CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

209210Orig1s000

STATISTICAL REVIEW(S)

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA #:	NDA 209210
Drug Name:	BYDUREON Bcise® (Exenatide extended-release injectable suspension) 2mg
Indication(s):	Glycemic control in adults with T2DM
Applicant:	AstraZeneca
Date(s):	Dec. 21. 2016 (Receipt date) Sep. 16. 2017 (Primary review completion date) Oct. 21. 2017 (PDUFA goal date)
Review Priority:	Standard
Biometrics Division:	DB II
Statistical Reviewer:	Yoonhee Kim, Ph.D, M.P.H
Concurring Reviewers:	Yun Wang, Ph.D
Medical Division:	DMEP
Clinical Team:	Mahtab Niyyati, MD Lisa Yanoff, MD
Project Manager:	Richard White

Keywords: Active control, Non-inferiority, Missing data, Sensitivity analysis, Trimmed means

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1 EXECUTIVE SUMMARY

1.1 Introduction

This statistical review evaluates original new drug application (NDA209210) by AstraZeneca for exenatide extended-release injectable suspension (proprietary name: BYDUREON BCise®). The proposed indication for BYDUREON BCise® in this review is as an adjunct to diet and exercise to improve glycemic control in adults with type 2 diabetes mellitus (T2DM) by once weekly administration, i.e. exenatide once-weekly suspension (hereafter; EQWS). EQWS contains the same active ingredient (i.e. exenatide) as the commercial products BYETTA (exenatide; NDA021773) and BYDUREON (exenatide extended-release; NDA022220). As an extension to the approved BYETTA (exenatide twice daily aqueous solution) and BYDUREON (exenatide once weekly aqueous suspension), EQWS in the current submission is a subcutaneously injectable, extended-release non-aqueous suspension formulation with once weekly dose of 2mg.

1.2 Brief Overview of Clinical Studies

This statistical review focuses on two Phase III clinical studies included in the submission: BCB118 and BCB120. Both studies used non-inferiority (NI) study design with 2 treatment arms and 3 treatment arms for BCB118 and BCB120 respectively. BCB118 is a randomized (EQWS:BYETTA; 3:2), open-label, long-term, parallel-group, comparator-controlled, multicenter study to compare the glycemic effects, safety, and tolerability of EQWS to BYETTA in subjects with T2DM. BCB120 is a randomized (EQWS:Sitagliptin:Placebo;3:2:1), long-term, open-label, 3-arm, multicenter study to compare the glycemic effects, safety, and tolerability of EQWS to Sitagliptin (active control) and placebo in subjects with T2DM.

1.3 Conclusion and Recommendations

The statistical reviewer recommends **approval** of the indication claimed in this NDA209210 application on condition of substantial changes of wording and the presentation of study results from both clinical studies on the proposed labeling (see section 5.3). This conclusion is based on collective evidence from the following statistical review issues and findings.

1.4 Statistical Issues and Findings

- **De-facto estimand**: De-facto estimand is defined to assess the study drug efficacy by measuring the values (endpoints) as actually taken in the intent-to-treat (ITT) population, i.e. randomized population notwithstanding the discontinuation of treatment. To evaluate the most likely true efficacy of EQWS in a regulatory setting, de-facto estimand is appropriate for estimating potential outcomes that would have been observed under different exposures such as: regardless of adherence to treatment, use of rescue therapy or discontinuation of study drug in the T2DM patients group. However, the applicant used on-treatment population (de-jure estimand) that excluded the subjects who used rescue therapy and/or discontinued study drug. Also, this clinical trial was not designed for continued follow-up of patients; thus,

the retrieved dropouts were not sufficient for imputing the high percentage of missing values in the data.

- **Statistical analysis methods with high missing rates in imbalanced randomization ratio:** EQWS is a reformulated drug by changing the drug's vehicle from aqueous format to non-aqueous format in an aqueous diluent. Thus, the purpose of submitted clinical studies was to establish the evidence that this new regimen EQWS is beneficial over a long period in the target population. In such a case, patients' dropout for different reasons in the treatment arm, comparator arm and/or placebo arm could depend on the drug effect. Indeed, almost double patients were dropped out from the placebo arm (38%) than the EQWS arm (22%) in study BCB120. Because high dropout rates are limitations in these studies, statistical analyses to deal with missing data using the best applicable estimand in T2DM clinical setting were critical to achieve the evidence for supporting the proposed indication. Nonetheless, the applicant used Mixed Model for Repeated Measures (MMRM) as a primary statistical analysis excluding post-rescued and/or discontinued observations and Last Observation Carried Forward (LOCF) as a sensitivity analysis: neither of them is sufficient in demonstrating the long-term benefits of EQWS in the ITT population.

In this review, to address abovementioned concerns given the limitation in the study design, the efficacy of EQWS was evaluated using de-facto estimand including post-rescued and/or discontinued observations for a primary analysis (ANCOVA model) and the primary results were validated through multiple sensitivity analyses including multiple imputation of pattern mixture model (PMM), and Trimmed Means method that estimates direct effects of EQWS (details in section 3.2.2) in a consideration of high and imbalanced dropout rates.

In totality, two clinical studies provided supportive evidence for the purported indication. In BCB118, the upper limit of 95% confidence interval for the difference between EQWS and BYETTA was below zero that supported the non-inferiority with the pre-specified non-inferiority (NI) margin of 0.4% as well as the superiority to BYETTA. In BCB120, EQWS was superior to placebo and showed non-inferior to Sitagliptin with the pre-specified 0.3% NI margin (Table 1).

Table 1. Summary of Trial Findings (primary efficacy analysis results from ANCOVA model in de-facto estimand)

Study	Comparison	Difference in LS mean changes in HbA1C from baseline to week 28 (%)	95% Confidence Interval (Lower, Upper) (%)
BCB118	EQWS 2mg vs. BYETTA 10 mcg	-0.36	-0.66, -0.14
BCB120	EQWS 2mg vs. Placebo	-0.49	-0.91, -0.07
BCB120	EQWS 2mg vs. Sitagliptin	-0.28	-0.62, 0.02

Source: Statistical reviewer's analysis

2 INTRODUCTION

2.1 Overview

2.1.1 Class and Indication

Exenatide is a glucagon-like peptide-1 (GLP-1) receptor agonist. Exenatide once-weekly suspension (EQWS; proprietary name BYDUREON BCise) is a subcutaneously injectable, extended-release non-aqueous suspension formulation that was developed as an extension to the currently marketed BYETTA (exenatide twice daily) and BYDUREON (exenatide once-weekly aqueous suspension) product line for the treatment of patients with T2DM. EQWS contains the same drug substance and extended release microspheres as BYDUREON (an aqueous formulation), but with a non-aqueous medium chain triglyceride (MCT) vehicle for use with an autoinjector. The proposed indication of EQWS is as an adjunct to diet and exercise to improve glycemic control in adults with type 2 diabetes mellitus, which is the same as the approved indication for BYDUREON. The proposed dose for EQWS is 2mg once weekly.

2.1.2 History of Drug Development

BYETTA was approved in the United States (US) on 28 April 2005 and in Europe on 20 November 2006 to improve glycemic control in adults with T2DM. BYETTA is administered twice daily (here after: BID) as a SC injection of 5µg or 10µg within 60 minutes before the 2 main meals of the day. BYDUREON was approved by the European Commission on 21 June 2011 and by the FDA on 27 January 2012 as an adjunct to diet and exercise to improve glycemic control in adults with T2DM. BYDUREON is administered once weekly and is provided as a powder to be combined immediately before dosing with an aqueous vehicle to form a suspension for injection.

The new extended-release non-aqueous suspension formulation (EQWS) was developed to simplify the administration steps for patients (users) using auto injector that was possible due to the non-aqueous vehicle in the aqueous diluent. In the current application, the applicant submitted three clinical studies; one phase II study (BCB110), and two Phase III studies (BCB118 and BCB120) for EQWS on December 2016.

2.1.3 Studies Reviewed

In this statistical review, two Phase III studies (BCB118 and BCB120) were selected for full statistical review and evaluation for the efficacy of EQWS to support the labeling in section 14. Table 2 is an overview of key factors of studies.

Table 2. Overview of Clinical Studies used in the Analysis for Statistical Review

	BCB118	BCB120
Title	A randomized, open-label, long-term, parallel-group comparator-controlled, multicenter study to compare the	A randomized, long-term, open-label, 3 arm multicenter study to compare the glycemic effects, safety, and tolerability of exenatide

	glycemic effects, safety and tolerability of exenatide once weekly suspension to exenatide twice daily in subjects with type 2 diabetes mellitus	once weekly suspension to Sitagliptin and placebo in subjects with type 2 diabetes mellitus
Study period	28-Jan-2013 to 19-Aug-2014	08-Feb-2013 to 04-Apr-2014
Study design and period	Phase III Randomized, open-label, multicenter, comparator-controlled 2-arm non-inferiority study 28 weeks + 24weeks (extension)	Phase III Randomized, open-label (oral agents blinded), multicenter, comparator- and placebo-controlled 3-arm non-inferiority study 28 weeks
Study population	Subjects with T2DM with diet and exercise alone or in combination with a stable regimen of oral antidiabetic medication, including metformin, sulfonylurea, pioglitazone, or a combination of any 2 of these agents	Subjects with T2DM with inadequate glycemic control while taking ≥ 1500mg metformin daily
Treatment groups and sample size	**EQWS 2mg**: 229subjects **BID(BYETTA) 5µg**: 148 subjects 3:2 randomization ratio stratified by diabetes management method at screening (diet/exercise alone, SU use, or non-SU use), screening HbA1c stratum (<9% or ≥9%), and renal function (normal, mild renal impairment, or moderate renal impairment)	**EQWS 2mg**:182 subjects **Sitagliptin 100mg**: 122 subjects **Placebo**: 61 subjects 3:2:1 randomization stratified by screening HbA1c stratum (<9% or ≥9%)
Background therapy	Diet/Exercise, metformin, SU, pioglitazone, or a combination	Metformin
Regimens	**28 weeks**: Subcutaneous (SC) administration of a 2mg dose of EQWS (Group A), or a 10 µ dose of exenatide BID (5µg for 4 weeks followed by 10 µg for the remaining 24 weeks of the controlled treatment period; Group B) **+ 24 weeks**: SC EQWS 2mg for all subjects (Groups A and B)	Subcutaneous (SC) administration of a 2 mg dose of EQWS once weekly, Sitagliptin 100mg by mouth once daily in the morning, or placebo by mouth once daily in the morning
Primary objectives	To compare the effect on glycemic control (HbA1c) of EQWS to that achieved by exenatide administered twice daily for 28 weeks in subjects with T2DM	To compare the effect on glycemic control (glycosylated hemoglobin, HbA1c) of EQWS to that achieved by Sitagliptin or placebo administered once daily for 28 weeks in subjects with T2DM
Primary endpoints	Change in HbA1c from baseline to Week 28	Change in HbA1c from baseline to Week 28

2.2 Data Sources

The applicant submitted materials for this review including data and clinical study report electronically in the electronic common technical document (eCTD) form and archived under the network path location: \\CDSESUB1\evsprod\NDA209210\209210.enx.

Data sets analyzed for this review were module 5 under EDR location: \\CDSESUB1\evsprod\NDA209210\0000

For BCB118, m5\datasets\d5553c00006-mb001003-bcb118\analysis\adam\dataset-interim in ADAM format and m5\datasets\d5553c00006-mb001003-bcb118\tabulations\sdtm-interim in SDTM format.

For BCB120, m5\datasets\d5553c00007-mb001-004-bcb120\analysis\adam\datasets in ADAM format and m5\datasets\d5553c00007-mb001-004-bcb120\tabulations\sdtm in SDTM format.

The original submitted data in BCB118 included duplicated subject pairs as independent subjects that must not be considered as a unique individual in the statistical analysis (details in section 3.2.3). The sponsor resubmitted corrected data set under EDR location: \\CDSESUB1\evsprod\NDA209210\0010.

Analysis ready format including primary endpoint (HbA1c) was adanlzl.xpt and including covariates per subject level was adsl.xpt under ADAM format EDR location for each study. Additional information and analysis results from the sponsor regarding sensitivity analyses and statistical analysis codes were found under following EDR locations:
\\CDSESUB1\evsprod\NDA209210\0003 (SAS programs for MMRM, GLM, CMH tests)
\\CDSESUB1\evsprod\NDA209210\0008 (SAS programs for ISE analysis)
\\CDSESUB1\evsprod\NDA209210\0010 (corrected ITT population flag for statistical analysis and post-hoc statistical analysis results (06 March 2017))

This reviewer ran independent coding for the descriptive statistics, plots and sensitivity analyses using R language.

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

Clinical study reports and the submitted data had average quality because there was inconsistency of primary results between clinical study report for BCB118 and integrated summary of efficacy report. By information requests from the agency, the applicant clarified that the interim datasets (short-term database locked at Week 28) were used for the statistical analysis for labeling claim. Also, the applicant removed six subjects from the analysis datasets: one miscoded EQWS subject (68-68022 miscoded as 67-68022), four subjects enrolled twice in different treatment arms (8-8006, 16-16006, 21-21010, and 50-50004), and one EQWS subject

who did not have baseline measurement (21-21007). The applicant submitted the updated primary analysis results using clarified data set. For BCB120, there was no issue about the data quality. This statistical reviewer was able to reproduce the applicant's primary efficacy results presented in the individual clinical study report and the applicant's responses in BCB120 as well as BCB118 with clarified data sets.

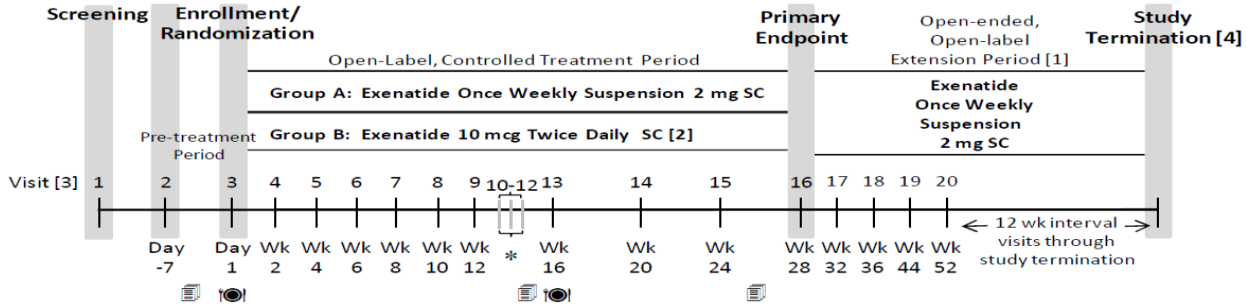
3.2 Evaluation of Efficacy

3.2.1 Study Design and Endpoints

Design features of two trials reviewed in this submission are displayed above in Table 2. Study design schematics are shown in Figures 1 and 2. Primary endpoint was change in HbA1c (%) from baseline to Week 28 in both studies. Key secondary endpoints were binary outcome percentage of patients with HbA1c < 7%, fasting plasma glucose (FPG), body weight changes (kg), and 2-hour postprandial plasma glucose at Week 28 in both studies.

Both studies used NI study design with active control (BID) arm in BCB118 and active control (Sitagliptin) and placebo arms in BCB120. Assay sensitivity is an essential property of an NI clinical trial. Assay sensitivity[1] is based on 1) historical evidence of sensitivity to drug effects, 2) similarity of the new NI trial to the historical trials (the constancy assumption), and 3) the quality of the new trial to rule out defects that would tend to minimize differences between treatments.

Figure 1. Study Design for BCB118

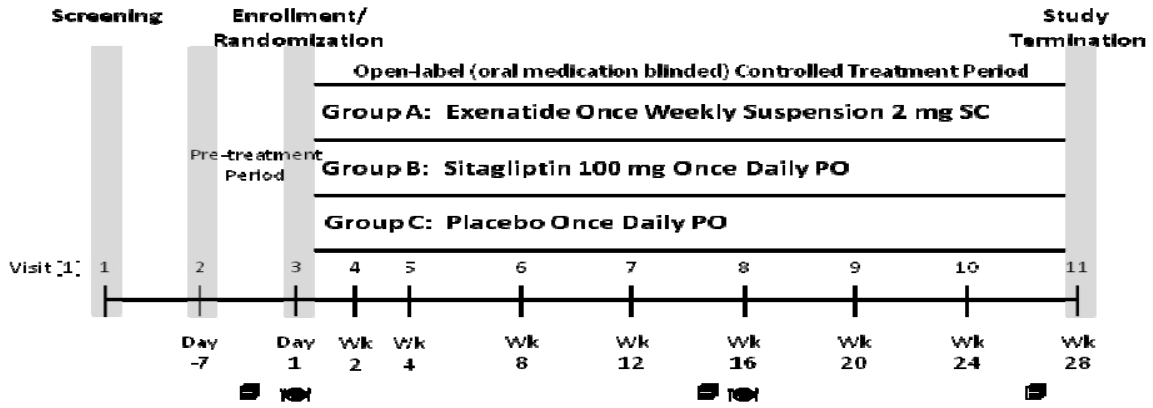


6 point self-monitored blood glucose profiles performed on any three days in week prior to subsequent visit.
Indicates meal test and postprandial assessments for a subset of subjects at select sites.
* Study site visits for PK assessments for a subset of subjects at select study sites.
[1] For the open-ended extension period, all subjects (Groups A and B) will receive 2 mg exenatide once weekly suspension treatment to Week 52.
[2] Subjects randomized to Group B will receive 4 weeks of 5 mcg exenatide BID followed by 24 weeks of 10 mcg exenatide BID.
[3] Visit 2 will occur within 14 days following Visit 1. Visit 3 will occur 7 (±2) days following Visit 2. Visits through Visit 20 will occur at the indicated interval (± 2 days), relative to Visit 3. Subsequent visits will occur at the indicated interval (± 3 days), relative to Visit 3.
[4] Subjects must complete a follow-up visit 10 weeks after the final treatment visit (whether the visit occurred during the controlled treatment period or the extension period).

Source: Sponsor's clinical study report (CSR page 36)

[1] Non-Inferiority Clinical Trials to Establish Effectiveness (Guidance for Industry) https://www.fda.gov/downloads/Drugs/Guidances/UCM202140.pdf

Figure 2. Study Design for BCB120



6 point self-monitored blood glucose profiles performed on any three days in the week prior to subsequent visit.

Indicates meal test and postprandial assessments for a subset of subjects at select sites.

PO = by mouth, SC = subcutaneous, Wk = week.

[1] Visit 2 will occur within 14 days following Visit 1. Visit 3 will occur 7 (± 2) days following Visit 2. Visits 4 through study termination will occur at the indicated interval (± 2 days), relative to Visit 3.

Subjects will return to the study site for a safety follow up visit 10 weeks after the final treatment visit.

Source: Sponsor's clinical study report(CSR page 26)

NI margins

NI margins to test non-inferiority of EQWS to active control were determined using historical evidence with the constancy assumption and pre-specified through IND 107,815. There was no issue with NI margins in this review, but brief description of historical evidence is as follows. The placebo-adjusted effect of BYETTA 10 µg in combination with metformin in HbA1c at Week 30 was -0.9 % [95% CI: -1.1, -0.7] using the label of BYETTA. Thus, the NI margin for BCB118 was 0.35% (~0.4%) that was 50% preserved drug effect of the upper 95% confidence interval. The placebo-adjusted effect of Sitagliptin was -0.69 % [95% CI: -0.84, -0.55] using meta-analysis of four placebo-controlled studies of Sitagliptin. Thus, the NI margin for BCB 120 was 0.3% that was 50% of the upper 95% CI limit.

Sample size

The applicant's sample size calculation was powered at > 95% power to demonstrate that EQWS is non-inferior to BID within 0.4% NI margin assuming a common standard deviation of 1.15% and 91% power to show the superiority of EQWS to BID in BCB118. Dropout rates were 23% (EQWS) and 26% (BID). Hence, moderate power of the study was sustained. For BCB120, applicant's sample size was powered at > 95% power to show superiority of Sitagliptin (n=122) to placebo (n=61) assuming a difference of -0.7% with a common standard deviation of 1.15% along with power at > 95% to show the EQWS(n=181) is non-inferior to Sitagliptin with the population mean difference of 0.7% between Sitagliptin and placebo and a 0.3% NI margin. Again, there were considerable dropout rates; 22% (EQWS), 20% (Sitagliptin), and 38% (Placebo).

For the trial with NI design, the quality of the trial is particularly important because poor quality could introduce bias towards the alternative hypothesis of non-inferiority[2]. Deficiency such as use of concomitant treatments whose effects may overlap with the drugs under study, or poor follow-up which may reduce the difference observed in the study, can potentially lead to a false conclusion of non-inferiority. These concerns regarding the NI trial design were considered in statistical methodologies of this review including analysis population, estimands and multiplicity detailed in following section 3.2.2.

3.2.2 Statistical Methodologies

3.2.2.1 Analysis population

The applicant used the modified ITT (mITT) population as the primary analysis population for efficacy. The mITT population consists of all randomized subjects received at least one dose of study drug excluding those who started glycemic rescue therapy or post-treatment follow-up. The applicant used evaluable population consisting of all mITT subjects who completed study procedures through Week 24 or beyond in compliance with the protocol and had adequate study drug exposure for sensitivity analysis.

The applicant's mITT population could not represent the target population in practice because of exclusion of subjects who initiated rescue therapy or with post-treatment follow-up. This mITT population can be used for de-jure estimand. In regulatory setting, de-facto estimand representing more realistic situation in clinical practice should be used. The purpose of submitted clinical studies was to establish the evidence that this new regimen EQWS is beneficial over a long period in the intent-to-treat (ITT) population as actually taken the drug in a realistic situation. Patients' dropout for different reasons in the treatment arm, comparator arm and/or placebo arm could depend on the drug effect. Furthermore, in NI trial design, follow-up data (i.e., retrieved data) is crucial to draw a true conclusion.

Due to the aforementioned reasons, in this review, ITT population, which includes all randomized and treated subjects with a baseline measurement regardless of initiation of rescue therapy and/or discontinuation of drug treatments, was used for primary statistical analysis. However, the applicant did not intend to continue measuring HbA1c after treatment discontinuation or initiation rescue therapy. Therefore, there were still remained considerable missing rates after including retrieved dropouts. Hereafter, this reviewer uses de-facto population term for representing ITT population that included retrieved dropouts in addition to mITT population for the analysis.

3.2.2.2 Estimands

[2] Choice of control group and related issues in clinical trials (ICH E10 Guidance for Industry) https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E10/Step4/E10_Guideline.pdf

The statistical parameters are referred to as estimands. The specification of estimands follows recommendations from the 2010 National Academy of Science report[3]. The applicant did not define an estimand per se; however, the applicant defined for the primary analysis based on mITT population is de-jure estimand. To evaluate the most likely true efficacy of EQWS in a regulatory setting, de-facto estimand is appropriate for estimating potential outcomes that would have been observed under different exposures regardless of adherence to treatment, use of rescue therapy or discontinuation of study drug in the T2DM patients group. An estimand in this review includes following four attributes:

[1] *Population:* ITT population including subjects who were randomized and treated with a baseline measurement regardless of initiation of rescue therapy and/or discontinuation of drug treatments;
[2] *Variable (endpoint)*: Changes in HbA1c from baseline to Week 28;
[3] *Intercurrent events*: regardless of whether or not switching to rescue medication had occurred; non-retrieved dropouts are considered relevant efficacy information (bad outcomes);
[4] *Population-level summary parameter*: difference in variable means between treatment arms for evaluating of non-inferiority followed by superiority.

3.2.2.3 Statistical hypotheses

Because trials in this submission used NI trial designs, statistical hypotheses testing procedures were important to conclude the non-inferiority as well as superiority conclusion. To protect the study-wise error rates for the primary endpoints and secondary endpoints in each study, a hierarchical testing strategy was followed, i.e., only if significance was achieved for a higher-order hypothesis in the hierarchical tests at the significance level α=0.05 two-sided (i.e., 0.025 one-sided), then hypothesis next in order would be tested. The applicant pre-specified hypotheses testing procedure in each study are as follows:

- **BCB118**

Primary endpoints

H_{01}: Non-inferiority of EQWS to BID on the primary endpoint at α=0.05 (two-sided)
H_{02}: Superiority of EQWS to BID on the primary endpoint at α=0.05 (two-sided)

Secondary endpoints

Family 1: H_{11} and H_{12} tested using the Hochberg procedure to control the Family Wise Error Rates at α=0.05 (two-sided)

H_{11}: Superiority of EQWS to BID on achieving HbA1c target value of < 7.0% at Week 28
H_{12}: Superiority of EQWS to BID on change in fasting plasma glucose concentrations from baseline to Week 28

[3] The prevention and Treatment of Missing Data in Clinical Trials https://www.nap.edu/catalog/12955/the-prevention-and-treatment-of-missing-data-in-clinical-trials

Family 2: H_{21} is tested at α=0.05 (two-sided)

H_{21}: Superiority of EQWS to BID on change in body weight from baseline to Week 28

Family 3: H_{31} is tested at α=0.05 (two-sided)

H_{31}: Superiority of EQWS to BID on change in 2-hour postprandial plasma glucose concentrations from baseline to Week 16 for subjects in the meal test cohort

- **BCB120**

Primary endpoints

H_{01}: Superiority of EQWS to Placebo on the primary endpoint at α=0.05 (two-sided)
H_{02}: Non-inferiority of EQWS to Sitagliptin on the primary endpoint at α=0.05 (two-sided)
H_{03}: Superiority of EQWS to Sitagliptin on the primary endpoint at α=0.05 (two-sided)

Secondary endpoints

H_{11}: Superiority of HbA1c goal for EQWS vs Sitagliptin at α=0.05 (two-sided)
H_{21}: Superiority of FPG for EQWS vs Sitagliptin at α=0.05 (two-sided)
H_{31}: Superiority of body weight for EQWS vs Sitagliptin at α=0.05 (two-sided)
H_{41}: Superiority of 2h PPG for EQWS vs Sitagliptin at α=0.05 (two-sided)

3.2.2.4 Statistical analysis

Statistical analysis methods by the applicant and the statistical reviewer are summarized in Table 3 and covariates in the statistical model for primary efficacy are summarized in Table 4.

Table 3. Statistical Methods for Primary Efficacy (Changes in HbA1c at Week 28)

By	Estimands (population)	Statistical Analysis Methods	BCB118	BCB120
Applicant	De-jure	MMRM	Imputation	Imputation
	De-facto	ANCOVA	No Imputation	No Imputation
		MMRM	Imputation	Imputation
		Multiple Imputation (PMM)	Imputation: *Return to baseline*	Imputation: *Washout model*
Statistical Reviewer	De-facto	ANCOVA	No Imputation	No Imputation
		Multiple Imputation (PMM): *Retrieved drop out analysis*	Imputation: *Return to baseline*	Imputation: *Washout model*
		Trimmed Means	Imputation on dropouts as bad scores	Imputation on dropouts as bad scores

Table 4. Covariates in Statistical Model for the Primary Efficacy (the changes in HbA1c)

	MMRM	ANCOVA, PMM, Trimmed Means
BCB118	**Fixed factors**: treatment, week of visit, treatment by week interaction, baseline HbA1c stratum (<9% or ≥9%), diabetes management method at screening (diet/exercise alone, SU use, or non-SU use), Renal function (normal, mild, or moderate renal impairment) **Covariates:** baseline HbA1c, baseline HbA1c stratum (<9% or ≥9%) by week, and baseline HbA1c by week interaction **Random Effect**: subject	**Covariates**: treatment, baseline HbA1c, baseline HbA1c stratum (<9% or ≥9%), diabetes management method at screening (diet/exercise alone, SU use, or non-SU use), Renal function (normal, mild, or moderate renal impairment)
BCB120	**Fixed factors**: treatment, week of visit, treatment by week interaction, baseline HbA1c stratum (<9% or ≥9%), and baseline HbA1c stratum by week interaction **Covariates:** Baseline HbA1c and Baseline HbA1c by week interaction **Random Effect**: subject	**Covariates:** treatment, baseline HbA1c stratum (<9% or ≥9%) and baseline HbA1c

The applicant's primary and sensitivity analyses

The applicant assessed the primary efficacy endpoint (change in HbA1c from baseline to Week 28) based on mITT population using mixed-effect model with repeated measures (MMRM) as a primary analysis. As aforementioned mITT population is not applicable for evaluating true drug efficacy in practice. In addition, MMRM assumes data are missing at random (MAR). The conclusions from the MMRM analysis may be subject to bias due to violation of the MAR assumption. The applicant conducted a supportive analysis by fitting a general linear model using LOCF data to handle missing data. However, either MMRM or LOCF approach using mITT population rarely estimates a valid causal effect in the presence of missing data.

Upon request by the agency, the applicant performed post-hoc sensitivity analyses to examine the impact of violation of the MAR assumption and in a different estimand (de-facto). The applicant performed ANCOVA including post-rescue, post-treatment discontinuation subjects and assumed missing not-at-random (MNAR) pattern for missing HbA1c at week 28. In addition, the pattern mixture model (PMM) based on "return to baseline" (Appendix A1) multiple imputation for missing data was applied in the active controlled study BCB 118; while the "wash out of treatment effect" (Appendix A2) multiple imputation for missing data was applied in the placebo controlled study BCB120. Details of methods are in the Appendices. For PMM, the applicant used all observations (of completers and retrieved dropouts) at Week 28 as references to impute missing values and performed 10,000 multiple imputations.

The statistical reviewer's primary and sensitivity analyses

This reviewer reproduced the applicant's primary analysis for primary endpoint (HbA1c) using MMRM for both studies. Additional efficacy analyses by the statistical reviewer were based on

the ITT population, which includes all randomized and treated subjects regardless of discontinuation of treatment or initiation of rescue therapy. As requested to the applicant, the statistical reviewer also performed sensitivity analyses to evaluate the impact of departure from MAR assumption by conducting PMM based on MNAR assumption in both studies. PMMs were based on "return-to-baseline" model for BCB118 and on "wash out of treatment effect" for BCB120. In contrast to the applicant's implementation of these models, this reviewer used retrieved dropouts only (i.e., available observations after treatment discontinuation or rescue therapy initiation) instead of all observed data (of completers and retrieved dropouts) as references to impute non-retrieved dropouts. This implementation was based on assumption of missing data of non-retrieved dropouts behave similar to those who discontinued treatment or initiated rescue therapy but had the HbA1c measured at Week 28. In BCB120, non-retrieved data in placebo group were imputed based on monotone regression pattern and implemented as sequential imputations using intermittent observed measurement in placebo group between baseline and Week 28. Multiple imputations were performed 100 times in each model.

- ***Trimmed Means Analysis***

Both studies had high dropout rates (ranged 20% to 38%) that are impactful for evaluating the efficacy of EQWS. Furthermore, both studies were designed with unequal number of subjects across treatment groups and smallest sample size in placebo group. Retrieved dropout rates including patients discontinued treatment or initiated rescue therapy adjunct to the exercise were 4% (EQWS) and 8% (BID) in BCB118 and 7% (EQWS), 8% (Sitagliptin), and 17% (Placebo) in BCB120. Imbalanced missing rates could induce low power of placebo group with retrieved dropouts including subjects who initiated rescue therapy. To some extent, effective rescue medication could mask or exaggerate the efficacy effects of the initially assigned treatments, and result in biased estimates of study drug effects.

To address the potential bias from the imbalanced sample size and missing rates in both studies, this reviewer also performed Trimmed means analysis[4] (Appendix A3, Figure 13) using ITT population by treating the missing data as bad outcomes and trimming out them with other observed bad outcomes to examine the robustness of primary efficacy results in lieu of imputations of actual values based on retrieved dropouts.

In T2DM clinical setting, the commonly used primary endpoint is changes of HbA1c from baseline. Patients may discontinue the study treatment due to lack of efficacy, adverse event or toxicity, which will result in missing values in primary endpoint. The treatment discontinuation reasons are not fatal under most circumstances. We can assume those drop-out patients have bad although unobserved outcomes (HbA1c), and there is no need to differentiate those bad outcomes caused by different reasons. Therefore, valid tests can be constructed based on the differences between good outcomes after treating all dropouts as equally bad outcomes. Estimating the average of the good outcomes would suffice in this clinical setting for evaluating the efficacy of study drug.

[4] Permutt and Li, Trimmed means for symptom trials with dropouts (*2017, Pharmaceut. Statist. 16; 20-28*)

Trimmed means analysis method estimates the difference between the means of the included upper (e.g., better half) proportion of the outcome distribution for each treatment arm. The estimate is a de-facto estimand. All missing values, as well as other observed bad outcomes worse than a certain proportioned outcome (e.g., median), are taken to be equally bad. Also, the fact of having more completers in a treatment arm is considered as a benefit of that treatment, which is taken into account by including completers with better outcomes from that treatment arm in Trimmed Means analysis method. Thus, abovementioned concerns about the imbalanced proportions of sample size and missing rates could be addressed by performing Trimmed means analysis. By this approach, the question of how well the drug works when it seems to work, even if there are complete measurements on patients in whom it did not work well could be addressed.

This reviewer conducted trimmed means analysis using 50% (better half) and 80% (adaptive proportions based on missing rates (20%) to maximize the utility of available observed data) better data for changes in HbA1c from baseline to Week 28 in both studies to estimate the population trimmed means. The 95% confidence interval of the difference in trimmed means for treatment groups and p-value for testing the null hypothesis of no difference in the trimmed means were calculated by 5,000 permutation tests using random shuffling the treatment assignments within the datasets. Empirical Cumulative Distribution Function (ECDF) plot was used to display the difference of trimmed means for treatment groups using the shaded area between the CDF curves and below the horizontal line at the height of fraction used (e.g., 50% or 80%). In this review, 80% fraction of the data was selected to evaluate the efficacy of EQWS.

3.2.3 Patient Disposition, Demographic, and Baseline Characteristics

Patient disposition and analysis populations with percent dropouts, reasons for dropouts and protocol violations for each study are tabulated in Tables 5 and 6. As mentioned before submitted clinical trials were not designed for following up patients who discontinued treatment; however, there were some retrieved dropouts.

Table 5. Analysis Population and Patient Disposition for BCB118

	BID	***EQWS***	***BID/EQWS***	***EQWS/EQWS***
	≤ 28 weeks		> 28 weeks	
Randomized	148	229		
Randomized & Treated (ITT population)	**146 (100%)**	**229 (100%)**		
Completed 28 week	116 (79.5%)	193 (84.3%)		
Observed N (%) at Week 28 for HbA1c				
mITT population (de-jure)	107 (73.3%)	176 (76.9%)		
Percent Dropouts	26.7 %	23.1 %		
mITT + retrieved dropouts (de-facto)	119 (81.5%)	182 (79.5%)		
Percent Dropouts	18.5 %	20.5 %		
Discontinuation Reasons				

Miscoded/Enrolled twice in different arms/No baseline measurement	2	4	0	0
ADMINISTRATIVE	2	0	0	0
ADVERSE EVENT	8	6	2	1
INVESTIGATOR DECISION	1	1	3	0
LOSS OF GLUCOSE CONTROL	0	1	0	0
LOST TO FOLLOW-UP	7	5	4	7
PROTOCOL VIOLATION	1	2	0	0
WITHDRAWAL BY SUBJECT	11	17	7	16

Source: Statistical Reviewer's analysis

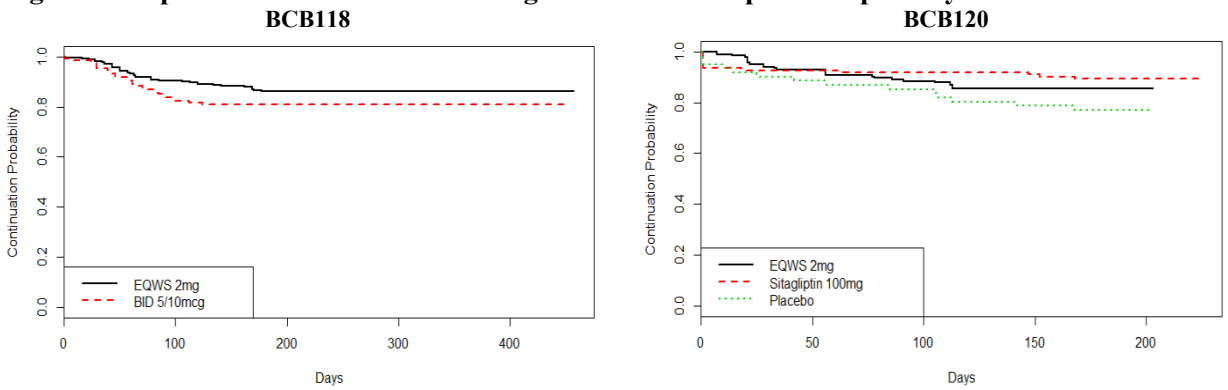
Table 6. Analysis Population and Patient Disposition for BCB120

	EQWS 2 mg	***Sitagliptin***	***Placebo***
Randomized	182	122	61
Randomized & Treated (ITT population)	**181 (100%)**	**122 (100%)**	**61 (100%)**
Completed 28-week	155 (85.6%)	109 (89.3%)	47 (77.0%)
Observed N (%) at Week 28 for HbA1c			
mITT population (de-jure)	141 (77.9%)	98 (80.3%)	38 (62.3%)
Percent Dropouts	22.1%	19.7%	37.7%
mITT + retrieved dropouts (de-facto)	153 (84.5%)	106 (86.8%)	48 (78.7%)
Percent Dropouts	15.5%	13.2%	21.3%
Discontinuation Reasons (N)			
ADMINISTRATIVE	1	0	0
ADVERSE EVENT	4	0	3
INVESTIGATOR DECISION	1	0	1
LOST TO FOLLOW-UP	7	6	3
PROTOCOL VIOLATION	1	0	0
WITHDRAWAL BY SUBJECT	14	7	7

Source: Statistical Reviewer's analysis

There were 377 and 365 subjects randomized, and 375 and 364 subjects treated in BCB118 and BCB120 respectively. The applicant used mITT population (283 subjects in BCB118 and 277 subjects in BCB120) excluding subjects discontinued treatment or initiated rescue therapy for efficacy analysis. The statistical reviewer named de-facto population in this review as observed subjects including subjects discontinued treatment or initiated rescue therapy for efficacy analysis under de-facto estimand using ITT population (375 subjects in BCB118 and 364 subjects in BCB120). Based on exposure days to treatment, Kaplan-Meier graph is plotted in Figure 3 for comparing proportion of subjects continued in the study between treatment groups.

Figure 3. Kaplan-Meier Curves Presenting the Time to Dropout Graphically



Source: Statistical Reviewer's analysis

Demographics (Gender, Age, Race and Ethnicity) at baseline are tabulated in Table 7. Note that both studies were conducted within USA only. No notable differences were found between treatment arms for demographics in each study. In BCB120 study, there were more Hispanic or Latino ethnicity compared to Non-Hispanic or Latino that was different from BCB118 study of more Non-Hispanic or Latino.

Baseline characteristics (HbA1c, BMI, Renal Function, Duration of Diabetes, and Diabetes management method at screening) by treatment arms in each study are summarized in Table 8. There was no notable difference between treatment arms in each study as well as across studies.

Table 7. Demographics at baseline

		BCB118		BCB120		
		EQWS (N=229)	***BID (N=146)***	***EQWS (N=181)***	***Sitagliptin (N=122)***	***Placebo (N=61)***
Gender	Male	148 (64.6)	92 (63)	89 (49.2)	66 (54.1)	37 (60.7)
n(%)	Female	81 (35.4)	54 (37)	92 (50.8)	56 (45.9)	24 (39.3)
Age	Mean (SD)	55.6 (9.98)	56.5 (9.04)	53.4(9.82)	54.3 (9.01)	53.4 (9.48)
	Age < 65, n(%)	182 (79.5)	118 (80.8)	154 (85.1)	106 (86.9)	54 (88.5)
	Age ≥ 65, n(%)	47 (20.5)	28 (19.2)	27 (14.9)	16 (13.1)	7 (11.5)
Race	White	168 (73.4)	110 (75.3)	148 (81.8)	98 (80.3)	50 (82.0)
n(%)	Black/African American	38 (16.6)	23 (15.8)	24 (13.3)	18 (14.8)	7 (11.5)
	Asian	17 (7.4)	8 (5.5)	9 (5.0)	2 (1.6)	3(4.9)
	Other	6 (2.6)	5 (3.5)	0	4 (3.2)	1 (1.6)
Ethnicity	Hispanic or Latino	54 (23.6)	34 (23.3)	111 (61.3)	77 (63.1)	32 (52.5)
n(%)	Not Hispanic or Latino	174 (76)	112 (76.7)	70 (38.7)	45 (36.9)	29 (47.5)

Source: Sponsor's clinical study reports Table S. 3.1.4.1.2 (eCTD module 5.3.5.1)

Table 8. Baseline Characteristics

		BCB118		BCB120		
	Category	***EQWS (N=229)***	***BID (N=146)***	***EQWS (N=181)***	***Sitagliptin (N=122)***	***Placebo (N=61)***
HbA1c (%)	Mean(SD)	8.47 (1.05)	8.51 (1.0)	8.42 (0.99)	8.5 (1.04)	8.5 (1.04)
N(%)	HbA1c < 9%	159 (69.4)	97 (66.4)	125 (69.1)	83 (68)	42 (68.9)
	HbA1c ≥ 9%	68 (29.7)	49 (33.6)	56 (30.9)	39 (32)	19 (31.1)
BMI(kg/m²)	Mean(SD)	33.08 (5.9)	33.4 (5.2)	32.08 (5.4)	31.62 (5.8)	31.5 (5.14)
N(%)	BMI < 30	68 (29)	44 (30.1)	74 (40.9)	55 (45.1)	28 (45.9)
	BMI ≥ 30	160 (69.9)	102 (69.9)	107 (59.1)	67 (54.9)	33 (54.1)
Duration of	N	219	136	178	121	61
Diabetes	Mean(SD)	8.5 (5.9)	8.5 (6.2)	8.5 (6.3)	7.9 (4.6)	8.7 (5.7)
Renal	Normal	85 (37.1)	55 (37.7)	102 (56.4)	64 (52.5)	35 (57.4)
Function	Mild Impairment	113 (49.3)	76 (52.1)	78 (43.1)	54 (44.3)	26 (42.6)
	Moderate Impairment	29 (12.7)	15 (10.3)	1 (0.6)	4 (3.3)	0
	Severe Impairment	0	0	0	0	0
Diabetes	SU	8 (3.5)	6 (4.1)	NA	NA	NA
manageme	SU + Metformin	76 (33.2)	52 (35.6)	NA	NA	NA
nt Method	SU + TZD	1 (0.4)	0	NA	NA	NA
at	SU + Metformin + TZD	4 (1.7)	2 (1.4)	NA	NA	NA
Screening	Diet and Exercise	31 (13.5)	17 (11.6)	NA	NA	NA
N(%)	Metformin	102 (44.5)	65 (44.5)	NA	NA	NA
	TZD	2 (0.9)	0	NA	NA	NA
	Metformin + TZD	5 (2.2)	4 (2.7)	NA	NA	NA

Source: Sponsor's clinical study reports Table S. 3.1.4.1.2 (eCTD module 5.3.5.1)

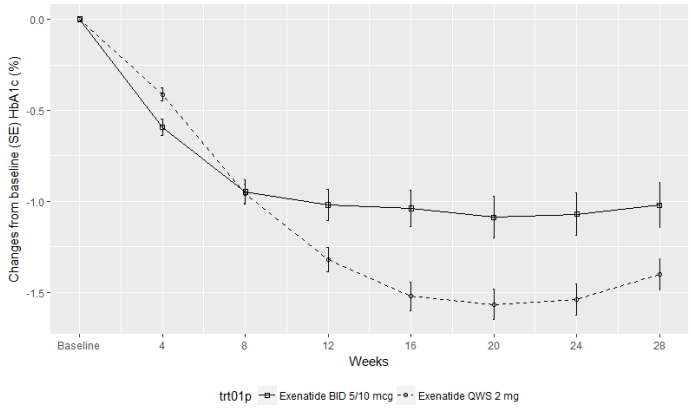
3.2.4 Results and Conclusions

3.2.4.1 Primary Endpoint: HbA1c (%)

Study BCB118

The longitudinal changes of HbA1c (%) from baseline until Week 28 in study BCB118 are illustrated in Figure 4. The LS means (SE) of changes of HbA1c(%) from baseline at Week 28 using de-facto population were -1.39 (0.09) and -1.03 (0.12) for EQWS 2mg and BID 10 µg respectively.

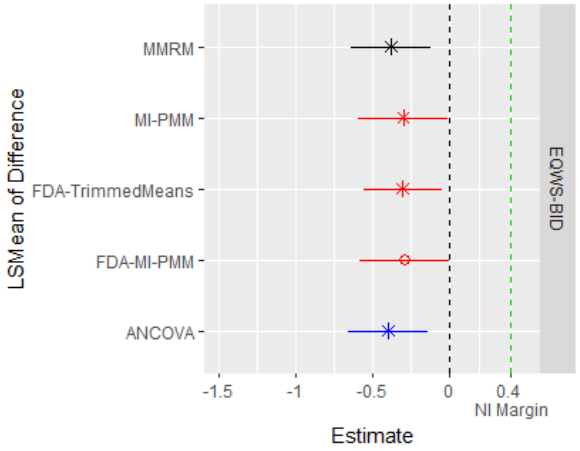
Figure 4. Longitudinal Changes of HbA1c (%) in BCB118



Source: Statistical Reviewer's analysis

Figure 5 is a forest plot to summarize the differnece (LS means and 95% CI) in HbA1c changes from baseline to Week 28 between EQWS and BID based on primary and post-hoc sesnsitivity analyses by the applicant and the statistical reviewer. Analyses included primary analysis (MMRM) using mITT population and sensitivity analysis with multiple imputations using PMM (MI-PMM) in ITT population by the applicant; ANCOVA using de-facto population (ANCOVA), multiple imputations using PMM (FDA-MI-PMM), and Trimmed Means analysis (FDA-TrimmedMeans) in ITT poluation by statistical reviewer.

Figure 5. Summary of Difference in HbA1c Change – Study BCB118



Method	Estimate (95% CI)	p-value
MMRM	-0.38(-0.64,-0.12)	0.0046
MI-PMM	-0.3(-0.59,-0.01)	0.0451
TrimmedMeans	-0.3(-0.56,-0.05)	0.03
FDA-MI-PMM	-0.29(-0.58,0)	0.0532
ANCOVA	-0.4(-0.66,-0.14)	0.0032

Source: Statistical Reviewer's analysis

Green dotted line indictaes NI margin (0.4) and black dotted line indicates zero for testing superiority. Symbol * or °, used to mark point estimate for difference in HbA1c change between EQWS and BID, indicates the significance of the superioirty testing (* if p-value is less than 0.05; °if p-value is greater or equal to 0.05) . The colored line of 95% CI is indicating the population used in the analysis (Black: mITT with imputations, Blue: de-facto, Red: ITT with imputations).

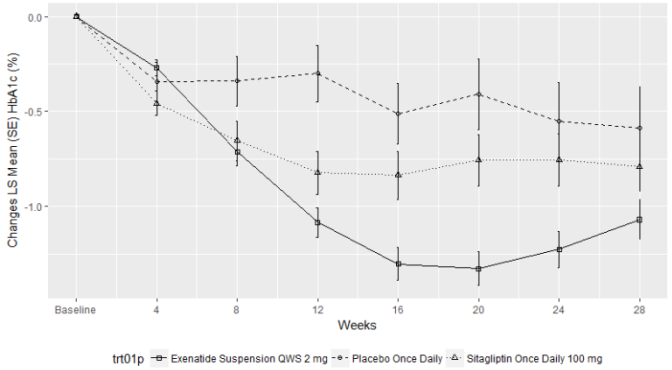
All analyses showed the non-inferioirty of EQWS to BID as the upper limit of 95% confidence interval for the difference between EQWS and BID was below pre-specified NI margin 0.4. The superioirity of EQWS to BID was also demonstrated in all analyses with significant p-values.

Study BCB120

The longitudinal changes of HbA1c (%) from baseline until Week 28 in study BCB120 are illustrated in Figure 6. The LS means (SE) of changes of HbA1c(%) from baseline at Week 28 using de-facto population were -1.07 (0.10), -0.79 (0.13) and -0.58 (0.21) for EQWS 2mg, Sitagliptin 100mg and Placebo respectively.

Note that differences of LS Means between Sitagliptin and placebo was -0.19% (95% CI: -0.62, 0.24) that was smaller than -0.7% of placebo-adjusted Sitagliptin effect from meta-analysis used for NI margin determination and power calculation described in previous section 3.2.1. Apparently Sitagliptin (active control) did not show the superiority to placebo in this study BCB120, which makes the non-inferiority testing of EQWS vs. Sitagliptin not meaningful anymore.

Figure 6. Longitudinal Changes of HbA1c (%) in BCB120

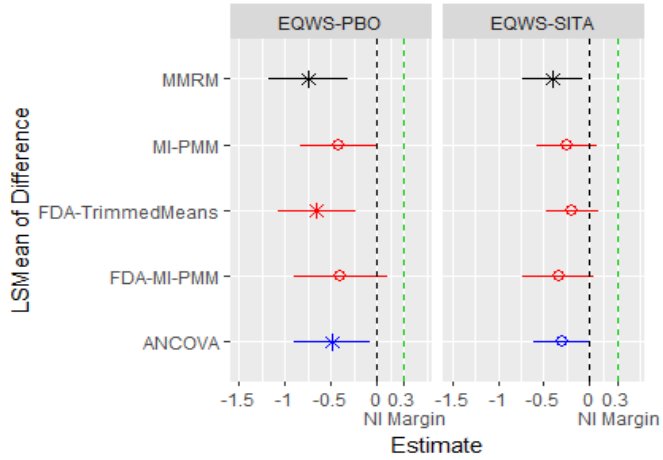


Source: Statistical Reviewer's analysis

Forest plots in Figure 7 (Left panel: between EQWS and Placebo, Right panel: between EQWS and Sitagliptin) summarize the differnece (LS means and 95% CI) in HbA1c changes from baseline to Week 28 between EQWS and BID based on primary and post-hoc sesnsitivity analyses by the applicant and the statistical reviewer. Analyses included primary analysis (MMRM) using mITT population and sensitivity analysis with multiple imputations using PMM (MI-PMM) in ITT population by the applicant; ANCOVA using de-facto population (ANCOVA), multiple imputations using PMM (FDA-MI-PMM), and Trimmed Means analysis (FDA-TrimmedMeans) in ITT poluation by statistical reviewer.

ANCOVA analysis results using de-facto population showed superioirty of EQWS to Placebo. Trimmed means analysis results are consistent with the ANCOVA analysis results. However, the superiority of EQWS to Sitagliptin was not demonstrated in any analysis using ITT population although the non-inferioirity of EQWS to Sitagliptin with the NI margin 0.3 was established.

Figure 7. Summary of Primary Efficacy Analysis Results for BCB120



	Method	Estimate (95% CI)	p-value
EQWS-PBO	MMRM	-0.75(-1.18,-0.32)	0.0008
EQWS-PBO	MI-PMM	-0.41(-0.84,0 02)	0.0563
EQWS-PBO	TrimmedMeans	-0.655(-1.08,-0.23)	0.0116
EQWS-PBO	FDA-MI-PMM	-0.39(-0.91,0.12)	0.1363
EQWS-PBO	ANCOVA	-0.49(-0.91,-0.07)	0.0213
EQWS-SITA	MMRM	-0.4(-0.73,-0.07)	0.0169
EQWS-SITA	MI-PMM	-0.25(-0.58,0 08)	0.141
EQWS-SITA	TrimmedMeans	-0.19(-0.48,0 09)	0.0547
EQWS-SITA	FDA-MI-PMM	-0.34(-0.74,0 05)	0.0889
EQWS-SITA	ANCOVA	-0.3(-0.62,0 02)	0.065

Source: Statistical Reviewer's analysis

Green dotted line indictaes NI margin (0.3) and black dotted line indicates zero for testing superiority. Symbol * or °, used to mark point estimate for difference in HbA1c change between EQWS and BID, indicates the significance of the superioirty testing (* if p-value is less than 0.05; °if p-value is greater or equal to 0.05) . The colored line of 95% CI is indicating the population used in the analysis (Black: mITT with imputations, Blue: de-facto, Red: ITT with imputations).

3.2.4.2 Secondary Endpoints

In BCB118, because hypothesis tests for primary endpoints were all statistically significant, hypothesis testing for secondary endpoints would be conducted in the following order to control the FWER at 0.05: 1) proportion of patients achieving HbA1c < 7.0% at Week 28 (patients with missing values at Week 28 counted as non-responders), and 2) change in fasting plasma glucose concentrations from baseline to Week 28. The proportions of subjects achieved HbA1c < 7.0% at Week 28 were 40% in EQWS group compared to 38% in BID group in ITT population. The difference in the proportions did not reach statistical significance. Any further formal hypothesis testing was stopped. For descriptive purpose, the mean changes from baseline to Week 28 for fasting serum glucose were -35 mg/dL and -27mg/dL for EQWS and BID respectively.

In BCB120, no further hypothesis testing for secondary endpoints was performed because hypothesis testing for primary endpoints failed to show the superiority of EQWS to Sitagliptin. For descriptive purpose, there were 41%, 31%, and 26% of subjects in ITT population achieved HbA1c < 7.0% at Week 28 in EQWS, Sitagliptin and Placebo group respectively. The mean changes from baseline to Week 28 for fasting serum glucose were -24 mg/dL, -19 mg/dL and 3.9 mg/dL for EQWS, Sitagliptin and Placebo respectively.

3.2.4.3 Conclusions

In BCB118, EQWS was both non-inferior and superior to BID. In BCB120 EQWS was superior to placebo and non-inferior to Sitagliptin. The totality of evidence indicated that EQWS has a treatment effect for glycemic control in T2DM patients, although, the non-inferiority of EQWS to Sitagliptin was not meaningful due to the fact that Sitagliptin was not superior to placebo in study BCB120.

3.3 Evaluation of Safety

An evaluation of safety of the phase 3 clinical trials is included in the FDA clinical review by Dr. Mahtab Niyatti of the Division of Metabolism and Endocrinology Products.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

Both trials BCB118 and BCB120 were conducted within USA only and there was no other geographic region involved. No subgroup analysis by region was conducted. Subgroup analyses (i.e., ANCOVA model using de-facto population and same covariates as primary analysis) for primary endpoints (change in HbA1c (%) from baseline to Week 28) were performed across subgroups defined by race (White vs. Black or African American vs. Others), gender (Female vs. Male), and age ($\geq$ 65 vs. < 65) in section 4.1. In addition, subgroup analyses by ethnicity (Hispanic/Latino vs. Non-Hispanic/Latino) and renal function at baseline (Normal vs. Impairment (normal, mild, and moderate)) were explored in section 4.2. Also, correlation between baseline HbA1c and the changes of HbA1c from baseline to Week 28 was demonstrated in graphics.

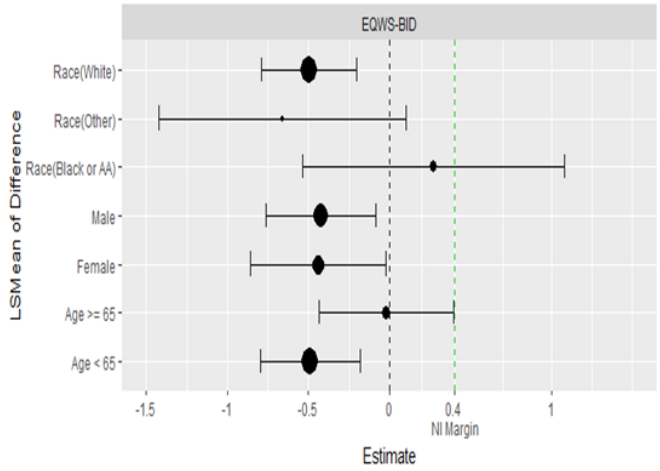
4.1 Gender, Race, and Age

Figures 8 and 9 show forest plots depicting LS mean differences with 95% confidence interval for HbA1c changes from baseline to Week 28 between EQWS and active control arm or placebo arm in subgroups of race, gender, and age in Studies BCB118 and BCB120 respectively. The size of circle indicates the point estimate is proportional to the number of subjects (N) in each subgroup for the analysis.

Majority of subgroups (White, both sexes and age <65) showed the favorable results of EQWS compared to control, which are consitent with the results from the primary analysis for HbA1c changes from baseline to Week 28 in BCB118. Subgroups with small number of subjects showed the wider 95% CI. In BCB120, most subgroups concurred the favorable direction to EQWS compared to placebo like the primary analysis results, except three subgroups (Black/AA, Other race groups and Age $\geq$65).

No interaction terms between subgroup and treatment group were significant using ANCOVA model in both studies.

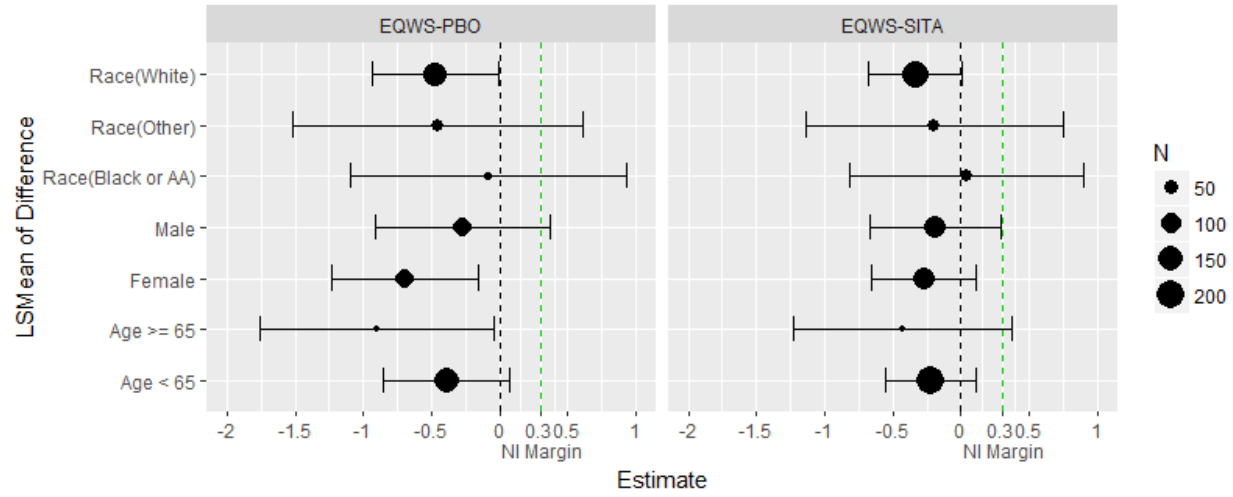
Figure 8. Subgroup analyses for Race, Gender and Age – Study BCB118



Subgroup	n	LS mean (95% CI)
Race(White)	228	-0.5(-0.79,-0.2)
Race(Other)	29	-0.66(-1.42,0.1)
Race(Black or AA)	44	0.27(-0.54,1.08)
Male	186	-0.42(-0.76,-0.08)
Female	115	-0.44(-0.86,-0.02)
Age >= 65	60	-0.02(-0.43,0.39)
Age < 65	241	-0.49(-0.8,-0.18)

Source: Statistical Reviewer's analysis

Figure 9. Subgroup analyses for Race, Gender, and Age – Study BCB120



EQWS-PBO		
Subgroup	**n**	**LS mean (95% CI)**
Race(White)	144	-0.47(-0.94,-0.01)
Race(Other)	34	-0.46(-1.52,0.61)
Race(Black or AA)	23	-0.09(-1.1,0.93)
Male	90	-0.27(-0.91,0.37)
Female	89	-0.69(-1.23,-0.16)
Age >= 65	21	-0.9(-1.76,-0.05)
Age < 65	158	-0.39(-0.85,0.07)

EQWS-SITA		
Subgroup	**n**	**LS mean (95% CI)**
Race(White)	190	-0.34(-0.68,0.01)
Race(Other)	35	-0.2(-1.14,0.74)
Race(Black or AA)	34	0.04(-0.82,0.9)
Male	122	-0.19(-0.67,0.29)
Female	116	-0.27(-0.66,0.11)
Age >= 65	21	-0.43(-1.23,0.37)
Age < 65	206	-0.22(-0.55,0.11)

Source: Statistical Reviewer's analysis

4.2 Other Special/Subgroup Populations

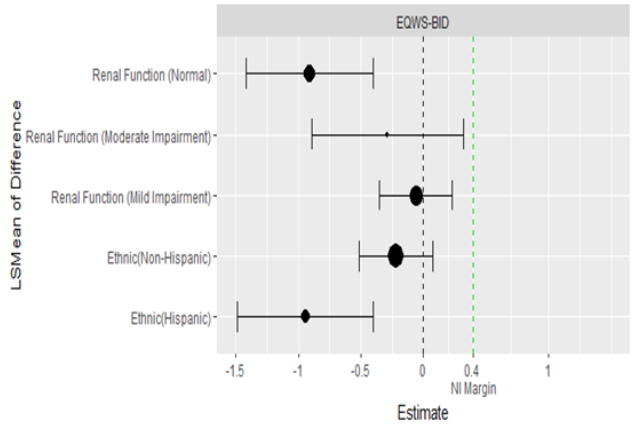
Figures 10 and 11 show forest plots depicting LS mean differences with 95% confidence interval for HbA1c changes from baseline to Week 28 between EQWS and active control arm or placebo arm in subgroups of renal function and ethnicity in Studies BCB118 and BCB120 respectively.

The size of circle indicates the point estimate is proportional to the number of subjects (N) in each subgroup for the analysis. In BCB120, due to small sample size in subgroup of mild impairment renal function, mild and moderate impairment renal function subgroups are combined.

There were no notable differences for favoribility to EQWS compared to comparators and placebo across subgroups in both studies.

Numerically significant interaction between normal renal function and treatment group and interaction between ethnicity and treatment group were detected in BCB118 (ANCOVA interaction term p-values: 0.004 and 0.028). Numerically significant interaction between mild/moderate renal function and Sitagliptin treatment group and interaction between non-Hispanic ethnicity and Sitagliptin treatment group were detected in BCB118 (ANCOVA interaction term p-values: 0.02 and 0.01).

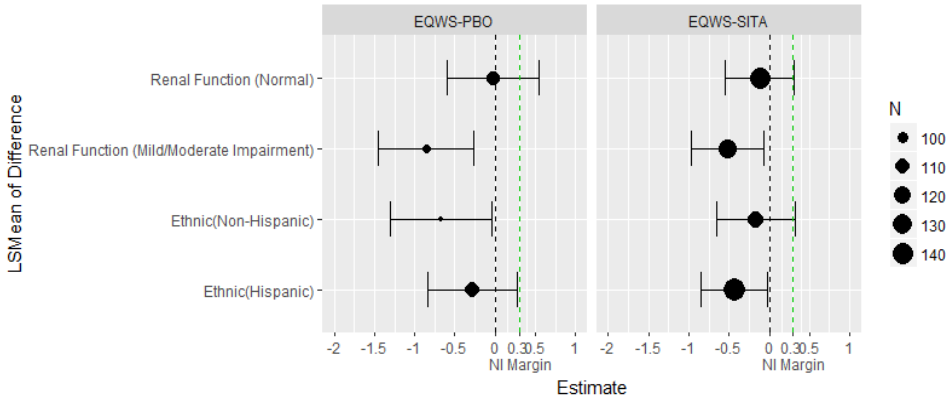
Figure 10. Subgroup Analyses for Renal Function and Ethnicity for BCB118



Subgroup	n	LS mean (95% CI)
Normal Renal Function	109	-0.91(-1.42,-0.4)
Moderate Impairment	34	-0.29(-0.89,0.32)
Mild Impairment	158	-0.06(-0.36,0.23)
Ethnic(Non-Hispanic)	228	-0.22(-0.51,0.07)
Ethnic(Hispanic)	73	-0.94(-1.48,-0.4)

Source: Statistical Reviewer's analysis

Figure 11. Subgroup Analyses for Renal Function and Ethnicity for BCB120



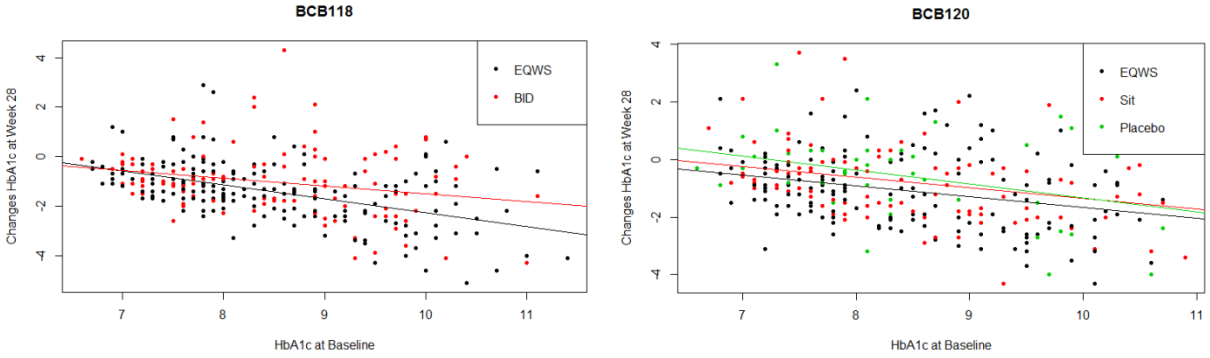
EQWS-PBO		
Subgroup	n	LS mean (95% CI)
Normal Renal Function	106	-0.03(-0.6,0.54)
Moderate/mild Impairment	95	-0.86(-1.45,-0.27)
Ethnic(Non-Hispanic)	92	-0.68(-1.31,-0.05)
Ethnic(Hispanic)	109	-0.29(-0.84,0.27)

EQWS-SITA		
Subgroup	n	LS mean (95% CI)
Normal Renal Function	135	-0.12(-0.55,0.31)
Moderate/mild Impairment	124	-0.52(-0.97,-0.07)
Ethnic(Non-Hispanic)	114	-0.17(-0.66,0.32)
Ethnic(Hispanic)	145	-0.44(-0.85,-0.02)

Source: Statistical Reviewer's analysis

Scatters plots in Figure 12 show the correlation between baseline HbA1c and the changes of HbA1c at Week 28 from baseline in both studies (Left panel:BCB118, Right panel:BCB120). In BCB118, larger separation of regression lines was seen in subjects with higher baseline HbA1c. On the contrary, in BCB120, regression lines for Sitagliptin and placebo groups are closer in higher baseline HbA1c subjects. However, regression line for EQWS showed consistent separation from regression lines for comparators (Sitagliptin and placebo) in parallel. Both scatter plots showed the subjects with higher baseline HbA1c tended to have the larger changes in HbA1c at Week 28.

Figure 12. Correlation between baseline HbA1c and the changes of HbA1c at Week 28



Source: Statistical Reviewer's analysis

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

There were several statistical issues identified in the clinical studies in this submission:

1) Estimand

- **Study Question:** To claim the efficacy of EQWS as an adjunct to diet and exercise to improve glycemic control in adults with T2DB in the label, difference in mean changes of HbA1c from baseline to Week 28 between EQWS and comparator(Sitagliptin)/placebo needed to be evaluated in a real world situation.

- **What the applicant used:** The applicant derived a de-jure estimand using mITT population that excluded the subjects who used rescue therapy and/or discontinued EQWS in the primary analysis.

- **What the statistical reviewer used:** To appropriate evaluate EQWS efficacy, de-facto estimand should be estimated using the ITT(de-facto) population, which include all randomized subjects regardless of initiation of rescue therapy or discontinuation of EQWS in the trials.

2) Non-Inferiority study design

Based on NI study guidance[5], adequate interpretation of NI study could be established with three steps. First, there should be reliable information about the effect the active control drug had in past studies. Second, there should be reason to believe the effect the active control drug has in the current study is similar to the effect observed in past studies. Third, the NI study provides reliable information about the effect of the test drug relative to the comparator.

- **NI margin:** In each NI trial in this submission, there was reliable information about the placebo-adjusted effects from the past studies of previously marketed drugs (BYETTA and Sitagliptin). Thus, we have no issues with the NI margins (0.4% and 0.3%) used in both clinical trials (BCB118 and BCB120).

- **Assay sensitivity:** Power calculation in the protocol by the applicant was based on the treatment effect of Sitagliption observed from previous studies. The placebo-adjusted effect of Sitagliptin was -0.7% based on previous studies. However, in BCB120 study, the difference between Sitagliptin and placebo was -0.19% that was less than -0.7% (Figure 6). Assay sensitivity was not sufficient in trial BCB120. Sitagliptin showed weaker than previous observed and statistically insignificant placebo-adjusted treatment effect. This lack of treatment effect for Sitagliptin makes it easier to claim non-inferiority/superiority of EQWS to Sitagliptin and this non-inferiority is not meaningful. However, placebo treatment arm was added in the NI trial design by the recommendation of the agency during the communication at the IND review stage. Thus, the efficacy of EQWS was able to be evaluated with the placebo arm in the same clinical trial by showing the superiority of EQWS compared to the placebo (Figure 7).

3) Missing data

- **High missing rates**: There were high missing rates and very few retrieved dropout because the applicant did not perform diligently follow-up for patients who discontinued treatment in both studies. Missing rates for ITT population that includes the subjects who initiated rescue therapy and/or discontinued treatment were not ignorable: 18.5% (BID)

[5] Guidance to Industry: Non-inferiority clinical trials to establish effectiveness
https://www fda.gov/downloads/Drugs/Guidances/UCM202140.pdf

and 20.5% (EQWS) in BCB118, and 21.3% (Placebo), 13.2% (Sitagliptin) and 15.5% (EQWS) (Tables 5 and 6).

- **Methods to address missing data (the applicant):** The applicant used MMRM for primary analysis based on mITT population, which may be inappropriate method to deal with missing data. MMRM assumes data were missing at random (MAR), and it would result in biased estimates if that assumption does not hold. In practice with T2DM clinical setting, strong assumption of MAR is very likely untrue. As post-hoc analysis, the applicant performed sensitivity analyses suggested by FDA using PMM in ITT population (Table 3).

- **Sensitivity analysis using Trimmed means analysis (the statistical reviewer):** The statistical reviewer performed sensitivity analyses using PMM multiple imputations as well as Trimmed means analysis. In BCB120, randomization ratio in EQWS and placebo was 3:1 and dropout rates were 15.5% and 21.3% that included retrieved dropouts (7% and 16% in EQWS and placebo respectively). Most of retrieved dropouts initiated rescue therapy (e.g. insulins) for a standard of care for T2DM. Multiple imputations using PMM based on retrieved dropouts produced conservative estimates of treatment effects for EQWS because of the strong effect of rescue therapy for the subjects who initiated rescue therapy in placebo group. Trimmed means analysis used same ITT population and took into consideration of different missing rates in unbalanced treatment groups to evaluate the efficacy of EQWS appropriately (Figure 7).

Overall, statistical findings provided the collective evidence that EQWS 2mg is superior to BYETTA, and EQWS 2mg is also superior to placebo for glycemic control in T2DM patients.

5.2 Conclusions and Recommendations

In conclusion, the efficacy results in the current NDA submission support the purported claim that EQWS improves the glycemic control in T2DM adult patients as an adjunct to diet and exercise.

This reviewer recommends the approval of EQWS (BYDUREON BCise® 2mg once weekly).

For future NI trial designs, this reviewer recommend the sponsor establish best applicable estimand in advance for proposed indication in the label. Also, thorough inspection of assay sensitivity of active control is necessary for obtaining the correct NI margin and sample size. Above all, prevention of missing data and follow-up of dropouts during the trial are always important for a successful clinical trial study.

5.3 Labeling Recommendations

Because we are concerned with using the most appropriate estimate of the treatment effect in clinical practice, the applicant's primary analysis results based on MMRM and mITT population should not be reported in the label. Rather, the statistical reviewer recommend the analysis

results based on ANCOVA using de-facto estimand (Table 1) should be included in the label to show the estimated difference of effect size between EQWS and comparator/placebo groups.

Current results in Section 14 should be modified to include statistically significant results only in the sequence of pre-specified multiple testing using de-facto estimand. The wording for BCB120 results needs to be changed from (b) (4) (b) (4) (b) (4) to "In this study, treatment with BYDUREON BCISE 2mg once weekly resulted in a statistically significant mean reduction in HbA1c compared to placebo."

This reviewer recommends the results about the mean change from mean baseline in body weight should be removed from the product label because these descriptive sentences could be misunderstood as the drug effect on reduction of body weight.

APPENDICES

A1. Return to baseline

Missing data at Week 28 of HbA1c measurement was imputed using return to baseline model in BCB118 for multiple imputations to implement the pattern mixture model. The values of missing data at Week 28 will be random selection of the normal distribution as following:

$$\sim N(Baseline\ of\ the\ missing\ subject, Vc * (1 + \frac{1}{n_c}))$$

V_c: Variance of HbA1c at Week 28 for completers* in both arms (EQWS and BID)
n_c: The number of completers* in both arms (EQWS and BID)

** the statistical reviewer used retreived dropouts (subjects who had observed values even though they were dropped out from the study regardless of initiation of rescue therapy or discontinuation of treatment) instead of completers (subjects who have observed HbA1c measurement).*

Multiple imputations were performed 100 times to generate imputed data sets for ANCOVA analysis.

A2. Washout model

Missing data at Week 28 of HbA1c measurement in EQWS arm and Sitagliptin arm was imputed using Washout model in BCB120 for multiple imputations to implement the pattern mixture model. The values of missing data at Week 28 will be random selection of the normal distribution as following:

$$\sim N(\hat{\beta}_0 + \hat{\beta}_1 * x_0, MSE * (1 + \frac{1}{n_{pc}} + \frac{(x_0 - \overline{x_{pc}})^2}{S_{xxpc}})$$

,where

x_0 : Baseline HbA1c measurement of the missing subject
n_{pc}: The number of completers* in placebo arm
x_{pc}: Mean of Baseline HbA1c measurments of completers* in placebo arm
S_{xxpc}: (n_{pc}-1) × variance of Baseline HbA1c measuremens of completers* in placebo arm
$\hat{\beta}_0, \hat{\beta}_1$, MSE obtained from fitted regression of

$$Y_{pc} = \hat{\beta}_0 + \hat{\beta}_1 * Baseline_{pc} + error$$

Y_{pc}: HbA1c measurements at Week 28 of completers* in placebo group

$Baseline_{pc}$: Baseline HbA1c measurements of completers* in placebo group
MSE:$Residuals\ of\ standard\ error^{2}$

** Statsitical reviewer used retreived dropouts (subjects who had observed values even though they were dropped out from the study regardless of initiation of rescue therapy or discontinuation of treatment) instead of completers (subjects who have observed HbA1c measurement).*

Missing data at Week 28 of HbA1c measurment in placebo arm was imputed using monotone missing pattern. Sequential imputation was performed using monotone reg option in SAS Proc MI. Multiple imputations were performed 100 times to generate imputed data sets for ANCOVA analysis.

A3. Trimmed means analysis

Trimmed means are defined as the mean of the upper (1- τ) fraction of the distribution:

$$\mu^{\tau} = \frac{1}{1-\tau} \int_{F^{-1}(\tau)}^{\infty} t dF(t)$$

,where
τ: The trimmed fraction
F(t): The cumulative distribution function Prob(Y ≤ t).

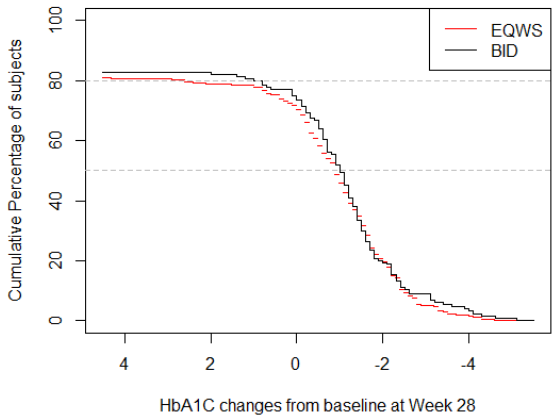
Difference in population trimmed means are the average of the difference for all quantiles:

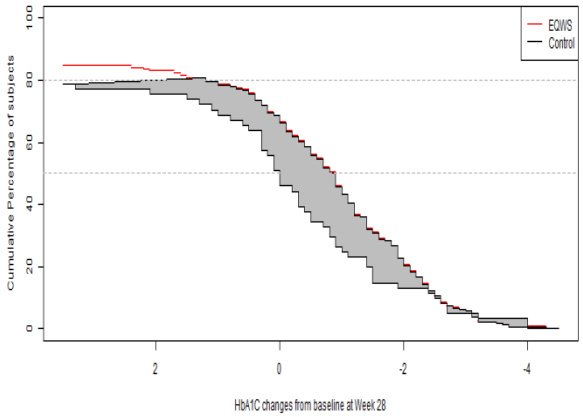
$$\mu_T^{\tau} - \mu_P^{\tau} = \frac{1}{1-\tau} \int_0^{1-\tau} \{S_T^{-1}(u) - S_P^{-1}(u)\}\, du$$

, where
S(t)=1-F(t)
$S^{-1}(u)$: the quantile function.

Difference in population trimmed means is proportional to the shaded area between the CDF curves and below the horizontal line at the height of 1- τ (Figure 13). In both BCB118 and BCB120, this reviewer chose the fraction, τ, as the maximum dropout rates (20%).

Figure 13. Population CDFs: Left panel- BCB118 (EQWS vs BID) and Right panel- BCB120 (EQWS vs Placebo) *Upper horizontal dotted line: the adaptive proportion of fraction (80%) ;Lower horizontal dotted line: 50%





Source: Statistical Reviewer's analysis

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

YOONHEE KIM
09/14/2017

YUN WANG
09/14/2017