg/day in each group). Mean baseline BMI and mean baseline IWQoL-Lite scores were similar in the two studies. The LOCF imputation was applied to similar proportions of subjects in both trials because of early discontinuation (8% in CANTATA-D, and between 5% and 14% in DURATION-2) [12, 32].

As shown in Table 2, approximately half of the subjects in the CANTATA-D study met all of the DURATION-2 inclusion and exclusion criteria. The most common inclusion and exclusion criteria from the DURATION-2 study that were not met by subjects in the CANTATA-D trial were prior exposure to sulfonylurea, weight fluctuations prior to the start of the study ($> 3\%$ weight change from screening to randomization), and restrictions on baseline BMI (< 25 or > 45 kg/m²) and HbA1c (7.1–11.0%). Subjects in the CANTATA-D trial were also eliminated from the “CANTATA-D D2 population” if they had comorbidities that excluded subjects from enrollment in DURATION-2.

Table 1 Comparison of trial baseline characteristics for DURATION-2 and CANTATA-D

Variable ^a	DURATION-2	CANTATA-D			
	LOCF imputation (<i>N</i> = 491)	Overall population LOCF imputation (<i>N</i> = 1020)	Complete case (no imputation) (<i>n</i> = 934)	D2 population LOCF imputation (<i>n</i> = 516)	Complete case (no imputation) (<i>n</i> = 479)
Age, years	52.3 ± 10.3	55.5 ± 9.3	55.6 ± 9.2	55.1 ± 9.1	55.2 ± 9.1
Gender, <i>n</i> (%)					
Female	237 (48.3)	543 (53.2)	496 (53.1)	277 (53.7)	255 (53.2)
Male	254 (51.7)	477 (46.8)	438 (46.9)	239 (46.3)	224 (46.8)
Race (DURATION-2 categories), <i>n</i> (%) ^b					
Asian	119 (24.2)	136 (13.3)	118 (12.6)	45 (8.7)	40 (8.4)
Black	52 (10.6)	39 (3.8)	30 (3.2)	13 (2.5)	10 (2.1)
Hispanic	143 (29.1)	294 (28.8)	282 (30.2)	170 (32.9)	162 (33.8)
White	168 (34.2)	551 (54.0)	504 (54.0)	288 (55.8)	267 (55.7)
Baseline weight, kg	88.0 ± 20.0	87.3 ± 21.8	86.9 ± 21.2	88.3 ± 18.3	88.3 ± 18.3
Baseline BMI, kg/m ²	32.0 ± 5.3	31.9 ± 6.3	31.8 ± 6.0	32.3 ± 4.7	32.3 ± 4.7
Baseline HbA1c, %	8.5 ± 1.2	7.9 ± 0.9	7.9 ± 0.9	7.9 ± 0.9	7.9 ± 0.9
Duration of T2DM, years	5.7 ± 4.7	6.9 ± 5.3	6.8 ± 5.3	6.1 ± 4.7	6.1 ± 4.7
Baseline metformin dose, mg	1523 ± 551	2148 ± 337	2154 ± 335	2171 ± 337	2175 ± 331
Baseline systolic BP, mmHg	126 ± 14	128 ± 13	128 ± 13	128 ± 13	129 ± 13
Proportion of subjects from countries in DURATION-2 (USA, Mexico, or India), <i>n</i> (%)	491 (100)	426 (41.8)	385 (41.2)	173 (33.5)	161 (33.6)
Frequency by country, <i>n</i> (%)					
India	Unknown	68 (6.7)	62 (6.6)	27 (5.2)	24 (5.0)
Mexico	Unknown	59 (5.8)	56 (6.0)	37 (7.2)	37 (7.7)
USA	Unknown	299 (29.3)	267 (28.6)	109 (21.1)	100 (20.9)
Baseline IWQoL-Lite total score	80.2 ^c	79.4 ± 17.7	79.4 ± 17.8	78.2 ± 17.2	78.01 ± 17.4

LOCF last observation carried forward, BMI body mass index, T2DM type 2 diabetes mellitus, BP blood pressure, IWQoL Impact of Weight on Quality of Life, SD standard deviation

^a Data are mean ± SD unless otherwise noted

^b Percentages may not total 100.0% due to rounding

^c Baseline SD not reported

Table 2 Common reasons for failure of DURATION-2 inclusion and exclusion criteria

	CANTATA-D	
	LOCF imputation (<i>N</i> = 1020)	Complete case (no imputation) (<i>n</i> = 934)
Fails DURATION-2 exclusion criteria, <i>n</i> (%)	504 (49.4)	455 (48.7)
Eligible for DURATION-2, <i>n</i> (%)	516 (50.6)	479 (51.3)
Any medication history exclusion	338 (33.1)	307 (32.9)
History of insulin use	43 (4.2)	40 (4.3)
History of SU or MET + SU strata	248 (24.3)	223 (23.9)
History of systemic steroid use	31 (3.0)	29 (3.1)
History of drug interacting with CYP2C8 enzyme system	34 (3.3)	31 (3.3)
Weight change > 3% from screening to randomization	169 (16.6)	148 (15.8)
Screening BMI < 25 or > 45 kg/m ²	136 (13.3)	119 (12.7)
Exclusion based on either blood glucose criteria	106 (10.4)	97 (10.4)
HbA1c < 7.1% or > 11.0% at screening	94 (9.2)	85 (9.1)
FPG ≥ 15.5 mmol/L at screening	12 (1.2)	12 (1.3)

LOCF last observation carried forward, *SU* sulfonylurea, *MET* metformin, *BMI* body mass index, *HbA1c* glycated hemoglobin, *FPG* fasting plasma glucose, *ANCOVA* analysis of covariance, *IWQoL* Impact of Weight on Quality of Life ANCOVA results from CANTATA-D, unadjusted and adjusted for baseline *IWQoL*-Lite total score, country, and race. Medically significant conditions were an exclusion criterion in both studies, and were therefore assumed to be similar.

Changes in IWQoL-Lite Total Score

Treatment estimates for DURATION-2 and CANTATA-D under different population assumptions (overall and CANTATA-D D2), imputation assumptions (LOCF and complete case), and adjustments for the differences in race and country distributions (adjusted and unadjusted) are summarized in Table 3. As shown, imputation strategy made little difference in mean change in IWQoL-Lite total score. However, the magnitude of the least squares mean change in IWQoL-Lite total score was smaller in the “CANTATA-D D2 population” compared with the overall population (2.3 versus 2.6 for canagliflozin, and 1.1 versus 1.6 for sitagliptin). This observation may have been a result of differences in the inclusion criteria of the two trials (Table 2). Adjustment for country and race distributions increased the magnitude

of the estimates compared with the unadjusted models.

Indirect Comparison Results

Figure 3 shows the results of the primary analysis. The probability that treatment with canagliflozin results in greater improvements in IWQoL-Lite total score than treatment with exenatide QW, sitagliptin, and pioglitazone was 60.0%, 89.9%, and 99.5%, respectively.

When analyses were restricted to the DURATION-2 inclusion and exclusion criteria, the probability that canagliflozin is better than exenatide QW, sitagliptin, and pioglitazone was 61.6%, 82.1%, and 98.9%, respectively. Further adjustment for race and country differences yielded probabilities of 74.6% versus exenatide QW, 85.9% versus sitagliptin, and 98.5% versus pioglitazone. Median differences in IWQoL-Lite

Table 3 ANCOVA results in DURATION-2 and CANTATA-D: mean change in IWQoL-Lite total score

Trial	Population	Imputation method	N	Treatment	Unadjusted		Adjusted	
					LSM (95% CI)	Difference versus SITA	LSM (95% CI)	Difference versus SITA
DURATION-2	LOCF		132	EXEN QW	5.15 (3.11, 7.19)	0.59	–	–
	imputation		130	PIO	1.20 (– 0.87, 3.28)	– 3.36	–	–
			139	SITA	4.56 (2.56, 6.57)	–	–	–
CANTATA-D	Overall	LOCF	688	Pooled CANA	2.60 (1.71, 3.49)	1.02	3.42 (2.03, 4.82)	1.44
	population	imputation	332	SITA 100 mg	1.58 (0.30, 2.87)	–	1.98 (– 0.08, 4.04)	–
		CC (no imputation)	634	Pooled CANA	2.57 (1.63, 3.52)	1.10	3.42 (1.92, 4.93)	1.81
			300	SITA 100 mg	1.47 (0.08, 2.86)	–	1.61 (– 0.66, 3.87)	–
	D2 population	LOCF	344	Pooled CANA	2.31 (0.92, 3.71)	1.17	4.24 (1.88, 6.67)	2.36
		imputation	172	SITA 100 mg	1.14 (– 0.93, 3.2)	–	1.88 (– 1.62, 5.39)	–
	CC (no imputation)	318	Pooled CANA	2.09 (0.62, 3.57)	1.03	4.13 (1.54, 6.72)	2.55	
		161	SITA 100 mg	1.06 (– 1.14, 3.25)	–	1.58 (– 2.37, 5.53)	–	

ANCOVA analysis of covariance, *IWQoL* Impact of Weight on Quality of Life, *LSM* least squares mean, *CI* confidence interval, *SITA* sitagliptin, *LOCF* last observation carried forward, *EXEN* exenatide, *QW* once weekly, *PIO* pioglitazone, *CANA* canagliflozin, *CC* complete case

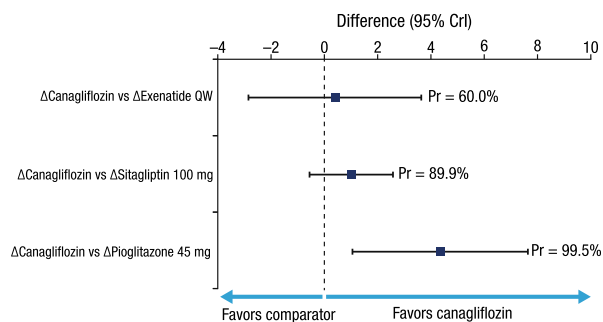


Fig. 3 Change in IWQoL-Lite total score from baseline to Week 26.* *IWQoL* Impact of Weight on Quality of Life, *CrI* credible interval, *Pr* probability. *Reported as median differences and 95% CrI for canagliflozin versus exenatide QW, sitagliptin, and pioglitazone

total score changes and the probability that canagliflozin outperforms exenatide QW, sitagliptin, and pioglitazone for this PRO measure were computed for the sensitivity analyses compared with the primary analysis. Comparisons between the primary analysis and sensitivity analyses with canagliflozin versus exenatide QW for different populations and imputation methods are shown in Table 4. In all cases, these differences were larger in the complete case analysis than in the LOCF analysis. Estimated differences were also higher in analyses that adjusted for distribution by country and race, which highlights the need to consider geographic and cultural differences that may affect PROs (e.g., differences in perceptions of body weight). The largest relative differences,

based on the Bayesian models, were seen when restricting the analysis to the “CANTATA-D D2 population”. Collectively, these findings suggest that differences in trial inclusion and exclusion criteria can impact PRO findings.

DISCUSSION

Weight management is an important component of T2DM management, and weight loss is associated with better health-related quality of life among people with T2DM [36, 37]. In the DURATION-2 study, mean body weight changes from baseline to week 26 were - 2.3 kg with exenatide QW, - 0.8 kg with sitagliptin, and + 2.8 kg with pioglitazone [31]. In the CANTATA-D study, body weight changes from baseline to week 26 were - 3.3 kg with canagliflozin 100 mg, - 3.6 kg with canagliflozin 300 mg, and - 1.1 kg with sitagliptin [12]. The current analysis evaluated changes in the impact of weight on health-related quality of life based on PRO data from these clinical studies.

Findings from this analysis demonstrated that, under a variety of assumptions, canagliflozin was associated with greater improvements in IWQoL-Lite total score compared with exenatide QW, sitagliptin, and pioglitazone. This is noteworthy because baseline IWQoL-Lite total scores in both studies were already relatively high, ranging from about 78 to 80 out of a possible 100 points [34].

Table 4 Comparison of primary and sensitivity analyses: differences in IWQoL-Lite total scores at week 26 with canagliflozin versus exenatide QW

Imputation	Population, adjustment	Median (95% CrI)	Probability (%)
LOCF	Overall, unadjusted	0.43 (- 2.84, 3.62)	60.0
	Overall, adjusted	0.84 (- 2.97, 4.55)	66.7
	CANTATA-D D2, adjusted	1.75 (- 3.40, 6.77)	74.6
Complete case (no imputation)	Overall, unadjusted	0.51 (- 2.82, 3.76)	61.7
	Overall, adjusted	1.22 (- 2.76, 5.09)	72.6
	CANTATA-D D2, adjusted	1.95 (- 3.64, 7.37)	75.2

IWQoL Impact of Weight on Quality of Life, *QW* once weekly, *CrI* credible interval, *LOCF* last observation carried forward

This study addresses considerations for the performance of indirect comparisons of PRO data in the setting of T2DM, which may play a role in health-care decision-making in the future, as health-care decision-makers increasingly acknowledge the importance of such outcomes as part of a range of treatment benefits. Examining health-related quality of life outcomes with indirect comparison methods should be considered a viable source of comparative information to aid in these decisions when data from head-to-head trials are unavailable. However, there are unique methodological challenges associated with indirect comparisons of health-related quality of life data. For example, data from indirect comparisons of PROs should be interpreted in the context of the study design features (i.e., parameters are sensitive to assumptions about populations, imputation strategies, and certain baseline characteristics). Unlike clinical variables, scores for PRO measures may be influenced by cultural differences (e.g., perceptions of body weight); therefore, determinants of health-related quality of life may be different than those used to describe clinical efficacy. The DURATION-2 inclusion criteria provided a much more homogeneous study population than the CANTATA-D study, such that roughly half of the subjects enrolled in CANTATA-D would have been ineligible to participate in DURATION-2. These differences were a necessary part of the trial designs; i.e., the inclusion criteria for DURATION-2 were set such that a sample size of 500 subjects was sufficiently powered to detect significant differences in HbA1c between treatment groups [31], whereas the broader CANTATA-D inclusion criteria required a sample size of 1260 (360 in each of three active treatment arms plus 180 in the placebo group) to be sufficiently powered to detect significant differences in HbA1c between groups [12].

Overall, findings from the current study support the importance of geographic and cultural factors in perceptions of body weight, since differences in IWQoL-Lite total score were observed when subject characteristics were restricted to the “CANTATA-D D2 population”. Because cultural differences can impact PRO results, future studies should report geographic and cultural characteristics as important

demographic variables. Researchers should also consider how study inclusion and exclusion criteria may affect the range of included perspectives of subjects enrolled in studies, especially given the importance of patient-centered care when collaborating on individualized T2DM management strategies [1]. Furthermore, systematic differences in effect modifiers are important for indirect comparisons, but this report found consistent results for adjusted and unadjusted analyses, as well as subgroups.

Only one study (DURATION-2) was identified from extensive searches for reports of PRO measurements that could be compared with data from the canagliflozin clinical development program. The only common PRO measure identified in the searches that could be used in this analysis was the IWQoL-Lite, which was designed to capture the impact of weight on quality of life and does not measure other factors that might influence a patient’s experience with treatment (e.g., side effects/tolerability). The lack of available common PRO data is an important gap in the assessment of treatments for T2DM, especially considering the strong associations between patient perceptions and clinical outcomes in this disease [38, 39]. Because HbA1c control is the primary endpoint used by regulatory agencies to determine acceptability for approval, randomized controlled trials in T2DM are typically powered to detect differences in HbA1c and are often insufficiently powered to detect differences in PRO scores. In addition, clinical reports of randomized controlled trials in T2DM tend to capture and report biomarker data (e.g., HbA1c) in a similar way (e.g., mean change by treatment arm), but there is far greater variability in PRO instruments employed and treatment effect summaries. For example, it should be noted that canagliflozin trials did not utilize the commonly used Diabetes Treatment Satisfaction Questionnaire [40] because this instrument is focused on concepts such as treatment flexibility and convenience, which are of limited relevance for a once-daily oral medication such as canagliflozin. Overall, the infrequent and inconsistent reporting of PRO data compared with clinical data likely contributes to the low

number of relevant studies for inclusion in synthesis efforts such as this one.

PRO score changes with the treatments of interest in the current analysis may be either over- or underestimated, in part, because of the small number of studies or other factors, and a more robust analysis could be conducted if subject-level data were available for both trials. Models adjusted for population differences are an important step toward unbiased comparisons, but this method assumes that covariates such as baseline IWQoL-Lite score impact final changes in a similar way for canagliflozin and exenatide QW. Though this assumption is not unreasonable, an evidence-based defense would require subject-level data from DURATION-2. Many pharmaceutical companies are developing data-sharing policies that would facilitate subject-level analyses in the future, and, especially in areas like PROs where suitable comparisons are limited, would result in indirect comparison methods requiring fewer assumptions.

In this analysis, as in many studies, missing data can present an important challenge. DURATION-2 publications summarize baseline data for the modified intention-to-treat analysis, but not for the subgroup of subjects who were included in the PRO analyses. An implicit assumption, which cannot be tested, is that subjects with missing PRO data were similar in DURATION-2 and CANTATA-D. LOCF is a sub-optimal imputation method, but more appropriate techniques such as mixed models for repeated measures (MMRM) require more frequent data collection, which, in turn, requires a higher prioritization for PRO outcomes. Fortunately, in this analysis, the rate of missing data was relatively low for both studies.

CONCLUSIONS

Findings from this analysis suggest that improvements in the impact of weight on health-related quality of life may be greater with canagliflozin compared with exenatide QW, sitagliptin, and pioglitazone. The impact of treatment on various aspects of health-related quality of life should not be overlooked when selecting T2DM treatments, as improvements in health-related

quality of life are associated with increased motivation to perform self-care behaviors and more persistent and consistent performance of disease management behaviors [41–43].

ACKNOWLEDGEMENTS

Funding. This analysis was funded by Janssen Global Services, LLC, and was based, in part, on data from a study funded by Janssen Research & Development, LLC. Article processing charges were funded by Janssen Global Services, LLC. Medical writing support was provided by Cherie Koch, PhD, of MedErgy, and was funded by Janssen Global Services, LLC. Canagliflozin has been developed by Janssen Research & Development, LLC, in collaboration with Mitsubishi Tanabe Pharma Corporation.

Authorship. All named authors meet the International Committee of Medical Journal Editors criteria for authorship for this manuscript, take responsibility for the integrity of the work as a whole, and have given final approval of the version to be published.

Disclosures. Shana B. Traina is a full-time employee of Janssen Global Services, LLC. April Slee is a full-time employee of Axio Research, which has received payment from Janssen for statistical support of the analyses reported in this manuscript. Suzy Van Sanden is a full-time employee of Janssen Research & Development. Joris Diels is a full-time employee of Janssen Research & Development.

Compliance with Ethics. This article does not contain any new studies with human participants performed by any of the authors.

Data Availability. The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Open Access. This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/>

by-nc/4.0/), which permits any non-commercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

REFERENCES

- Inzucchi SE, Bergenstal RM, Buse JB, et al. Management of hyperglycemia in type 2 diabetes, 2015: a patient-centered approach: update to a position statement of the American Diabetes Association and the European Association for the Study of Diabetes. *Diabetes Care*. 2015;38:140–9.
- Food and Drug Administration. Guidance for industry diabetes mellitus: developing drugs and therapeutic biologics for treatment and prevention. 2008. <http://www.fda.gov/downloads/Drugs/Guidances/ucm071624.pdf>.
- Food and Drug Administration. Public workshop: diabetes outcome measures beyond hemoglobin A1c (HbA1c). 2016. <http://www.fda.gov/Drugs/NewsEvents/ucm499281.htm>.
- Traina SB, Slee A, Woo S, Canovatchel W. The importance of weight change experiences for performance of diabetes self-care: a patient-centered approach to evaluating clinical outcomes in type 2 diabetes. *Diabetes Ther*. 2015;6:611–25.
- Powers MA, Bardsley J, Cypress M, et al. Diabetes self-management education and support in type 2 diabetes: a joint position statement of the American Diabetes Association, the American Association of Diabetes Educators, and the Academy of Nutrition and Dietetics. *Diabetes Care*. 2015;38:1372–82.
- Traina S, Guthrie R, Slee A. The impact of weight loss on weight-related quality of life and health satisfaction: results from a trial comparing canagliflozin with sitagliptin in triple therapy among people with type 2 diabetes. *Postgrad Med*. 2014;126:7–15.
- Centers for Disease Control and Prevention. Prevalence of overweight and obesity among adults with diagnosed diabetes—United States, 1988–1994 and 1999–2002. *MMWR Morb Mortal Wkly Rep*. 2004;53:1066–8.
- Wing RR, Lang W, Wadden TA, et al. Benefits of modest weight loss in improving cardiovascular risk factors in overweight and obese individuals with type 2 diabetes. *Diabetes Care*. 2011;34:1481–6.
- Van Gaal L, Scheen A. Weight management in type 2 diabetes: current and emerging approaches to treatment. *Diabetes Care*. 2015;38:1161–72.
- Noria SF, Grantcharov T. Biological effects of bariatric surgery on obesity-related comorbidities. *Can J Surg*. 2013;56:47–57.
- Rosenthal N, Meininger G, Ways K, et al. Canagliflozin: a sodium glucose co-transporter 2 inhibitor for the treatment of type 2 diabetes mellitus. *Ann N Y Acad Sci*. 2015;1358:28–43.
- Lavalle-González FJ, Januszewicz A, Davidson J, et al. Efficacy and safety of canagliflozin compared with placebo and sitagliptin in patients with type 2 diabetes on background metformin monotherapy: a randomised trial. *Diabetologia*. 2013;56:2582–92.
- Scherthaner G, Gross JL, Rosenstock J, et al. Canagliflozin compared with sitagliptin for patients with type 2 diabetes who do not have adequate glycemic control with metformin plus sulfonylurea: a 52-week, randomized trial. *Diabetes Care*. 2013;36:2508–15.
- Cefalu WT, Leiter LA, Yoon KH, et al. Efficacy and safety of canagliflozin versus glimepiride in patients with type 2 diabetes inadequately controlled with metformin (CANTATA-SU): 52 week results from a randomised, double-blind, phase 3 non-inferiority trial. *Lancet*. 2013;382:941–50.
- Mearns ES, Sobieraj DM, White CM, et al. Comparative efficacy and safety of antidiabetic drug regimens added to metformin monotherapy in patients with type 2 diabetes: a network meta-analysis. *PLoS One*. 2015;10:e0125879.
- Zaccardi F, Webb DR, Htike ZZ, Youssef D, Khunti K, Davies MJ. Efficacy and safety of sodium-glucose cotransporter 2 inhibitors in type 2 diabetes mellitus: systematic review and network meta-analysis. *Diabetes Obes Metab*. 2016;18:783–94.
- Shyangdan DS, Uthman OA, Waugh N. SGLT-2 receptor inhibitors for treating patients with type 2 diabetes mellitus: a systematic review and network meta-analysis. *BMJ Open*. 2016;6:e009417.
- Jansen JP, Fleurence R, Devine B, et al. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. *Value Health*. 2011;14:417–28.
- Hoaglin DC, Hawkins N, Jansen JP, et al. Conducting indirect-treatment-comparison and network-meta-analysis studies: report of the ISPOR Task Force on Indirect Treatment Comparisons

- Good Research Practices: part 2. *Value Health*. 2011;14:429–37.
20. Dias S, Welton NJ, Sutton AJ, et al. NICE DSU technical support document 2: a generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials. 2012. Updated April 2012. <http://www.nicesdu.org.uk>.
 21. Haute Autorité de Santé. Summary report: indirect comparisons—methods and validity. Saint-Denis La Plaine: Haute Autorité de Santé; 2009.
 22. Jansen JP, Trikalinos T, Cappelleri JC, et al. Indirect treatment comparison/network meta-analysis study questionnaire to assess relevance and credibility to inform health care decision making: an ISPOR-AMCP-NPC Good Practice Task Force report. *Value Health*. 2014;17:157–73.
 23. Jansen JP, Crawford B, Bergman G, Stam W. Bayesian meta-analysis of multiple treatment comparisons: an introduction to mixed treatment comparisons. *Value Health*. 2008;11:956–64.
 24. Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol*. 2011;64:163–71.
 25. Alonso J, Bartlett SJ, Rose M, et al. The case for an international patient-reported outcomes measurement information system (PROMIS(R)) initiative. *Health Qual Life Outcomes*. 2013;11:210.
 26. Reaney M, Gwaltney C. Measuring and interpreting patient-reported outcome data from clinical trials of diabetes medication. *J Diabetes Res Clin Metab*. 2014. <https://doi.org/10.7243/2050-0866-3-7>.
 27. Brundage M, Blazeby J, Revicki D, et al. Patient-reported outcomes in randomized clinical trials: development of ISOQOL reporting standards. *Qual Life Res*. 2013;22:1161–75.
 28. Reaney M, Celeste Elash CA. Patient reported outcomes in the development of new medications for type 2 diabetes: a review of recently approved products. *Value Health*. 2015;18:A715.
 29. Puhan MA, Soesilo I, Guyatt GH, Schunemann HJ. Combining scores from different patient reported outcome measures in meta-analyses: when is it justified? *Health Qual Life Outcomes*. 2006;4:94.
 30. Johnston BC, Patrick DL, Busse JW, Schunemann HJ, Agarwal A, Guyatt GH. Patient-reported outcomes in meta-analyses—part 1: assessing risk of bias and combining outcomes. *Health Qual Life Outcomes*. 2013;11:109.
 31. Bergenstal RM, Wysham C, MacConell L, et al. Efficacy and safety of exenatide once weekly versus sitagliptin or pioglitazone as an adjunct to metformin for treatment of type 2 diabetes (DURATION-2): a randomised trial. *Lancet*. 2010;376:431–9.
 32. Best JH, Rubin RR, Peyrot M, et al. Weight-related quality of life, health utility, psychological well-being, and satisfaction with exenatide once weekly compared with sitagliptin or pioglitazone after 26 weeks of treatment. *Diabetes Care*. 2011;34:314–9.
 33. Traina SB, Ho KF. Weight- and health-related quality of life with canagliflozin versus sitagliptin in subjects with type 2 diabetes mellitus on background metformin. *Value Health*. 2013; 6:A448
 34. Kolotkin RL, Crosby RD, Kosloski KD, Williams GR. Development of a brief measure to assess quality of life in obesity. *Obes Res*. 2001;9:102–11.
 35. Kolotkin RL, Crosby RD, Williams GR. Health-related quality of life varies among obese subgroups. *Obes Res*. 2002;10:748–56.
 36. Grandy S, Fox KM, Bazata DD. Association of self-reported weight change and quality of life, and exercise and weight management behaviors among adults with type 2 diabetes mellitus: the SHIELD study. *Cardiol Res Pract*. 2012;2012:892564.
 37. Rubin RR, Wadden TA, Bahnson JL, et al. Impact of intensive lifestyle intervention on depression and health-related quality of life in type 2 diabetes: the Look AHEAD Trial. *Diabetes Care*. 2014;37:1544–53.
 38. Allison DB, Downey M, Atkinson RL, et al. Obesity as a disease: a white paper on evidence and arguments commissioned by the Council of the Obesity Society. *Obesity (Silver Spring)*. 2008;16:1161–77.
 39. Gregory RS, Handelsman Y, Pezalla EJ, Pikelny D. Practical perspectives on the management of overweight and obesity. *Am J Manag Care*. 2014;20:S64–75.
 40. Bradley C. Diabetes Treatment Satisfaction Questionnaire (DTSQ). In: Bradley C, editor. *Handbook of psychology and diabetes*. London: Gordon & Breach; 1994. p. 111–32.
 41. Pi-Sunyer FX. The impact of weight gain on motivation, compliance, and metabolic control in patients with type 2 diabetes mellitus. *Postgrad Med*. 2009;121:94–107.
 42. Traina S, Weiser J, Katz EG, Palak A. Satisfaction with health and motivation to follow a diabetes diet and exercise vary by weight status among people with type 2 diabetes mellitus (T2DM).

-
- In: Poster presented at: 71st scientific sessions of the American diabetes association (ADA); June 24–28, San Diego, CA. Poster 858-P; 2011.
43. Traina S, Johnson P, Bancroft T, Watkins A, Adams S. Motivation to perform diabetes self-care activities and satisfaction with health vary by weight status among people living with type 2 diabetes mellitus (T2DM). In: Poster presented at: 72nd Scientific Session of the American Diabetes Association (ADA); June 8–12, Philadelphia, PA. Poster 813-P; 2012.