NCT03712449

Study ID: CMO-MA-FAS-0579

Title: Canada Global HARMONY: Prospective, Multi-site, Study to Evaluate Subject Satisfaction With Facial Appearance Overall and the Aesthetic and Psychosocial Impact of Combined Facial Treatment

Statistical Analysis Plan Date: 25Aug2020

[redacted]

1 TITLE PAGE

[redacted]

Canada Global HARMONY: Prospective, Multi-site Study to Evaluate Subject Satisfaction with Facial Appearance Overall and the Aesthetic and Psychosocial Impact of Combined Facial Treatment

STATISTICAL ANALYSIS PLAN - Clinical Study Report

Version: 2.0
Draft 0.1: 08/25/2020

SIGNATURE PAGE

Written by:

[redacted]

August 26, 2020

Date (DD/MM/YYYY)

Medical Science Statistics
Allergan Pharmaceutical

Reviewed and Approved at Allergan Pharmaceutical by:

[redacted]

August 26, 2020

Date (DD/MM/YYYY)

AVP, Medical Science Statistics
Allergan Pharmaceutical

2 TABLE OF CONTENTS

1 TITLE PAGE 1
2 TABLE OF CONTENTS 4
3 LIST OF ABBREVIATIONS 6
4 INTRODUCTION 7
5 OBJECTIVES 9
5.1 Primary Objective 9
5.2 Secondary Objectives 9
6 PATIENT POPULATIONS 10
6.1 Full Analysis Populations 10
6.2 Evaluable Population 10
6.3 Subgroups 10
7 SUBJECT DISPOSITION 11
8 DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS 12
8.1 Demographics 12
8.2 Baseline characteristics 12
8.3 Medical History 12
8.4 Concomitant Medications 12
8.5 Clinical Scales for Treatment Inclusion Only 12
9 EXTENT OF EXPOSURE AND TREATMENT COMPLIANCE 13
10 EFFICACY ANALYSES 14
10.1 FACE-Q™: Description 14
10.2 Primary Efficacy Parameter 15
10.3 Secondary Efficacy Parameters 15
10.3.1 Self-Perception of Age Questionnaire 15
[redacted]
11 SAFETY ANALYSES 18
11.1 Adverse Events 18
12 HEALTH OUTCOMES ANALYSES 20
13 INTERIM ANALYSIS 21
14 DETERMINATION OF SAMPLE SIZE 22
15 STATISTICAL SOFTWARE 23
16 DATA HANDLING CONVENTIONS 24

16.1 Visit Time Windows 24
16.2 Handling of dropouts missing data 24
16.2.1 Handling of missing or incomplete dates 24
16.3 Missing Severity Assessment for Adverse Events 24
17 DATA COLLECTED BUT NOT ANALYZED 25
18 CHANGES TO ANALYSES SPECIFIED IN PROTOCOL 26
19 REFERENCES 27
1. *Development and Psychometric Evaluation of the FACE-Q Satisfaction with Appearance Scale A New Patient-Reported Outcome Instrument for Facial Aesthetics Patients, Andrea L. Pusic, MD, MD, MHS, FRCSC, Anne F. Klassen.* 27
2. *Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods, Health Technology Assessment 2009; Vol. 12 No. 12, J Hobart and S Cano* 27

3 LIST OF ABBREVIATIONS

Abbreviation	Definition
AE	adverse event
BMI	body mass index
CFLs	crow's feet lines (also known as lateral canthal lines in BOTOX product literature)
DCA	deoxycholic acid
eCRF	electronic case report form
EOS	End of Study
FACE-Q	A clinically useful, patient-reported outcome measure (PROM) of satisfaction and quality of life following elective surgical and nonsurgical facial rejuvenation.
FWS	facial wrinkle scale
GAIS	global aesthetic improvement scale
GLs	glabellar lines
ICF	informed consent form
ISR	injection site reaction
PAAQ	periorbital aesthetic appearance questionnaire
SAE	serious adverse event
SAP	statistical analysis plan
SMF	submental fat
SMFRS	submental fat rating scale
SOPS	SkinMedica overall photodamage scale
SPA	self-perception of age
SSRS	subject self-rating scale
USADE	unanticipated serious adverse device effect

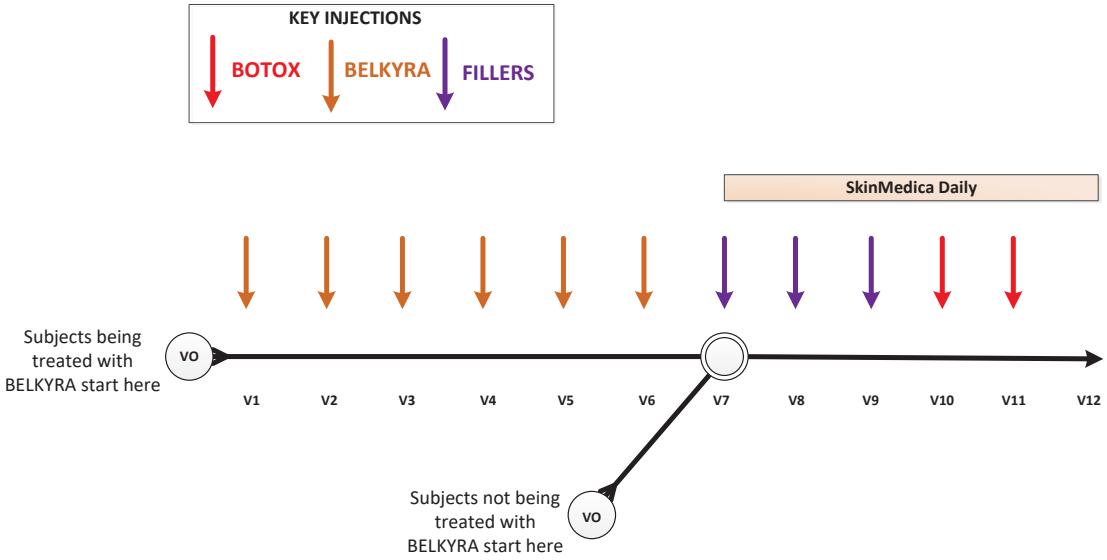
4 INTRODUCTION

The objective of the initial USA HARMONY study was to observe the safety and overall aesthetic impact of treating subjects with several modalities in an attempt to achieve subjects' desired clinical outcome. The Canada Global HARMONY study will expand upon the products tested, the facial areas treated, and the quantification of the psychological and emotional impact of the comprehensive aesthetic treatment provided and will ultimately be global in scope.

In the Global HARMONY study, subjects will be evaluated by the investigator to receive BOTOX Cosmetic, BELKYRA, Juvéderm facial fillers (JUVÉDERM VOLITE, JUVÉDERM VOLBELLA with Lidocaine, JUVÉDERM VOLIFT with Lidocaine, and JUVÉDERM VOLUMA with Lidocaine), and SkinMedica products. Selected subjects will receive staged treatment with the multiple products and will be evaluated using validated patient-reported outcome measurements to determine the psychological and emotional impact of an integrated treatment approach to reduce visible signs of aging.

The clinical hypothesis of this study is that subjects treated with a combination of BOTOX Cosmetic, BELKYRA, JUVÉDERM facial fillers, and SkinMedica products will have a significantly higher satisfaction with their overall facial appearance at the final study visit when compared with baseline, pre-treatment satisfaction scores, as measured by the mean change from baseline on the FACE-Q Satisfaction with Facial Appearance Overall Scale.

Overall Study Design



All subjects will sign an informed consent form (ICF), and begin the study at the screening visit (V0). Subjects receiving BELKYRA treatment will enter the study at V1 continuing on through V12. There will be no less than 1 month between repeat BELKYRA treatments and the final BELKYRA treatment. Subjects not receiving BELKYRA following screening (V0) will enter the

study at V7 to begin treatment with fillers, SkinMedica, and BOTOX Cosmetic as indicated in the study design diagram above.

All adverse events (AEs) including injection site reactions (ISRs) will be collected from when the ICF is signed through to the last study visit. Serious adverse events (SAEs) will be reported directly to Allergan and SkinMedica within 24 hours.

5 OBJECTIVES

5.1 Primary Objective

The primary objective of this study is to quantify the psychological, social and emotional impact of comprehensive aesthetic treatment with a portfolio of Allergan products by measuring the change in subject's satisfaction with facial appearance from baseline (before any treatment) to the final study visit (FACE-Q™: Satisfaction with Facial Appearance). These Allergan products include BOTOX® Cosmetic, BELKYRA®, Juvéderm HA facial fillers, and SkinMedica® topical nonprescription skin care products, as described in Section 9.

5.2 Secondary Objectives

Secondary objectives are:

- To determine the subject's self-appraisal of expectations of life change FACE-Q™: Expectations), age-related facial appearance (FACE-Q™: Aging Appraisal), psychological function (FACE-Q™: Psychological Function), social function (FACE-Q™: Social Function) and overall satisfaction with skin (FACE-Q™ Satisfaction with Skin Questionnaire).
- To quantify the investigator's and subject's assessment of aesthetic improvement.

6 PATIENT POPULATIONS

6.1 Full Analysis Populations

Full analysis population is defined as all subjects who have met all eligibility criteria at the screening visit (V0) and received any product used as treatment in this study.

6.2 Evaluable Population

Evaluable population is defined as all subjects as defined in the Full Analysis Population who have had at least one post-treatment efficacy assessment at the final visit (V12).

The Full Analysis Population will be used for subject disposition, demographics and baseline characteristics summary will be used for safety analyses. The Evaluable Population will be used on demographics and key efficacy endpoints as well.

6.3 Subgroups

Two subgroups will be examined separately:

1. Gender: As male and female subjects may have differing expectations/perceptions of aesthetic improvement, they will be examined separately.
2. BELKYRA: As it is expected that approximately two-thirds of the enrolled population will not be treated with BELKYRA, subjects treated and not treated with BELKYRA will be examined separately.

7 SUBJECT DISPOSITION

The number and percentage of subjects of the Full Analysis Population will be summarized by treatment group and overall. The number and percentage of patients for the following categories will be displayed:

Enrolled
Treated
  BELKYRA
  Non-BELKYRA
Completed the Study
  Yes
  No
Primary Reason for Discontinuation:
  Adverse Event
  Termination by Investigator/Sponsor
  Lost to Follow-Up
  Personal Reasons
  Protocol Violation
  Other

Screen-failure patients (i.e., patients screened but not treated) and the associated reasons for failure to enroll will be tabulated overall for the all screened patients. The number and percentage of patients who complete the V12 and of patients who prematurely discontinue during the same period will be presented for each treatment group and pooled across treatment groups for the Full Analysis Population. The reasons for premature discontinuation from the V12 as recorded on the termination pages of the electronic case report forms will be summarized (number and percentage) by treatment group for the Full Analysis Population. All patients who prematurely discontinue during the V12 will be listed by discontinuation reason for the Full Analysis Population.

8 DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS

8.1 Demographics

Age, Fitzpatrick skin type, height, weight, BMI at baseline will be summarized descriptively with Full Analysis population. Women of childbearing potential comprises women who have experienced menarche and who have not undergone successful surgical sterilization (hysterectomy, bilateral salpingectomy, or bilateral oophorectomy) or who are not post-menopausal. Number and percentage of female subjects of childbearing potential and in pregnancy are reported.

The demographic information will be listed.

8.2 Baseline characteristics

All FACE-Qs, SPA, PAAQ will be reported.

8.3 Medical History

A summary of medical history will be presented by system organ class (SOC) and preferred term (PT) using Medical Dictionary for Regulatory Affairs® (MedDRA) Version 21.0 or higher.

In addition, medical and surgical history will be listed.

8.4 Concomitant Medications

Therapy considered necessary for the subject's welfare may be given at the discretion of the investigator, including but not limited to medications for other conditions (i.e., hypertension, diabetes, etc.), treatment of AEs, estrogens, androgens, anti-androgenic agents, vitamins, iron supplements, folate, and herbal supplements. Medications should be taken consistently throughout the study and at the investigator's discretion.

Concomitant medications encompass all medicinal products that the subject was taking prior to the day 0 visit that are ongoing at the visit, in addition to all medications that have a start date on or after the day 0 visit date.

Concomitant medications will be summarized descriptively using frequency tables by ATC class and preferred name by treatment group and overall.

In addition, individual listing will be generated for all medication.

8.5 Clinical Scales for Treatment Inclusion Only

FWS, SMFRS and other inclusion/exclusion criteria will not be reported.

9 EXTENT OF EXPOSURE AND TREATMENT COMPLIANCE

Subject study product exposure and compliance will be summarized using descriptive statistics for each product, combination of multiple products, and overall exposure respectively during the treatment period.

10 EFFICACY ANALYSES

The primary efficacy endpoint is the change from baseline to the final visit in the overall Rasch-transformed scores of the FACE-Q™ Satisfaction with Facial Appearance Scale. The scale data will be analyzed by using paired t-test. If normality is unmet, Wilcoxon signed-rank test is used to replace paired t-test. If skewness of FACE-Q change from baseline is not close to zero, the change data of FACE-Q is not considered as normally distributed. In this case, one-way nonparametric procedure with Wilcoxon option will be used to replace paired t-test.

One-sample paired t-test is performed on changes at end of study. The following statistics will be provided;

Label of Statistics	Description
N	Number of subjects of overall or a subgroup
Mean	Mean change of FACE-Q™ Scale
SD	Standard deviation of Mean Change
SEM	Standard error of Mean Change
Median	Median of Changes
Minimum, Maximum	Minimum, Maximum (of Changes)
Pr > |t|	p value of Mean Change through paired t-test, or that of Wilcoxon signed-rank test
95% CIs	95% Confidence Intervals of the Mean, lower and upper limits

The efficacy analyses will be based on the Evaluable population. Baseline for efficacy is defined as the last nonmissing assessment before the first dose of study treatment. All statistical tests will be 2-sided hypothesis tests performed at the 5% level of significance for main effects. All confidence intervals will be 2-sided 95% confidence intervals, unless stated otherwise.

Hypothesis testing will be performed at a 0.05 significance level (2-sided). Since BELKYRA and without BELKYRA are the major treatment cohorts, unless specified, all endpoints will be summarized by those two groups (BELKYRA and without BELKYRA) and overall respectively. However, this study is not designed to compare the efficacy and safety outcomes between subjects treated with BELKYRA and those treated without BELKYRA. All discussion and interpretation related to study outcomes will be mainly focused on the overall results.

10.1 FACE-Q™: Description

The FACE-Q™ was developed in response to a lack of well-defined and reliable PRO instruments for facial aesthetic patients. It is made up of a set of more than 40 independently functioning scales and checklists measuring concepts and symptoms important to facial

aesthetic patients in relation to different facial areas. The FACE-Q™ can be used with any facial aesthetic patient (surgical and nonsurgical) to measure what patient's think about their appearance, quality of life, adverse effects, and process of care. Each scale provides a standalone score from 0 to 100, with higher scores indicating a better outcome. Depending on the surgical or nonsurgical procedure, only those FACE-Q™ scales and/or checklists relevant to a patient or procedure(s) need be completed.

Primary Efficacy Endpoint is FACE-Q™ Satisfaction with Facial Appearance.

The other efficacy endpoints include the following FACE-Q™ questionnaires;

1. FACE-Q™ Expectations Scale
2. FACE-Q™ Aging Appraisal
3. FACE-Q™ Psychological Function Scale
4. FACE-Q™ Social Function Scale
5. FACE-Q™ Satisfaction with Skin

10.2 Primary Efficacy Parameter

The primary endpoint is the change from baseline in the Rasch-transformed score of the FACE-Q™ Satisfaction with Facial Appearance Overall Scale to the final study visit.

The primary efficacy parameter will be summarized by treatment group and overall, as well as by subgroup being defined in Section 6.3 Subgroup. Two treatment groups are defined as BELKYRA and without-BELKYRA. Overall efficacy is summarized by both groups.

Subgroups of gender and treatment are treated as separate cohorts. There is no between groups analysis to be performed.

10.3 Secondary Efficacy Parameters

The same analysis procedures and methods used for the primary efficacy endpoint will be applied for the secondary efficacy endpoints, which include FACE-Q Expectations, FACE-Q Aging Appraisal, FACE-Q Psychological Function Scale, FACE-Q Social Function scale and FACE-Q Satisfaction with Skin.

10.3.1 Self-Perception of Age Questionnaire

Subject's assessment of age-related facial appearance by the self-perception of age (SPA) measure at baseline before any treatment to final study visit (V12); responders are defined as having achieved a younger category. No statistical testing will be performed for SPA data distribution.

Summaries of responders are provided in frequency and percentage at baseline and final visit for all three questions (a. I look my current age, b. I look years younger and c. I look years older). A shift table to summarize changes into an improvement category from baseline to the study end is provided in frequency and percentage. No statistical testing for shift tables is performed.

Subgroups in analysis are treated as individual cohorts. No comparison between treatments is performed.

10.3.2 Other Efficacy Endpoints

In general, summaries are provided in frequency and percentage for both of assessments by investigators and subjects, and by subgroups at V12. No statistical testing will be performed for other efficacy endpoints.

10.3.2.1 Global Aesthetic Improvement Scale (GAIS)

There are two assessment categories in GAIS at V12; Subject rating and Investigator rating. Descriptive statistics in frequency and percentage of scales at end of the study, in the 5-point scale from -2 to +2 are provided for both of assessments by investigators and subjects, and by subgroups as well.

10.3.2.2 Periorbital Aesthetic Appearance Questionnaire (PAAQ)

Descriptive summaries are provided in frequency and percentage for nine (9) questions by treatment and subgroups at baseline and end of the study, with five categories of answers for each question. Change in subject's satisfaction with appearance of periorbital area as measured by the periorbital aesthetic appearance questionnaire (PAAQ). Transformed PAAQ scores are summarized in the same manner of the primary endpoint.

11 SAFETY ANALYSES

The safety analysis will be performed using the Full Analysis Population. The safety parameters will include adverse events (AEs). The last nonmissing safety assessment before the first dose of study treatment will be used as the baseline for all analyses of that safety parameter.

Categorical variables will be summarized by number and percentage of patients. No statistical test will be performed.

11.1 Adverse Events

All Adverse events (AEs) will be classified by Primary System Organ Class (SOC) and Preferred Term (PT) according to the Medical Dictionary for Regulatory Activities (MedDRA) Version 21.0 or newer.

TEAE is defined as an AE that was not present prior to treatment with investigational products but appeared following treatment or was present at treatment initiation but worsened after treatment or become SAE.

An SAE is defined as any AE that either led to death; resulted in a life-threatening illness or injury; resulted in a permanent impairment of a body structure or a body function; required in-patient hospitalization or prolongation of existing hospitalization; resulted in medical or surgical intervention to prevent permanent impairment to body structure or a body function; led to fetal distress, fetal death or a congenital abnormality or birth defect.

AE summary tables will be presented for TEAEs only and will include the following:

- All TEAEs
- TEAEs related to study drug
- TEAEs leading to study discontinuation
- Serious/Severe TEAEs
- Serious/Severe TEAEs related to study drugs
- TEAEs leading to death.

All TEAEs will be summarized by SOC, PT and treatment group using frequency counts and percentages (i.e., number and percentage of subjects with an event). In addition, an overall summary for the categories above will be prepared by treatment group and overall.

All TEAEs will be summarized by SOC, PT, maximum severity, and treatment group using frequency counts and percentages (i.e., number and percentage of subjects with an event)

An AE will be considered as related if the event is Possibly, Probably or Definitely related. If the relationship to study drug is missing, the AE will be treated as related. For AE with severity missing, the missing severity will be imputed as severe.

For the incidence at the subject level by SOC and PT, if a subject experiences more than one event within the same SOC and PT, only one occurrence will be included in the incidence. For the incidence at the subject level by SOC, PT, and severity, if a subject experiences more than one event within the same SOC and PT, only the most severe occurrence will be included in the incidence.

For the incidence at the subject level by SOC, PT, and relationship to investigational product, if a subject experiences more than one event within the same SOC and PT, only the most related occurrence will be included in the incidence.

For the summary of number of events, if a subject experiences more than one event within the same SOC and PT, all events will be counted.

All AEs will be listed, and additional listings will be presented for AEs leading to discontinuation, serious AEs and AEs resulted in deaths.

12 HEALTH OUTCOMES ANALYSES

Health outcome analysis is not applicable to this study.

13 INTERIM ANALYSIS

Interim analysis is not applicable to this study.

14 DETERMINATION OF SAMPLE SIZE

A sample size of 48 subjects will provide 90% power to detect whether change from baseline to end-of-study visit in the FACE-Q Satisfaction with Facial Appearance Overall Scale Rasch-transformed score is different from zero, assuming a mean change of 12.0 and a standard deviation of 20.0. This calculation is based on a 1-sample, 2-sided t-test at 5% significance level. The assumptions of mean and standard deviation are based on FACE-Q results from Allergan HARMONY Study GMA-CMB-14-001. With a 20% drop-out rate (for early withdrawals and/or subjects who failed to duly complete the FACE-Q Satisfaction with Facial Appearance Overall Scale assessment at end of study), 60 subjects will be enrolled. From the 60 subjects planned to enroll in this study, approximately 20 subjects will be treated with BELKYRA.

Subjects will be assigned to study treatment only if they meet the inclusion criteria and none of the exclusion criteria. Subjects who discontinue treatment will not be replaced.

EXACT software (version 6) was used to confirm the sample size in the protocol.

15 STATISTICAL SOFTWARE

Statistical analyses will be performed using

16 DATA HANDLING CONVENTIONS

16.1 Visit Time Windows

All subjects will sign an informed consent form (ICF), and begin the study at the screening visit (V0).

Subjects receiving BELKYRA treatment will enter the study at V1 and continuing through V12. There will be no less than 1 month between repeat BELKYRA treatments and the final BELKYRA treatment. Subjects not receiving BELKYRA following screening (V0) will enter the study at V7 to begin treatment with fillers, SkinMedica, and BOTOX Cosmetic as indicated in the study design diagram above.

16.2 Handling of dropouts missing data

After the subjects exit the study, the values for missed future visits will not be imputed. For subjects who discontinued from the study, the efficacy data collected on the exit visit will be mapped to scheduled visit.

16.2.1 Handling of missing or incomplete dates

Missing or incomplete dates will not to be imputed. There are cases where dates are missing for which visits and results are also missing, imputation will only be done on visits and results, but not on dates.

16.3 Missing Severity Assessment for Adverse Events

If severity is missing for an AE that started before the date of the first dose of study treatment, an intensity of mild will be assigned. If severity is missing for an AE that started on or after the date of the first dose of study treatment, an intensity of severe will be assigned. The imputed values for severity assessment will be used for the incidence summary; the values will be shown as missing in the data listings.

17 DATA COLLECTED BUT NOT ANALYZED

This section is not applicable to this study.

18 CHANGES TO ANALYSES SPECIFIED IN PROTOCOL

Change to analyses specified in protocol is not applicable to this study.

19 REFERENCES

1. *Development and Psychometric Evaluation of the FACE-Q Satisfaction with Appearance Scale A New Patient-Reported Outcome Instrument for Facial Aesthetics Patients, Andrea L. Pusic, MD, MD, MHS, FRCSC, Anne F. Klassen.*

2. *Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods, Health Technology Assessment 2009; Vol. 12 No. 12, J Hobart and S Cano*