1. TITLE PAGE

Project:	CDISCPILOT01 – Initial Case Study of the CDISC SDTM/ADaM Pilot Project
Case Study Title:	Safety and Efficacy of the Xanomeline Transdermal Therapeutic System (TTS) in Patients with Mild to Moderate Alzheimer's Disease
Investigational Product:	Xanomeline Transdermal
Indication:	Alzheimer's Disease
Brief Description of Case Study:	This study was a prospective, randomized, multi-center, double-blind, placebo-controlled, parallel-group study. The objectives of the study were to evaluate the efficacy and safety of transdermal xanomeline, 50 cm^2 and 75 cm^2, and placebo in subjects with mild to moderate Alzheimer's disease.
Study Sponsor:	CDISC Pilot Project
Protocol No.:	CDISCPILOT01
Study Phase:	2
Study Initiation Date:	06 July 2012 (Date of first subject visit)
Study Completion Date:	05 March 2015 (Date of last subject completion)
Principal Investigators	Due to the nature of this CDISC Pilot Project, a list of principal investigators is not provided in this study report.
Good Clinical Practice	This study was conducted in accordance with International Conference on Harmonisation (ICH) Good Clinical Practice (GCP) regulations/guidelines. Essential documents will be retained in accordance with ICH GCP.
Report Date:	27 June 2006

2. SYNOPSIS

Name of Sponsor: CDISC Pilot Project	**Name of Finished Product:** Transdermal Xanomeline	**Name of Active Ingredient:** Xanomeline
Case Study Title: Safety and Efficacy of the Xanomeline Transdermal Therapeutic System (TTS) in Patients with Mild to Moderate Alzheimer's Disease		
Investigators and Study Centers: This study was conducted at 17 centers. Due to the nature of this CDISC Pilot Project, a list of investigators is not provided.		
Publications: Not applicable		
Study Period: 06 July 2012 to 05 March 2015	**Development Phase:** Phase 2	
Objectives: The objectives of the study were to evaluate the efficacy and safety of transdermal xanomeline, 50 cm^2 and 75 cm^2, and placebo in subjects with mild to moderate Alzheimer's disease.		
Methodology: This was a prospective, randomized, multi-center, double-blind, placebo-controlled, parallel-group study. Subjects were randomized equally to placebo, xanomeline low dose, or xanomeline high dose. Subjects applied 2 patches daily and were followed for a total of 26 weeks.		
Number of Subjects Planned: 300 subjects total (100 subjects in each of 3 groups)		
Number of Subjects Enrolled: 254 subjects were randomized (86 placebo, 84 xanomeline low dose, 84 xanomeline high dose) **Sex:** 111 (44%) Male; 143 (56%) Female **Mean (SD) Age:** 75.1 (8.25) years **Ethnicity (Race):** 218 (86%) Caucasian; 23 (9%) African Descent; 12 (5%) Hispanic; 1 (<1%) Other		
Diagnosis and Main Criteria for Eligibility: Subjects were males or females of non-childbearing potential, 50 years of age or older, had probable Alzheimer's disease according to the National Institute of Neurologic and Communicative Disorders and Stroke–Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) criteria, and an Mini-Mental State Examination (MMSE) score of 10 to 23.		
Investigational Product, Dose and Mode of Administration, Batch Number: Xanomeline transdermal patches of 50 cm^2 or 25 cm^2 in area, with 54 mg and 27 mg of xanomeline, respectively. Two patches were applied daily. Xanomeline high dose group received an active patch of each size for a total dose of 81 mg and the xanomeline		

low dose received an active large patch and a placebo small patch for a total dose of 54 mg. Due to the nature of this CDISC Pilot Project, batch numbers are not provided in this study report.

Duration of Treatment: 26 weeks of treatment

Reference Therapy, Dose and Mode of Administration, Batch Number:
Matching placebo transdermal patches of 50 cm^2 or 25 cm^2 in area. Placebo group received a placebo patch of each size. Due to the nature of this CDISC Pilot Project, batch numbers are not provided in this study report.

Criteria for Evaluation:
Primary Efficacy Endpoints:

- Alzheimer's Disease Assessment Scale - Cognitive Subscale, total of 11 items [ADAS-Cog (11)] at Week 24
- Video-referenced Clinician's Interview-based Impression of Change (CIBIC+) at Week 24

Secondary Efficacy Endpoints:

- ADAS-Cog (11) at Weeks 8 and 16
- CIBIC+ at Weeks 8 and 16
- Mean Revised Neuropsychiatric Inventory (NPI-X) from Week 4 to Week 24

Safety Endpoints:

- Adverse events
- Vital signs (weight, standing and supine blood pressure, heart rate)
- Laboratory evaluations

Statistical Methods:
Unless otherwise noted, hypothesis testing was evaluated at a significance level of 0.05. Summary statistics for continuous variables included the number of observations, mean, standard deviation (SD), median, minimum, and maximum. Summary statistics for the categorical variables included frequency and percentage.

The number of subjects randomized, the number of subjects in each analysis dataset, and the disposition of subjects were tabulated by treatment group. Specific reasons for early study discontinuation (protocol completed, lack of efficacy, and adverse event) were compared using a Fisher's exact test.

The baseline characteristics were summarized by treatment group and across all treatment groups. The treatment groups were compared by analysis of variance (ANOVA) for continuous variables and by Pearson's chi-square test for categorical variables.

The primary analysis of the ADAS-Cog (11) or CIBIC+ at Week 24 used the efficacy population with LOCF imputation for any missing values at Week 24. For ADAS-Cog (11), an analysis of covariance (ANCOVA) model was used to test for dose response with the baseline score, site, and treatment included as independent variables. A

supportive analysis for the ADAS-Cog (11) used a likelihood-based repeated measures (MMRM) analysis. For CIBIC+, an ANOVA model was used to test for dose response with site and treatment included as independent variables. Similar analyses were performed at Weeks 8 and 16 for ADAS-Cog (11) and CIBIC+. Summary statistics for ADAS-Cog (11) were also generated for each visit using the efficacy population with LOCF imputation.

The primary analysis of mean NPI-X total score from Week 4 to Week 24 used the efficacy population. For this endpoint, an ANCOVA model was used to test for dose response with the baseline score, site, and treatment included as independent variables.

Average daily dose and cumulative dose at end of study was computed for each subject based on the planned dose and the actual number of days in the study and was summarized for each treatment group.

Treatment emergent adverse events and serious adverse events were summarized by system organ class (SOC) and preferred term (PT). The incidence of treatment emergent events grouped under preferred terms for each active treatment were compared to placebo using Fisher's exact test. Additional analysis of dermatological adverse events was conducted. The time to the first dermatological event was compared across the treatment groups using Kaplan-Meier methods.

Hematology and clinical chemistry values were summarized at each visit week. The number of subjects with no abnormal measure during treatment and those with at least one abnormal measure during treatment were summarized for each lab analyte. Fisher's exact test was used to analyze the incidence of abnormal (high or low) measures during the post-randomization phase. A display summarizing shifts from baseline by week in terms of abnormality based on threshold range was provided. The data were summarized comparing baseline and on drug categorization for each treatment group for each week for each laboratory analyte. Shift tables summarizing whether a subject's status changed from baseline during the treatment period were provided for changes based on threshold ranges and changes based on Hy's Law. Cochran-Mantel-Haenszel (CMH) tests, stratifying by status at baseline, were performed.

Vital sign data and weight were summarized by treatment group. The number and percent of subjects receiving each concomitant medication were summarized.

Summary of Results:

Disposition:

A total of 254 subjects were randomized and entered the double-blind treatment phase. The number of subjects randomized to each treatment arm was: 86 to placebo, 84 to the xanomeline low dose treatment group and 84 to the xanomeline high dose treatment group. Of the 254 subjects randomized to treatment, 118 completed the treatment phase (Week 24), and 110 completed the study through Week 26. A statistically significantly

($p < 0.0001$) higher number of subjects in the xanomeline low dose and high dose groups (67% and 64%, respectively) prematurely discontinued from the study prior to Week 24 as compared to the placebo group (30%). The most common reason for discontinuation was adverse event (9% placebo subjects, 52% xanomeline low dose subjects, 46% xanomeline high dose subjects), with a statistically significant association between discontinuation due to adverse event and treatment group ($p < 0.0001$).

Efficacy Results:

A statistically significant dose response was not seen for either of the primary efficacy endpoints, changes from baseline in ADAS-Cog (11) at Week 24 and CIBIC+ at week 24. Adjusted means for these 2 endpoints were similar for all 3 treatment groups. Additional analyses at earlier time points showed similar results. Subgroup analyses by gender, a sensitivity analysis for missing data, and a repeated measures analysis for ADAS-Cog (11) also indicated lack of treatment response. The secondary efficacy endpoint of the mean NPI-X values from Week 4 through Week 24 also did not demonstrate a statistically significant dose response.

Safety Results:

Over 90% of subjects receiving active therapy reported at least 1 adverse event compared to 75.6% of subjects receiving placebo. This difference is due largely to a disproportionate number of dermatologic type events that occurred in the xanomeline treatment groups. Approximately 73% of the subjects in either of the xanomeline groups experienced at least one dermatologic adverse event of interest compared to 33.6% of the placebo subjects. There was a statistically significant difference ($p < 0.001$) in the time to first dermatologic event between the treatment groups. There were 3 deaths (2 in placebo group, 1 in the xanomeline low dose group) observed during the study. None of the deaths were judged related to treatment. Aside from the deaths, there were 3 serious adverse events reported in 3 subjects (2 in xanomeline high dose and 1 in the xanomeline low dose group) and all were related to the nervous system.

The association between treatment group and the number of abnormal values beyond the normal range was significant for three laboratory analytes: albumin ($p = 0.042$), urea nitrogen ($p = 0.023$), and eosinophils ($p = 0.001$). The association between clinically significant changes from the previous visit and treatment was statistically significant for aspartate aminotransferase ($p = 0.045$) and eosinophils ($p = 0.010$). The analysis of shifts from baseline to most abnormal value could not be calculated on 19 of the analytes. Of the remaining 11 analytes, only eosinophils showed a statistically significant association with treatment group ($p = 0.044$). There was no significant association with treatment group in the Hy's law analyses examining shifts in transaminase levels, and transaminase and total bilirubin levels between baseline values and values while on treatment.

Changes from baseline in vital signs (SBP, DBP, and pulse), at the Week 24 and end of treatment assessments, were generally small decreases. Changes from baseline in

weight, at the Week 24 and end of treatment assessments, however, were generally small with no treatment-related pattern of increases or decreases.

Conclusions:
A statistically significantly higher proportion of subjects in the active treatment groups withdrew prematurely from the study as compared to the placebo group. This is largely due to the higher proportion of subjects in the active treatment groups experiencing a dermatologic event and subsequently resulting in premature withdrawal from the study. This further hindered the study's ability to demonstrate efficacy.

A statistically significant dose response was not seen for both of the primary efficacy endpoints, change from baseline in ADAS-Cog (11) at Week 24 and CIBIC+ at Week 24, and for the secondary efficacy endpoint, mean NPI-X values from Week 4 to Week 24. Adjusted means for all 3 endpoints were similar across all treatment groups.

There were an increased number of dermatologic adverse events reported in the active treatment groups as compared to the placebo group. There were 3 serious adverse events. In addition, there were 3 deaths that were deemed unrelated to treatment.

For the laboratory data, subjects in both the xanomeline low and high dose groups showed more observations above normal range than the placebo group. Albumin was more often lower than the normal range for subjects in the placebo and xanomeline low dose group. Subjects in the xanomeline treatment groups had statistically significantly more values above the normal range than subjects in the placebo group for both urea nitrogen and eosinophils. There was a statistically significant association between clinically significant changes from the previous visit and treatment group for aspartate aminotransferase and eosinophils. Shifts from baseline for eosinophils were statistically significant with both xanomeline treatment groups showing more changes from normal to above normal than the placebo group. There was no significant association with treatment group in the Hy's law analysis examining shifts in liver function tests between baseline values and values while on treatment.

There were only minor changes from baseline in vital signs and weight at Week 24.

Report Date: 27 June 2006

3. TABLE OF CONTENTS

Page

3.1. List of Tables

3.2. List of Figures

4. LIST OF ABBREVIATIONS

AD	Alzheimer's Disease
ADaM	Analysis Dataset Model
ADAS-Cog	Alzheimer's Disease Assessment Scale - Cognitive Subscale
AE	adverse event
ALT	alanine aminotransferase (also known as SGPT [serum glutamic pyruvic transaminase])
ANCOVA	analysis of covariance
ANOVA	analysis of variance
AST	aspartate aminotransferase (also known as SGOT [serum glutamic oxalacetic transaminase])
BMI	body mass index
BPM	beats per minute
CDISC	Clinical Data Interchange Standards Consortium
CI	confidence interval
CIBIC+	Video-referenced Clinician's Interview-based Impression of Change
cm^2	centimeters squared – measure of area
CMH	Cochran-Mantel-Haentzel
CNS	central nervous system
DBP	diastolic blood pressure
ECG	Electrocardiogram
FDA	Food and Drug Administration
GGT	gamma-glutamyl transferase (also known as GGPT [gamma glutamyl transpeptidase]; SGGT [serum gamma gutamyl transferase]; YGGT)
IEC	Independent Ethics Committee
IRB	Institutional Review Board
ITT	Intent-to-Treat
kg	kilograms
LLN	lower limit of normal
LOCF	last observation carried forward
MCH	mean corpuscular hemoglobin
MCHC	mean corpuscular hemoglobin concentration
MCV	mean corpuscular volume
mg	milligrams
mmHg	millimeters of mercury
MMSE	Mini-Mental State Examination
NINCDS-ADRDA	National Institute of Neurologic and Communicative Disorders and Stroke–Alzheimer's Disease and Related Disorders Association (developed criteria for the diagnosis of Alzheimer's disease)
NPI-X	Revised Neuropsychiatric Inventory
PT	preferred term

SBP	systolic blood pressure
SD	standard deviation
SDTM	Study Data Tabulation Model
SOC	system organ class
TTS	Transdermal Therapeutic System
ULN	upper limit of normal
XAN	xanomeline

5. ETHICS

5.1. Independent Ethics Committee (IEC) or Institutional Review Board (IRB)

Due to the nature of this CDISC Pilot Project, this section is not included in this study report.

5.2. Ethical Conduct of the Study

Due to the nature of this CDISC Pilot Project, this section is not included in this study report.

5.3. Subject Information and Consent

Due to the nature of this CDISC Pilot Project, this section is not included in this study report.

6. INVESTIGATORS AND STUDY ADMINISTRATIVE STRUCTURE

Due to the nature of this CDISC Pilot Project, this section is not included in this study report.

7. INTRODUCTION

The objective of the CDISC SDTM/ADaM Pilot Project is to demonstrate the effective transformation of legacy data into CDISC SDTM domains and ADaM datasets and their associated metadata. The Project team will produce a "pilot submission" that will be delivered to FDA reviewers for their evaluation in a mock review, assessing whether data submitted to the FDA using the CDISC Standard will meet the needs and expectations of both medical and statistical FDA reviewers. This abbreviated study report documents the analysis results of the legacy data for this first pilot submission.

The legacy data being used in CDISCPILOT01 were provided by Eli Lilly and Company (Legacy Sponsor) for the purposes of this CDISC Pilot Project. The data were de-identified and documents were redacted prior to release to the CDISC Pilot Project team. De-identification included changing dates and shifting them into the future. All chronological relationships and sequences were maintained within the data elements for a subject (e.g., no change in the relationship of timing of adverse events with respect to dosing). The submission did not reproduce all of the Legacy Sponsor's analyses and reports. Instead only the more common elements of a submission were addressed. These included safety data, the primary efficacy endpoints and a few secondary efficacy endpoints. Deviations from the protocol-specified analyses are described in the statistical analysis plan created specifically for this study as part of the CDISC Pilot Project (Appendix 9).

8. STUDY OBJECTIVES

The primary objectives of this study were:

- To determine if there is a statistically significant relationship (overall Type 1 error rate, α=0.05) between the change in both the ADAS-Cog (11) and CIBIC+ scores, and drug dose (0, 50 cm^2 [54 mg], and 75 cm^2 [81 mg]).
- To document the safety profile of the xanomeline TTS.

A secondary objective of this study was:

- To assess the dose-dependent improvement in behavior. Improved scores on the Revised Neuropsychiatric Inventory (NPI-X) will indicate improvement in these areas.

9. INVESTIGATIONAL PLAN

9.1. Overall Study Design and Plan

This study was a prospective, randomized, multi-center, double-blind, placebo-controlled, parallel-group study. The objectives of the study were to evaluate the efficacy and safety of transdermal xanomeline, 50 cm^2 and 75 cm^2, and placebo in subjects with mild to moderate Alzheimer's disease.

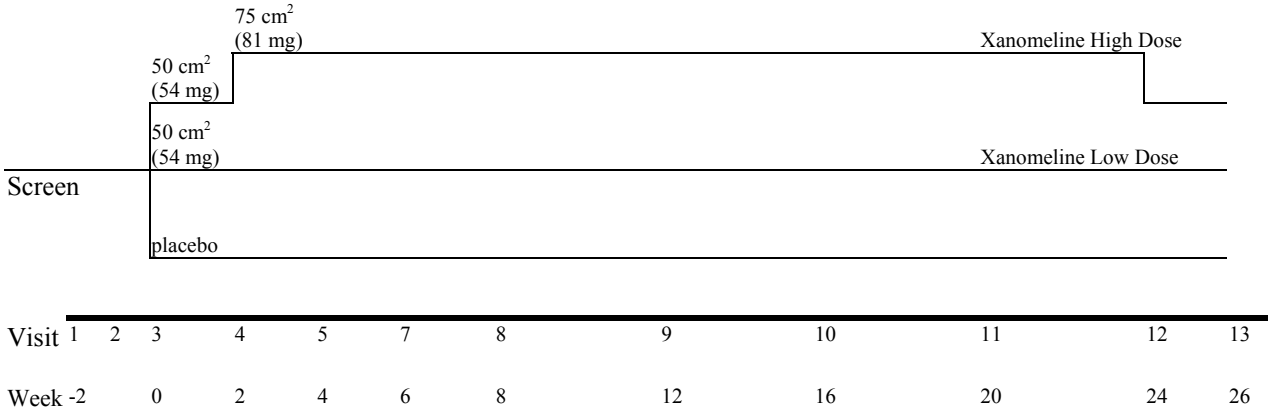
Xanomeline or placebo was administered daily in the morning, with the application of two adhesive patches, one 50 cm^2 in area, the other 25 cm^2 in area. Doses were measured in terms of the xanomeline base, and were 54 mg for the 50 cm^2 patch and 27 mg for the 25 cm^2 patch. Placebo was identical in appearance to the primary study material. The total doses being compared are therefore 0 (both patches placebo), 54 mg (large patch active drug, small patch placebo), and 81 mg (both patches active drug). The treatment groups referred to throughout this report will be "xanomeline high dose," "xanomeline low dose," and "placebo".

Subjects were males or females of non-childbearing potential, 50 years of age or older, had probable mild to moderate Alzheimer's disease according to the NINCDS-ADRDA criteria, and an MMSE score of 10 to 23. The duration of treatment was 26 weeks, with 24 weeks of active treatment. Approximately 300 subjects were to be enrolled and randomized equally to xanomeline high dose, xanomeline low dose, or placebo.

Subjects were assessed for efficacy using the Alzheimer's Disease Assessment Scale - Cognitive Subscale of 11 items [ADAS-Cog (11)], video-referenced Clinician's Interview-based Impression of Change (CIBIC+), and Revised Neuropsychiatric Inventory (NPI-X). Safety assessments include reporting of adverse events, laboratory values, and vital signs.

The schema for this study is illustrated in Figure 9-1. Additional study design details are described in the study protocol (Appendix 1).

Figure 9-1. Study Schema



9.2. Discussion of Study Design, Including the Choice of Control Groups

This study was designed to evaluate the safety and efficacy of the low and high dose xanomeline relative to placebo in subjects with mild to moderate Alzheimer's disease. Subjects were randomized to 1 of 3 treatment groups: placebo, xanomeline low dose and xanomeline high dose. Two patches were administered to each subject in a double-blind fashion to minimize investigator and subject bias. Placebo was the control group used in this study.

9.3. Selection of Study Population

Subjects were included in the study if they were males or females of non-childbearing potential, 50 years of age or older, had probable mild to moderate Alzheimer's disease according to the NINCDS-ADRDA criteria, had an MMSE score of 10 to 23, had a Hachinski Ischemic Scale score of ≤ 4, and CNS imaging compatible with Alzheimer's disease within the past year. Subjects were excluded from the study if they had previously participated in a xanomeline study, had used an investigational or approved Alzheimer's therapeutic medication within 30 days of prior to enrollment, serious illness requiring hospitalization within 3 months prior to screening, have certain concurrent or historical medical conditions, or were concurrently or historically using certain medications. Details of the inclusions and exclusion criteria are included in the study protocol (Appendix 1).

9.4. Treatments

Subjects were randomized to 1 of the 3 treatment groups: placebo, xanomeline low or high dose. Xanomeline or placebo was administered daily in the morning, with the application of two adhesive patches, one 50 cm^2 in area, the other 25 cm^2 in area. Doses were measured in terms of the xanomeline base, and were 54 mg for the 50 cm^2 patch and 27 mg for the 25 cm^2 patch. Placebo was identical in appearance to the primary study material. Xanomeline high dose group received an active patch of each size and the xanomeline low dose received an active large patch and a placebo small patch. The total doses being compared are therefore 0 (both patches placebo), 54 mg (large patch active drug, small patch placebo), and 81 mg (both patches active drug). The treatment groups referred to throughout this report will be "xanomeline high dose," "xanomeline low dose," and "placebo".

9.5. Efficacy and Safety Variables

Efficacy in this study was assessed using Alzheimer's Disease Assessment Scale - Cognitive Subscale (ADAS-Cog), Video-referenced Clinician's Interview-based Impression of Change (CIBIC+), and Revised Neuropsychiatric Inventory (NPI-X). The ADAS-Cog is an established measure of cognitive function in Alzheimer's disease. This study will specifically use an 11-item subscale of the ADAS-Cog, denoted as ADAS-Cog (11). The CIBIC+ is an assessment of the global clinical status relative to baseline and utilizes semi-structured interviews with the subject and the caregiver. The NPI-X is an assessment of change in psychopathology in subjects with dementia and is administered to the designated caregiver. Methods for scoring the ADAS-Cog (11), CIBIC+, and NPI-X are noted in the statistical analysis plan (Appendix 9).

Safety in this study was assessed with the reporting of adverse events, laboratory measures, and vital signs. The list of laboratory measures is noted in the study protocol (Appendix 1).

The schedule of efficacy and safety assessments is listed in the study protocol (Appendix 1). Additional efficacy and safety assessments were utilized in this protocol, but were not included in this study report due to the nature of this CDISC Pilot Project.

9.6. Data Quality Assurance

Due to the nature of this CDISC Pilot Project, this section is not included in this study report.

9.7. Statistical Methods Planned in the Protocol and Determination of Sample Size

9.7.1. Statistical Methods

This section summarizes the statistical methods used to analyze the data for this report. Unless otherwise noted, hypothesis testing was evaluated at a significance level of 0.05. Summary statistics for continuous variables included the number of observations, mean, standard deviation (SD), median, minimum, and maximum. Summary statistics for the categorical variables included frequency and percentage. Additional details can be found in the statistical analysis plan (Appendix 9).

9.7.1.1. Endpoints

The primary efficacy endpoints were:

- Alzheimer's Disease Assessment Scale - Cognitive Subscale, total of 11 items [ADAS-Cog (11)] at Week 24
- Video-referenced Clinician's Interview-based Impression of Change (CIBIC+) at Week 24

The secondary efficacy endpoints were:

- Alzheimer's Disease Assessment Scale - Cognitive Subscale, total of 11 items [ADAS-Cog (11)] at Weeks 8 and 16
- Video-referenced Clinician's Interview-based Impression of Change (CIBIC+) at Weeks 8 and 16
- Mean Revised Neuropsychiatric Inventory (NPI-X) from Week 4 to Week 24

The safety endpoints were:

- Adverse events
- Vital signs (weight, standing and supine blood pressure, heart rate)
- Laboratory evaluations

9.7.1.2. Disposition and Baseline Characteristics

The number of subjects randomized, the number of subjects in each analysis dataset, and the disposition of subjects were tabulated by treatment group. Specific reasons for early study discontinuation (protocol completed, lack of efficacy, and adverse event) were compared using a Fisher's exact test.

The following baseline characteristics were summarized by treatment group and across all treatment groups: age, age category (<65, 65-80, >80), sex, race, Mini-Mental State (MMSE), duration of Alzheimer's disease, years of education, weight, height, BMI, and BMI category (BMI<25, BMI 25-<30, BMI>=30). The treatment groups were compared by analysis of variance (ANOVA) for continuous variables and by Pearson's chi-square test for categorical variables.

9.7.1.3. Efficacy

The primary analysis of the ADAS-Cog (11) at Week 24 used the efficacy population with LOCF imputation for any missing values at Week 24. An ANCOVA model was used with the baseline score, site, and treatment included as independent variables. Treatment was included as a continuous variable, and results for a test of dose response were produced. If the test for dose response was statistically significant, pairwise comparisons among the 3 groups were to be performed and evaluated at a significance level of 0.05. Similar analyses were performed at Weeks 8 and 16. Summary statistics were generated for each visit using the efficacy population with LOCF imputation.

A supportive analysis for the ADAS-Cog (11) used a likelihood-based repeated measures (MMRM) analysis. In this analysis for the change from baseline in the ADAS-Cog (11) at Week 24, the independent variables included in the model were the fixed, categorical effects of treatment, site, time (week), and treatment by time interaction along with the continuous effects of baseline ADAS-Cog (11) score and baseline ADAS-Cog (11) score by time interaction.

The primary analysis of CIBIC+ at Week 24 used the efficacy population with LOCF imputation for any missing values at Week 24. For this endpoint, an ANOVA model was used with site, and treatment included as independent variables. Treatment was included as a continuous variable, and results for a test of dose response were produced. If the test for dose response was statistically significant, pairwise comparisons among the 3 groups were to be performed and evaluated at a significance level of 0.05. Similar analyses were performed at Weeks 8 and 16.

The primary analysis of mean NPI-X total score from Week 4 to Week 24 used the efficacy population. This endpoint was calculated as the mean of all available total scores between Weeks 4 and 24, inclusive. For this endpoint, an ANCOVA model was used with the baseline score, site, and treatment included as independent variables. Treatment was included as a continuous variable, and results for a test of dose response were produced. If the test for dose response was statistically significant, pairwise comparisons among the 3 groups were to be performed and evaluated at a significance level of 0.05.

9.7.1.4. Safety

Average daily dose and cumulative dose at end of study (Week 26 or early termination) was computed for each subject based on the planned dose and the actual number of days in the study and was summarized for each treatment group.

Adverse events were coded according to MedDRA. Due to the nature of the CDISC Pilot Project, the higher level terms and higher level group terms of the MedDRA coding were masked. In addition, no numeric MedDRA codes are included in the databases. Treatment emergent adverse events and serious adverse events were summarized by SOC (system organ class) and preferred term (PT). The incidence of treatment emergent events grouped under preferred terms for each active treatment were compared to placebo using Fisher's exact test.

Additional analysis of dermatological adverse events was conducted. A category of special events was created to identify the events that were considered dermatological events. These events were determined by a thorough review of blinded coded adverse event terms and all preferred terms that were considered to be dermatologic in nature, such as rash, pruritus, or dermatitis, were flagged as adverse events of special interest. The time to the first dermatological event was compared across the treatment groups using Kaplan-Meier methods. Graphical displays of the survival curves were produced.

Hematology and clinical chemistry values were summarized at each visit week, for each analyte, for each treatment group. Four assessments of abnormality were identified for each laboratory analyte:

- Values outside the normal range
- Values significantly beyond the normal range (i.e., outside the threshold range)
- Values differing significantly from values at the previous scheduled visit,
- Abnormal values as defined by Hy's Law

The number of subjects with no abnormal measure during treatment and those with at least one abnormal measure during treatment were summarized for each lab analyte. Two tables were provided – one defining abnormal as beyond normal range (i.e., below LLN or above ULN) and the other defining abnormal as a clinically significant change from the previous visit. Fisher's exact test was used to analyze the incidence of abnormal (high or low) measures during the post-randomization phase.

A display summarizing shifts from baseline by week in terms of abnormality based on threshold range was provided. The data were summarized using sets of 3-by-3 matrices comparing baseline and on drug categorization for each treatment group for each week for each laboratory analyte.

Shift tables summarizing whether a subject's status changed from baseline during the treatment period were provided for changes based on threshold ranges and changes based on Hy's Law. Two variations of the modified Hy's Law criteria were used in the

assessment. The first considered subjects with transaminase (ALT or AST) elevations of greater than 1.5 times ULN as abnormal. The second further narrowed the assessment of abnormality to require total bilirubin elevations to be greater than 1.5 times ULN in addition to transaminase elevations of greater than 1.5 times ULN. In these tables a subject was categorized as normal or abnormal (i.e., outside the threshold range) at baseline. During the treatment phase, the most extreme value was used to categorize a subject as normal or abnormal during the treatment phase. The shift table shows the number of subjects whose on treatment categorization was the same or shifted from the baseline categorization. The treatment period was defined as any planned visit after Week 0 (Visit 3), up to and including Week 24 (Visit 12). Cochran-Mantel-Haenszel (CMH) tests, stratifying by status at baseline, were performed.

Vital sign data (blood pressure supine, blood pressure standing 1 minute, blood pressure standing 3 minutes, heart rate supine, heart rate standing 1 minute, and heart rate standing 3 minutes) at baseline, week 24 and end of treatment (last visit on or before week 24 visit) was summarized by treatment group. Change from baseline was also summarized.

Weight data at baseline and Week 24 (with and without including early terminations) was summarized by treatment group. Change from baseline was also summarized.

The concomitant medication data were coded using a publicly available sample of WHO Drug. The data were matched to a preferred term and an anatomical class (ATC level 1). Drugs not matching those in the sample were considered "uncoded" for the purposes of this submission. Due to the nature of this CDISC Pilot Project, drugs were matched to only one class. The number and percent of subjects receiving each concomitant medication were summarized. Concomitant medications were reported by anatomical class and ingredient. Medications were sorted in descending order of total incidence across treatment groups for anatomical class and in descending order of total incidence for the ingredient within each anatomical class. If the total incidence for any two or more ingredients is equal, the events were presented in alphabetical order.

9.7.2. Determination of Sample Size

Approximately 100 subjects were to be randomized to each of the 3 treatment groups. Previous experience with the oral formulation of xanomeline suggested that this sample size had 90% power to detect a 3.0 mean treatment difference in ADAS-Cog ($p < 0.05$, two-sided), based on a standard deviation of 6.5. Furthermore, this sample size had 80% power to detect a 0.36 mean treatment difference in CIBIC+ ($p < 0.05$, two-sided), based on a standard deviation of 0.9.

9.8. Changes in Study Conduct or Planned Analyses

The protocol was amended 3 times. For the first 2 amendments, changes were made to the ambulatory ECG assessments. These changes included shortening the duration of the ambulatory ECG monitoring from 48 to 24 hours prior to visit 3, adding monitoring prior to visit 4, removing monitoring prior to visit 6, and subsequently removing the monitoring prior to visit 4. During the time the additional monitoring prior to visit 4 was required, a visit (designated as visit 3e) was added for the placement of the ambulatory ECG equipment. Other changes in the study conduct are not described in this study report due to the nature of this CDISC Pilot Project.

Changes to the protocol-specified analyses are described in the statistical analysis plan (Appendix 9). These changes included omission of secondary or sensitivity analyses, omission of some efficacy endpoints, omission of some safety endpoints, and the inclusion of additional types of analyses.

10. STUDY SUBJECTS

10.1. Disposition of Subjects

Table 14-1.01 provides a summary of the subjects in each analysis population. A total of 306 subjects have demographic information in the study tabulation database (Figure 10-1). Fifty-two (52) subjects were not randomized, and thus were screen failures. The remaining 254 subjects were randomized and entered the double-blind treatment phase. These subjects comprise the Intent-to-Treat population. The number of subjects randomized to each treatment arm was: 86 to placebo, 84 to the xanomeline low dose treatment group and 84 to the xanomeline high dose treatment group. Of the 254 subjects randomized to treatment, 118 completed the treatment phase (Week 24), and 110 completed the study through Week 26.

Figure 10-1. Subject Disposition

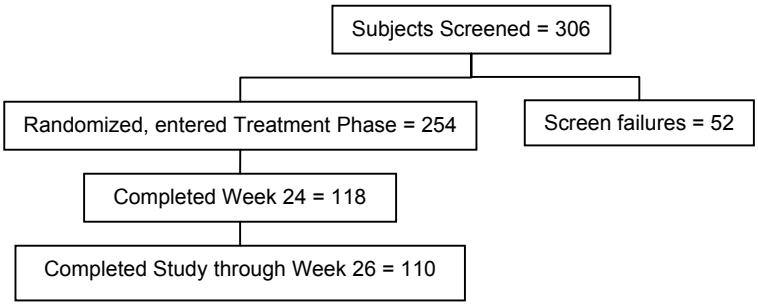


Table 14-1.03 summarizes the disposition of the population by site. The table provides the number of subjects randomized and completing Week 24, per treatment group, for each site. Of the 17 sites participating in the study, 7 met the pre-specified criteria for small sample sizes, requiring them to be grouped together for the purposes of analyses including site as a covariate, as shown in Table 14-1.03.

A total of 136 subjects prematurely discontinued from the study prior to Week 24. Table 14-1.02 provides a summary of the reasons for premature discontinuation for these subjects. There was a statistically significant association between the number of subjects discontinuing the study and the treatment group ($p < 0.0001$), with 30% of placebo subjects terminating early as compared to 67% of the xanomeline low dose subjects, and 64% of the xanomeline high dose subjects. The most common reason for discontinuation was adverse event (9% placebo subjects, 52% xanomeline low dose subjects, 46% xanomeline high dose subjects), with a statistically significant association between discontinuation due to adverse event and treatment group ($p < 0.0001$).

10.2. Protocol Deviations

Due to the nature of this CDISC Pilot Project, this section is not included in this study report.

11. EFFICACY EVALUATION

11.1. Data Sets Analyzed

Table 14-1.01 provides the summary of study populations analyzed in this protocol. Subjects were randomly assigned to treatment groups at Week 0 (Visit 3). All subjects who were randomized comprised the Intent-to-Treat (ITT) Population. A total of 254 subjects (86 in the placebo group; 84 in the xanomeline low dose group; 84 in the xanomeline high dose group) were randomized.

All subjects randomized and known to have taken at least one dose of randomized drug were included in the Safety Population. Of the 254 subjects randomized, all took a dose of the randomized study drug and were included in the Safety Population. Note that the first patches were applied at the randomization visit.

All subjects who were randomized and took at least one dose of randomized drug (i.e., were in the Safety Population), and have at least one post-baseline measure for both ADAS-Cog and CIBIC+ were included in the Efficacy Population. Twenty (20) of the subjects in Safety Population did not have a post-randomization ADAS-Cog assessment and a post-randomization CIBIC+ assessment, thus 234 subjects (79 placebo; 81 xanomeline low dose; 74 xanomeline high dose) comprised the Efficacy Population.

The 118 subjects (60 placebo; 28 xanomeline low dose; 30 xanomeline high dose) who were in the Efficacy Population and completed their Week 24 visit (Visit 12) comprised the Completers Population. Eight (8) subjects discontinued between Week 24 and Week 26, leaving 110 subjects (58 placebo; 25 xanomeline low dose; 27 xanomeline high dose) who completed the study.

11.2. Demographics and Other Baseline Characteristics

11.2.1. Demographic Characteristics

Table 14-2.01 summarizes demographic characteristics of age, sex, race and education level for all ITT subjects in the study. The study population was similar across all treatment groups (Table 11-1).

Table 11-1. Demographic Characteristics

	Placebo (N=86)	Xan Low (N=84)	Xan High (N=84)	Total (N=254)
Age (years), mean (range)	75.2 (52-89)	75.7 (51-88)	74.4 (56-88)	75.1 (51-89)
Gender (%)				
Male	38%	40%	52%	44%
Female	62%	60%	48%	56%
Race (%)				
White/Caucasian	87%	86%	85%	86%
Other	13%	14%	15%	14%
Education (years), mean (range)	12.6 (6-21)	13.2 (3-24)	12.5 (6-20)	12.8 (3-24)

Source: Table 14-2.01

11.2.2. Baseline Characteristics

Table 14-2.01 summarizes the height, weight, and body mass index (BMI) data at randomization for all ITT subjects in the study. The height was similar across all treatment groups and ranged from 135.9 to 195.6 cm, with a mean of 163.9 cm. The weight at baseline differed across treatment groups (p-value = 0.003), with placebo subjects having a mean weight of 62.8 kg (SD 12.77; range 34.0-86.2), xanomeline low dose subjects having a mean weight of 67.3 kg (SD 14.12; range 45.4-106.1), and xanomeline high dose subjects having a mean weight of 70.0 kg (SD 14.65; range 41.7-108.0).

Table 14-2.01 also summarizes the duration of subjects' disease and the Mini-Mental State Examination (MMSE) score at screening.

The duration of subjects' disease was similar across all treatment groups and ranged from 2.2 to 183.1 months with a mean duration of 43.9 months. The percentage of subjects with disease duration of < one year was 5%, and the percentage of subjects with disease duration of ≥ one year was 95%.

The MMSE score at screening was similar across all treatment groups and ranged from 10 to 24, with a mean score of 18.1.

11.3. Measurements of Treatment Compliance

Due to the nature of this CDISC Pilot Project, this section is not included in this study report.

11.4. Efficacy Results

There are 2 co-primary efficacy endpoints in this study: the change from baseline in ADAS-Cog (11) at Week 24 and the CIBIC+ score at Week 24. Hypothesis testing for both of these endpoints considered whether there is a statistically significant dose response among the 3 treatment groups. The significance level for each of these endpoints is 0.05.

11.4.1. ADAS-Cog (11)

Table 14-3.01 presents the summary statistics for baseline and Week 24 values for ADAS-Cog (11) and the change from baseline in ADAS-Cog (11) at Week 24. These results are based on the efficacy population with LOCF imputation. The mean scores across all 3 treatment groups are similar at baseline with the similar worsening in scores by Week 24. Accordingly, the dose response for the change from baseline at Week 24, adjusted for site and baseline score is not statistically significantly different from 0. Pairwise comparisons among the treatment groups are also presented in Table 14-3.01, but should not be considered since the dose response analysis was not statistically significant.

Similar analyses were performed at Weeks 8 and 16 (Table 14-3.03 and Table 14-3.05). Likewise, no dose response was seen at these earlier time points.

To check the robustness of the imputation method, a sensitivity analysis using only the observed values was performed (Table 14-3.07). Although the xanomeline low dose group had a smaller mean change from baseline at Week 24, the dose response in this analysis was also not statistically significantly different from 0.

Subgroup analyses by gender were also performed using the LOCF-imputed values (Table 14-3.08 and Table 14-3.09). The males in the xanomeline low dose group had a mean change from baseline that was higher than the other 2 groups, but the dose response was not statistically significantly. The dose response in the female subgroup was trending in a positive direction ($p = 0.094$) indicating that higher doses may result in smaller changes by Week 24. Mean (SD) changes from baseline at Week 24 for the placebo, xanomeline low dose, and xanomeline high dose were 3.0 (5.57), 1.7 (5.54), and 1.1 (4.77), respectively.

Summary statistics for the actual values and the change from baseline in ADAS-Cog (11) are presented in Table 14-3.10 for both the observed values by visit window and LOCF-imputed values.

A repeated measures analysis was performed on ADAS-Cog (11) using postbaseline changes from baseline (Table 14-3.11) to examine treatment effect over time. Covariates included in the model were treatment, site, time, treatment by time interaction, baseline

score, and baseline by time interaction. Adjusted means for the change from baseline at Week 24 were similar across the 3 treatment groups. Pairwise comparisons among the 3 treatment groups were not statistically significantly different.

11.4.2. CIBIC+

Table 14-3.02 presents the summary statistics for the Week 24 values for CIBIC+. These results are based on the efficacy population with LOCF imputation. The mean scores across all 3 treatment groups are similar at Week 24. Accordingly, the dose response for the CIBIC+ score at Week 24, adjusted for site is not statistically significantly different from 0. Pairwise comparisons among the treatment groups are also presented in Table 14-3.02, but should not be considered since the dose response analysis was not statistically significant.

Similar analyses were performed at Weeks 8 and 16 (Table 14-3.04 and Table 14-3.06). Likewise, no dose response was seen at these earlier time points.

An analysis not specified in the protocol was performed considering CIBIC+ scores as a categorical variable (Table 14-3.13). This analysis used the LOCF-imputed values and compared treatment groups at Weeks 8, 16, and 24. Distributions of scores between the 3 treatment groups were not statistically significantly different.

11.4.3. NPI-X

Table 14-3.12 presents the summary statistics for the baseline NPI-X score and for the mean of the observed values from Weeks 4 to 24, inclusive. The mean NPI-X value at baseline for the xanomeline high dose group was slightly larger than the other 2 groups. The dose response for the mean of the Week 4 through Week 24 values, adjusted for site and baseline value, was not statistically significantly different from 0. Pairwise comparisons among the treatment groups are also presented in Table 14-3.12, but should not be considered since the dose response analysis was not statistically significant.

11.5. Efficacy Conclusions

A statistically significant dose response was not seen for either of the primary efficacy endpoints: changes from baseline in ADAS-Cog (11) at Week 24 and CIBIC+ at Week 24. Adjusted means for these 2 endpoints were similar for all 3 treatment groups. Additional analyses at earlier time points showed similar results. Subgroup analyses by gender, a sensitivity analysis for missing data, and a repeated measures analysis for ADAS-Cog (11) also indicated lack of treatment response. The secondary efficacy endpoint of the mean NPI-X values from Week 4 through Week 24 also did not demonstrate a statistically significant dose response.

12. SAFETY EVALUATION

12.1. Extent of Exposure

For the purposes of this submission, planned exposure was summarized. It was assumed that while the subject was in the study the subject took the randomized drug as planned, with the first dose occurring on the date of randomization (Week 0), and the new dose levels occurring on the day following the Week 2 visit and the Week 24 visit. The date of last dose was that indicated on the CRF. If no date of last dose was available, the date of discontinuation was assumed to be the date of last dose.

A total of 254 subjects received randomized drug during the study. Eighty-six (86) subjects received placebo, 84 received xanomeline low dose and 84 received xanomeline high dose. The mean daily dose was 54.0 mg and 71.6 mg for the low dose and high dose treatment groups, respectively, as shown in Table 14-4.01.

12.2. Adverse Events

12.2.1. Brief Summary of Adverse Events

A summary of total adverse events across all body systems showed an increase in adverse events associated with randomized drug with over 90% of subjects receiving active therapy reporting at least one adverse event compared to 75.6% of subjects receiving placebo (Table 14-5.01). However, this difference is due largely to a disproportionate number of dermatologic type events that occurred in the xanomeline treatment groups. Therefore, with the exception of dermatologic irritation (discussed below), the overall adverse event profile does not suggest that there is a specific hazard associated with either dose of xanomeline.

The number of serious adverse events reported during this study were minimal (Table 14-5.02) and do not suggest that there is any pattern attributable to active therapy. There were 3 deaths observed during the course of the study, yet none of these were flagged as being serious events.

12.2.2. Display of Adverse Events

The most commonly reported adverse events, those reported in ≥ 5% of subjects in any treatment group are summarized in Table 12-1 in the order that they appear in Table 14-5.01.

Table 12-1. Most Common AE's (≥ 5% Subjects in any Treatment Group)

	Placebo (N=86) n (%)	Xan Low Dose (N=84) n (%)	Xan High Dose (N=84) n (%)
Sinus Bradycardia	2 (2.3%)	7 (8.3%)*	8 (9.5%)*
Vomiting	3 (3.5%)	3 (3.6%)	7 (8.3%)
Nausea	3 (3.5%)	3 (3.6%)	6 (7.1%)
Diarrhoea	9 (10.5%)	4 (4.8%)	4 (4.8%)
Application Site Pruritus	6 (7.0%)	22 (26.2%)*	22 (26.2%) *
Application Site Erythema	3 (3.5%)	12 (14.3%)*	15 (17.9%)*
Application Site Irritation	3 (3.5%)	9 (10.7%)*	9 (10.7%)*
Application Site Dermatitis	5 (5.8%)	9 (10.7%)	7 (8.3%)
Application Site Vesicles	1 (1.2%)	4 (4.8%)	6 (7.1%)*
Fatigue	1 (1.2%)	5 (6.0%)*	5 (6.0%)*
Nasopharyngitis	2 (2.3%)	4 (4.8%)	6 (7.1%)
Upper Respiratory Tract Infection	6 (7.0%)	1 (1.2%)*	3 (3.6%)
Dizziness	2 (2.3%)	8 (9.5%)*	11 (13.1%)*
Headache	3 (3.5%)	3 (3.6%)	5 (6.0%)
Cough	1 (1.2%)	5 (6.0%)*	5 (6.0%)*
Pruritus	8 (9.3%)	21 (25.0%)*	26 (31.0%)*
Erythema	8 (9.3%)	14 (16.7%)	14 (16.7%)
Rash	5 (5.8%)	13 (15.5%)*	9 (10.7%)
Hyperhidrosis	2 (2.3%)	4 (4.8%)	8 (9.5%)*
Skin Irritation	3 (3.5%)	6 (7.1%)	5 (6.0%)
Blister	0	5 (6.0%) *	1 (1.2%)

Source: Table 14-5.01

* $p < 0.150$ versus placebo

12.2.3. Analysis of Adverse Events

The adverse event profile of the treatment groups was generally similar across the treatment groups. With the exception of skin related adverse events, discussed below, the most commonly reported events (reported by ≥ 5% of subjects in any treatment group) were reported in numerous body systems and did not show a consistent treatment dependent pattern. There were noted differences ($p < 0.150$) between placebo and at least one of the active therapy groups in the incidence of sinus bradycardia (2.3% placebo, 8.3% low dose, and 9.5% high dose), fatigue (1.2% placebo, 6.0% xanomeline low dose, 6.0% xanomeline high dose), dizziness (2.3% placebo, 9.5% xanomeline low dose, 13.1% xanomeline high dose), cough (1.2% placebo, 6.0% xanomeline low dose, 6.0% xanomeline high dose), and hyperhiridosis (2.3% placebo, 4.8% xanomeline low dose, 9.5% xanomeline high dose). Upper respiratory tract infection was reported more frequently in the placebo group (7.0%) compared with the xanomeline low dose group (1.2%, $p < 0.150$) but not the xanomeline high dose group (3.6%).

12.3. Analysis of Death, Other Serious Adverse Events, and Other Significant Adverse Events

12.3.1. Deaths

There were 3 deaths reported during the conduct of this study with 2 subjects in the placebo group and 1 subject randomized to xanomeline low dose. None of these deaths were recorded as serious adverse events yet two of the deaths were clearly of serious nature (myocardial infarction and sudden death) and the other death being a suicide. None of the deaths were judged related to treatment.

12.3.2. Other Serious Adverse Events

There were a total of 3 serious adverse events reported in 3 individual subjects, 2 of whom were in the xanomeline high dose and 1 in the xanomeline low dose group (Table 14-5.02). All 3 events were events related to the nervous system. Due to the low numbers of serious adverse events, it is not possible to make any conclusions regarding the relationship of these events and treatment group.

12.3.3. Dermatologic Adverse Events

A special category of dermatologic events was created prior to unblinding. This category combines all adverse events that were considered to be of dermatologic importance after a thorough medical review of the coded adverse events. The preferred terms listed in Table 12-2 were included in this category of special interest adverse events.

Table 12-2. Preferred Terms Included in Special Interest Category of Dermatologic Events

System Organ Class	Preferred Term
General Disorders and Administration Site Conditions	Application Site Pruritus
	Application Site Erythema
	Application Site Irritation
	Application Site Dermatitis
	Application Site Vescicles
	Application Site Pain
	Application Site Perspiration
	Application Site Swelling
	Application Site Discharge
	Application Site Reaction
	Application Site Urticaria
	Application Site Bleeding
	Application Site Desquamation
	Application Site Discolouration
	Application Site Induration
	Application Site Warmth
Skin and Subcutaneous Tissue Disorders	Pruritus
	Erythema
	Rash
	Skin Irritation
	Rash Pruritic
	Actinic Keratosis
	Blister
	Pruritus Generalised
	Rash Maculo-papular
	Skin Odour Abnormal
	Urticaria
	Dermatitis Cintact
	Drug Eruption
	Rash Erythematous
	Skin Exfoliation
	Skin Ulcer

As shown on the table at the bottom of Figure 14-1, there was a disproportionate number of subjects in the xanomeline treatment groups who experienced a special interest dermatologic adverse event with 74% (n = 62) of xanomeline low dose and 73% (n = 61) of xanomeline high dose subjects with at least one event of special interest compared to 34% (n = 29) of placebo subjects. This high rate of dermatologic events most likely contributed to the larger proportion of subjects in the active treatment groups who discontinued the study due to an adverse event (Table 14-1.02).

An analysis of the time to the first dermatologic event indicated that the median time to first event was significantly different ($p<0.0001$) between treatment groups with a median time of 33 days in the xanomeline low dose group (95% CI: 27 – 48 days) and 36 days in the xanomeline high dose group (95% CI: 24 – 46 days) compared to placebo, in which the median time to first event was not estimable (Figure 14-1).

12.3.4. Narratives of Deaths, Other Serious Adverse Events and Certain Other Significant Adverse Events

Narratives for subjects who died while on-study are included below. Narratives for subjects who reported other serious adverse events or other significant adverse events are not included in this study report due to the nature of this CDISC Pilot Project.

Subject 01-701-1211, a 76-year-old woman with Alzheimer's dementia, began receiving xanomeline low dose on 15 November 2012. The subject was diagnosed with Alzheimer's dementia in 2010. The subject also had non-insulin dependent diabetes mellitus, which was considered to be mild in severity. The subject experienced sudden death on 14 January 2013. This death was considered by the investigator to be unrelated to study medication. No action was taken for this event and study medication was not discontinued prior to death. Eight other adverse events were reported for this subject yet none were classified as being serious and none appeared to contribute to the sudden death.

Subject 01-701-1445, a 75-year-old man with Alzheimer's dementia, began receiving placebo on 11 May 2014. The subject was diagnosed with Alzheimer's dementia in 2012. The subject also suffered from mild forms of non-insulin dependent diabetes mellitus, chronic obstructive pulmonary disease, and benign prostatic hyperplasia. This subject committed suicide on 31 October 2014. This death was considered by the investigator to be unrelated to study medication. No action was taken for this event and study medication was not discontinued prior to death. No other adverse events were reported by this subject.

Subject 01-701-1083, an 89-year-old woman with Alzheimer's dementia, began receiving placebo on 22 July 2013. The subject was diagnosed with Alzheimer's dementia in 2011. This subject had previous history of cardiovascular disease: implanted cardiac pacemaker, first degree atrioventricular block, and a history of myocardial infarction and

angina pectoris in 2006. The subject experienced a fatal myocardial infarction on 2 August 2013. This event required hospitalization. The event was considered by the investigator to be possibly related to study medication. No action was taken for this event and study medication was not discontinued prior to death. No other adverse events were reported by this subject.

12.4. Clinical Laboratory Evaluation

Table 14-6.01 to Table 14-6.05 summarize the findings for the laboratory data analysis. Commonly used laboratory analytes (as defined by the CDISC coding document) that were measured with sufficient frequency were reported. Hematology analytes that meet this criterion are: hemoglobin, hematocrit, mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), leukocyte count, lymphocytes, monocytes, eosinophils, basophils, platelet count, and erythrocyte count. Clinical chemistry analytes that meet this criterion are: sodium, potassium, chloride, total bilirubin, alkaline phosphatase, gamma-glutamyl transferase (GGT, GGPT, SGGT, YGGT), alanine aminotransferase (ALT, SGPT), aspartate aminotransferase (AST, SGOT), urea nitrogen, creatinine, uric acid, inorganic phosphorus, calcium, non-fasting glucose, total protein, albumin, cholesterol, and creatine phosphokinase. Urinalysis and other lab data were not summarized, but were included in the tabulation datasets.

The hematology and clinical chemistry measures were summarized for baseline and Weeks 2, 4, 6, 8, 12, 16, 20, 24, and 26 (visits 1, 4, 5, 7, 8, 9, 10, 11, 12, and 13, respectively). The baseline values were those collected at Week -2 (visit 1). Table 14-6.01 presents summaries of these results by week. This table provides descriptive statistics (n, mean, and standard deviation) for the measured value in standard units as well as the change from baseline. The change from baseline laboratory value was calculated as the difference between the baseline lab value and the endpoint value (i.e., the value at the specified visit) or the end of treatment observation. These results are shown by treatment group.

Table 14-6.02 and Table 14-6.03 summarize each laboratory value, including the baseline value, as categorized with reference to the lab normal range as

- "L" - less than or equal to the lower limit of normal (LLN)
- "N" – Greater than the LLN and less than the upper limit of normal (ULN)
- "H" – Greater than or equal to the ULN

Laboratory values were assigned a flag of abnormal (high or low) if the value was outside the threshold range (defined as significantly beyond the normal range, i.e., > 1.5 times ULN or < 0.5 times LLN) or if the value was significantly different from the value observed at the preceding scheduled visit (i.e., absolute value of the change from

previous value is larger than the 50% of the normal range, LLN to ULN). Shift tables summarizing whether a subject's status changed from baseline at each week of treatment and during the treatment period are provided based on threshold ranges and changes based on Hy's Law. Two variations of the modified Hy's Law criteria were used in the assessment. The first considered subjects with transaminase (ALT or AST) elevations of greater than 1.5 times ULN as abnormal. The second further narrowed the assessment of abnormality to require total bilirubin elevations to be greater than 1.5 times ULN in addition to transaminase elevations of greater than 1.5 times ULN. In these tables subjects were categorized as normal or abnormal (i.e., outside the threshold range) at baseline. During the treatment phase, the most extreme value was used to categorize a subject as normal or abnormal during the treatment phase. The shift table shows the number of subjects whose on treatment categorization was the same or shifted from the baseline categorization. The treatment phase is defined as any planned visit after Week 0 (Visit 3), up to and including Week 24 (Visit 12).

The number of subjects with no abnormal measure during treatment and those with at least one abnormal measure during treatment are summarized for each lab analyte. Table 14-6.02 provides an analysis with abnormal as beyond normal range (i.e., below LLN or above ULN). Table 14-6.03 provides the analysis based on abnormal as a clinically significant change from the previous visit.

As shown in Table 14-6.02, only three laboratory analytes were statistically significantly associated with treatment group: albumin ($p = 0.042$), urea nitrogen ($p = 0.023$), and eosinophils ($p = 0.001$). There were four additional analytes that were nearly significant: chloride ($p = 0.058$), hematocrit ($p = 0.052$), hemoglobin ($p = 0.093$), and MCV ($p = 0.077$). Albumin was more often lower than normal range for subjects in the placebo and xanomeline low dose groups. Subjects in the xanomeline treatment groups had statistically significantly more values above the normal range than subjects in the placebo group for urea nitrogen [placebo=9 (11%), xanomeline low dose group =22 (27%), xanomeline high dose group=12 (15%)] and eosinophils [placebo=0, xanomeline low dose group =11 (13%), xanomeline high dose group=7 (9%)].

Table 14-6.03 shows the number of clinically significant changes from previous visit by treatment group for each analyte. The association between clinically significant change from the previous visit and treatment was statistically significant for aspartate aminotransferase ($p = 0.045$) and eosinophils ($p = 0.010$). Nearly significant were protein ($p = 0.062$) and monocytes ($p = 0.081$).

Table 14-6.04 summarizes shifts from baseline by week in terms of abnormality based on threshold range. The data in this table were summarized using sets of 3-by-3 tables comparing baseline and on drug categorization for each treatment group for each week for each laboratory analyte. Because no subjects were abnormally low at baseline, only the normal at baseline and high at baseline data are shown.

In Table 14-6.05, a CMH test, stratifying by status at baseline, is shown. This test was performed to assess the association of significant shifts from baseline for each analyte with treatment group. The number of subjects with no abnormal measure during treatment and those with at least one abnormal measure during treatment was summarized for each lab analyte. Nineteen of the 30 analytes could not be analyzed because they had less than two non-missing levels. Of the 11 remaining analytes, only shifts from baseline to treatment phase for eosinophils were statistically significantly related to treatment group ($p = 0.044$). For eosinophils, the number of subjects in each group that showed a shift from normal to high were 0 in the placebo group, 5 (6%) in the xanomeline low dose group, and 6 (8%) in the xanomeline high dose group.

Finally, Table 14-6.06 reports the results of the Hy's Law analysis. Of the cases where transaminase was greater than 1.5 times ULN, 3 subjects in the placebo group, 1 in the xanomeline low dose, and 3 in the xanomeline high dose shifted from normal at baseline to above normal during treatment. In addition, 2 subjects in the placebo group and 1 in the xanomeline low dose group had an elevated transaminase at baseline that remained high during treatment, and 1 subject in the xanomeline low dose group had an elevated transaminase at baseline that became normal during treatment. There were no subjects in the xanomeline treatment groups and 1 subject in the placebo group who had both transaminase levels greater than 1.5 times ULN and total bilirubin greater than 1.5 times ULN. A CMH test for an association with treatment group was not statistically significant for either Hy's law assessment.

12.5. Vital Signs and Weight

Vital signs and weight were to be collected for all subjects at each visit. For the purposes of this submission, vital signs were summarized only for baseline and Week 24 and end of treatment. End of treatment is defined to be the last on-treatment assessment of the specified measure.

12.5.1. Vital Signs Summary

Table 14-7.01 presents a summary of the vital sign data (systolic blood pressure [SBP], diastolic blood pressure [DBP] and pulse) collected at baseline (Week 0), Week 24, and end of treatment.

Table 14-7.02 summarizes the changes in vital signs at Week 24 and end of treatment as compared to values at the baseline visit (Week 0).

Table 12-3 presents summary statistics for SBP, DBP, and pulse after standing for 3 minutes. These assessments are representative of the vitals taken at other positions and time points in terms of the relationships among treatment groups. Measurements were also taken after the subject had been lying down for 5 minutes and after the subject had

been standing for 1 minute. The largest change from baseline occurred in the assessments taken after the subject had been standing for 3 minutes.

Vital sign values at baseline were comparable across treatment groups. Changes from baseline, at the Week 24 and end of treatment assessments, were generally small decreases.

Table 12-3. Summary of Change from Baseline in Vital Signs

	Treatment Group					
	Placebo (N=86)		Xan Low (N=84)		Xan high (N=84)	
Parameter	n	Mean	n	Mean	n	Mean
SBP (mmHg) after standing for 3 minutes						
Baseline	85	136.5	84	136.4	84	138.8
Change at Week 24	58	-1.0	27	-0.1	30	-9.0
Change at End of Treatment	83	-2.5	83	-3.5	81	-8.3
DBP (mmHg) after standing for 3 minutes						
Baseline	85	77.7	84	76.6	84	79.6
Change at Week 24	58	-2.3	27	-1.6	30	-2.1
Change at End of Treatment	83	-2.7	83	-1.8	81	-2.6
Pulse (BPM) after standing for 3 minutes						
Baseline	85	74.6	84	72.3	84	74.0
Change at Week 24	58	-1.5	27	-2.1	30	-2.7
Change at End of Treatment	83	-1.0	83	-0.7	81	-1.9

Source: Table 14-7.01 and Table 14-7.02

12.5.2. Weight Summary

Table 14-7.03 presents a summary of the weight data collected at baseline (Week 0), Week 24, and end of treatment. It also summarizes the changes in weight at Week 24 and end of treatment as compared to values at the baseline visit (Week 0).

As mentioned in Section 11.2.2, weight at baseline differed across treatment groups. Changes from baseline, at the Week 24 and end of treatment assessments, however, were generally small with no treatment-related pattern of increases or decreases (Table 12-4).

Table 12-4. Summary of Change from Baseline in Weight

	Treatment Group					
	Placebo (N=86)		Xan Low (N=84)		Xan high (N=84)	
Parameter	N	Mean	n	Mean	n	Mean
Weight (kg)						
Baseline	86	62.8	83	67.3	84	70.0
Change at Week 24	59	0.1	27	-0.3	30	1.0
Change at End of Treatment	84	0.2	83	-0.4	81	0.1

Source: Table 14-7.03

12.6. Concomitant Medications

Table 14-7.04 shows that use of concomitant medications was similar across the 3 treatment groups. Concomitant medications were taken by 77 (90%), 74 (88%), and 78 (93%) subjects in the placebo, xanomeline low dose, and xanomeline high dose groups, respectively.

The most common concomitant medication (based on coded term) used by xanomeline subjects in this study was hydrocortisone, which was taken by 2 (2%), 13 (15%), and 8 (10%) subjects in the placebo, xanomeline low dose, and xanomeline high dose groups, respectively.

12.7. Safety Conclusions

Over 90% of subjects receiving active therapy reported at least 1 adverse event compared to 75.6% of subjects receiving placebo. This difference is due largely to a disproportionate number of dermatologic type events that occurred in the xanomeline treatment groups. Approximately 73% of the subjects in either of the xanomeline groups experienced at least one dermatologic adverse event of interest compared to 33.6% of the placebo subjects. There was a statistically significant difference ($p < 0.001$) in the time to first dermatologic event between the treatment groups. There were 3 deaths (2 in placebo group, 1 in the xanomeline low dose group) observed during the study. None of the deaths were judged related to treatment. Aside from the deaths, there were 3 serious adverse events reported in 3 subjects (2 in xanomeline high dose and 1 in the xanomeline low dose group) and all were related to the nervous system.

The association between treatment group and the number of abnormal values beyond the normal range was significant for three laboratory analytes: albumin ($p = 0.042$), urea nitrogen ($p = 0.023$), and eosinophils ($p = 0.001$). The association between clinically significant changes from the previous visit and treatment was statistically significant for aspartate aminotransferase ($p = 0.045$) and eosinophils ($p = 0.010$). The analysis of shifts from baseline to most abnormal value could not be calculated on 19 of the analytes. Of the remaining 11 analytes, only eosinophils showed a statistically significant association with treatment group ($p = 0.044$). There was no significant association with treatment group in the Hy's law analyses examining shifts in transaminase levels, and transaminase and total bilirubin levels between baseline values and values while on treatment.

Changes from baseline in vital signs (SBP, DBP, and pulse), at the Week 24 and end of treatment assessments, were generally small decreases. Changes from baseline in weight, at the Week 24 and end of treatment assessments, however, were generally small with no treatment-related pattern of increases or decreases.

13. DISCUSSION OF STUDY RESULTS AND CONCLUSIONS

A statistically significantly higher proportion of subjects in the active treatment groups withdrew prematurely from the study as compared to the placebo group. This is largely due to the higher proportion of subjects in the active treatment groups experiencing a dermatologic event and subsequently resulting in premature withdrawal from the study. This further hindered the study's ability to demonstrate efficacy.

A statistically significant dose response was not seen for both of the primary efficacy endpoints, change from baseline in ADAS-Cog (11) at Week 24 and CIBIC+ at Week 24, and for the secondary efficacy endpoint, mean NPI-X values from Week 4 to Week 24. Adjusted means for all 3 endpoints were similar across all treatment groups.

There were an increased number of dermatologic adverse events reported in the active treatment groups as compared to the placebo group. There were 3 serious adverse events. In addition, there were 3 deaths that were deemed unrelated to treatment.

For the laboratory data, subjects in both the xanomeline low and high dose groups showed more observations above normal range than the placebo group. Albumin was more often lower than the normal range for subjects in the placebo and xanomeline low dose group. Subjects in the xanomeline treatment groups had statistically significantly more values above the normal range than subjects in the placebo group for both urea nitrogen and eosinophils. There was a statistically significant association between clinically significant changes from the previous visit and treatment group for aspartate aminotransferase and eosinophils. Shifts from baseline for eosinophils were statistically significant with both xanomeline treatment groups showing more changes from normal to above normal than the placebo group. There was no significant association with treatment group in the Hy's law analysis examining shifts in liver function tests between baseline values and values while on treatment.

There were only minor changes from baseline in vital signs and weight at Week 24.

14. SUMMARY TABLES AND FIGURES