CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

021825Orig1s000

STATISTICAL REVIEW(S)

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

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/Serial Number:	21-825/ SE 0056
Drug Name:	Deferiprone (Ferriprox)
Indication(s):	Treatment of iron overload
Applicant:	ApoPharma Inc
Date(s):	Letter Date: April 14 2011
	Stamp Date: April 14 2011
	PDUFA Goal Date:
Review Priority:	Standard
Biometrics Division:	Division of Biometrics V
Statistical Reviewer:	Qing Xu, Ph.D
Concurring Reviewers:	Mark Rothmann, Ph.D., Statistical Team Leader
	Rajeshwari Sridhara, Ph.D., Director, DBV
Medical Division:	Division of Hematology Products
Clinical Team:	George Shashaty, M.D., Clinical Reviewer
	Kathy M Robie Suh, M.D., Clinical Team Leader
Project Manager:	Mara Bauman Miller

Keywords:

Bias, Confidence interval, Covariate, Logistic regression, Meta-analysis, Missing data, Pooling

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1. EXECUTIVE SUMMARY

Deferiprone (Ferriprox) is an orally administered iron chelator that is being developed for the indication of the treatment of patients with transfusional iron overload when current chelation therapy is inadequate. NDA 21825 for Ferriprox (deferiprone) was submitted on January 29, 2009 and a Complete Response letter was sent to the sponsor on November 30, 2009. This resubmission of NDA 21825/SE0056 including study LA36-0310 was designed as an analysis of existing data from studies previously conducted to evaluate the efficacy of Ferriprox. No new data were collected and the original purpose of collecting the data and their application did not change. Efficacy data for study LA 36-0310 were derived from 12 of 17 studies. The primary efficacy endpoint was the change in serum ferritin concentration from baseline within one year of Ferriprox therapy. Ferriprox therapy was considered successful in individual patients who experienced a ≥20% decline in serum ferritin concentration within one year of therapy.

The sponsor's efficacy analysis for serum ferritin by pooling 12 studies showed that the overall success rate was 52% with 95% CI of (45%, 58%). As the lower limit of the 95% CI is larger than 20%, the protocol defined endpoint was met for this trial. However, this study has several serious limitations including lack of randomization, lack of control group, high rate of missing data and ignoring the variation between studies by simple pooling, all of which can introduce biases to the primary outcome. Therefore, it is unclear whether the efficacy shown in the study is solely due to the Ferriprox therapy, and the interpretation of these analysis results should be taken cautiously.

The Oncology Drug Advisory Committee (ODAC) Meeting discussed NDA 21825/SE0056 study results on September 14, 2011. For the question:

1. Is there a favorable benefit/risk profile for deferiprone in the treatment of patients in who current chelation therapy is inadequate?
- Committee voted: No 2, Yes 10.

The results for the AC member's questions are given in the Appendix.

2. INTRODUCTION

2.1 Overview

The original NDA 21825 was submitted to the Agency on January 29, 2009. The submission of the NDA included a single randomized controlled trial. The primary endpoint in that trial was the change in cardiac iron as measured by cardiac MRI T2* assessment after one year of treatment with Ferriprox. The comparator drug was deferoxamine, which at the time of the study was the only approved drug for the indication. The agency was concerned that the primary endpoint

measured was not a validated surrogate for clinical utility. At the conclusion of the review, the Agency sent a Complete Response (CR) letter to the sponsor describing the deficiencies of the data submitted. In the CR letter dated November 30, 2009, the Agency recommended that the sponsor perform "adequate and well-controlled" trials with Ferriprox to support the application for approval of the NDA. Subsequently, the FDA communicated to ApoPharma that it would consider accelerated approval for the Ferriprox NDA with a single arm study in patients intolerant of or not responding to existing therapy, based on pre-existing data from Ferriprox clinical program.

This resubmission includes clinical study report of LA36-0310. Study LA36-0310 was designed as an analysis of existing data from studies previously conducted to evaluate the efficacy of Ferriprox. Efficacy data in LA36-0310 were derived from 12 of 17 studies. Four additional studies submitted to the NDA, LA109902, LA14-9907, LA20-BA, and LA21-BE, were not included in the current analysis as they did not assess efficacy. Another submitted study LA17-9701, was not included as the sponsor did not receive a complete database for the program.

The determination of which patients fulfilled the study criteria was conducted by an Independent Committee that reviewed the relevant data from the patients enrolled in studies previously submitted to the FDA to determine eligibility. Eligible patients met the following criteria:

1. Patients were treated with Ferriprox;
2. At least a single baseline value for serum ferritin or LIC or cardiac MRI T2*was available; and
3. Follow-up assessment of serum ferritin or LIC or cardiac MRI T2* was carried out after initiation of Ferriprox therapy and within one year therapy.

2.2 Data Sources

The applicant submitted this NDA including the data to the FDA CDER Electronic Document Room (EDR).The analysis dataset was not adequate and required an information request. The clinical study reports and datasets are located at the following location:
\\CDSESUB1\EVSPROD\NDA021825\021825.ENX

3. STATISTICAL EVALUATION

3.1 Data and Analysis Quality

Study LA36-0310 analyzed data from clinical studies that were conducted to support NDA 21-825. Patients were selected from studies submitted to the Agency as part of the NDA, and data integrated up to 11 May 2010 for ongoing clinical studies. An integrated lab dataset including all

serum ferritin, LIC and cardiac MRI T2* data, and data on demographics, disposition, medical history, and exposure, were sent to an Independent Committee responsible for selecting patients for analysis. The integrated datasets were prepared according to standard CDISC principles and were accompanied by a proper description of each field (i.e., metatables). Patient selection was based on inclusion/exclusion criteria previously established and agreed upon with the Agency. The Biostatistics group at ApoPharma subsequently assessed the serum ferritin, LIC and cardiac MRI T2* data captured during treatment with Ferriprox for up to one year for analysis of its efficacy in the cohort of patients selected by the independent committee. The number of Ferriprox-treated patients who met the defined criteria for successful treatment outcomes was determined and a success rate was calculated. The complete patient dataset, the dataset indicating the patient selection, the corresponding analysis dataset and the SAS program for determination of responders is presented in NDA 21-825 resubmission.

With the updates from the sponsor upon the statistical reviewer's information request, the reviewer was able to perform all analyses using the submitted data. No additional data submission was needed.

3.2 Evaluation of Efficacy

3.2.1 Study Objective

The objective of this study was to evaluate the efficacy of oral administration of Ferriprox in the treatment of iron overload in patients in whom previous chelation has failed. Chelation failure was defined as iron accumulation above a boundary level, defined by high serum ferritin or LIC or low cardiac MRI T2*levels.

3.2.2 Study Design, Endpoints and Analysis Population

Overall Study Design

In order to evaluate Ferriprox as a second line treatment for transfusion iron overload, Study LA36-0310 was designed to analyze existing data from studies previous conducted to evaluate the efficacy of Ferrriprox. Efficacy data in LA36-0310 were derived from 12 of 17 studies. Four additional studies submitted to the NDA, LA109902, LA14-9907, LA20-BA, and LA21-BE, were not included in the current analysis as they did not assess efficacy. Another submitted study LA17-9701, was not included as the sponsor did not receive a complete database for the program. An integrated lab dataset including all serum ferritin, LIC and cardiac MRI T2* data, and data on demographics, disposition, medical history, and exposure, were sent to an Independent Committee responsible for selecting patients for analysis.

Primary Endpoint: Change in serum ferritin concentration from baseline within one year of Ferriprox therapy (and up to 3 months following the anniversary date or the date of medication

termination). Ferriprox therapy was considered successful in individual patients who experienced a ≥20% decline in serum ferritin concentration within one year of therapy.

Secondary Endpoints: Change in cardiac MRI T2* and LIC from baseline within one year of Ferriprox therapy (and up to 3 months following the anniversary date or the date of medication termination). Ferriprox therapy was considered successful in individual patients who experienced a ≥20% increase in cardiac MRI T2* or a ≥20% decline in LIC within one year of therapy.

The Intent-To-Treat (ITT) Population: was the primary population for the efficacy analyses for this study. The ITT population was defined by sponsor as patients who had taken at least one dose of Ferriprox and had at least one post-baseline measurement of that efficacy measure. The ITT population included data from all randomized patients in the studies. For studies in which there was no patient randomization (Studies LA-02/06; LA-03; LA-04/06B; LA11; LA12 9907; LA15-0002; LA28-CMP; LA30-0307 and Borgna-Pignatti study), data from all patients who had received Ferriprox therapy were included.

The Per-Protocol (PP) population: comprised those patients who had completed their study of origin or at least one year of Ferriprox therapy for long-term studies, and had no missing data for the end-of-study measurement or the last scheduled measurement at the end of the first year, respectively, for that efficacy measure.

Reviewer's Comment:
During the review of the submissions, the Agency raised a concern of selection bias due to the lack of randomization. In order to minimize such bias, the Agency suggested sponsor using an Independent Committee to select subjects from the entire integrated dataset following approved inclusion and exclusion criteria. The agreement was reached between Agency and sponsor.

3.2.3 Patient Disposition, Demographic and Baseline Characteristics

A total of 746 patients with serum ferritin, LIC and MRI T2* data were analyzed by an Independent Committee for study eligibility. Of these , 264 were deemed eligible based on the serum ferritin criterion, 117 based on the LIC criterion and 39 based on the cardiac MRI T2* criterion. These data are shown in the Table 1, Table 2 and Table 3.

Table 1 Number of eligible patients by study for serum ferritin-ITT population

Study	Total N	N for eligible patients
LA-01	35	8 (23%)
LA-02/06	151	65 (43%)
LA-03	24	8 (33%)
LA-04/06B	157	56 (36%)
LA08-9701	25	7 (28%)
LA-11	23	12 (52%)
LA12-9907	69	19 (28%)
LA15-002	29	18 (62%)
LA16-0102	29	5 (17%)
LA28-CMP	8	3 (38%)
LA30-0307	100	36 (36%)
Borgna-Pignatti	96	27 (28%)
Total	**746**	**264 (35%)**

Data source: sponsor's clinical report table 6.1-1

Table 2 Number of eligible patients by study for liver iron concentration-ITT population

Study	Total N	N for eligible patients
LA-01	35	15 (43%)
LA-02/06	62	0 (0%)
LA-03	25	12 (48%)
LA-04/06B	100	11 (11%)
LA08-9701	29	21 (72%)
LA12-9907	75	35 (47%)
LA16-0102	28	20 (71%)
LA28-CMP	2	0 (0%)
Borgna-Pignatti	2	0 (0%)
Total	**358**	**114 (32%)**

Data source: sponsor's clinical report table 6.1-2

Table 3 Number of eligible patients by study for cardiac MRI T2* -ITT population

Study	Total N	N for eligible patients
LA-01	1	0 (0%)
LA-02/06	1	0 (0%)
LA-04/06B	72	10 (14%)
LA16-0102	29	29 (100%)
LA28-CMP	2	0 (0%)
Total	**105**	**39 (37%)**

Data source sponsor's clinical report table 6.1-3

Reviewer's comments:

This study is a retrospective study designed to analyze existing data from studies conducted previously to evaluate the efficacy of Ferrriprox. Efficacy data in LA36-0310 were derived from 12 of 17 studies; the study involves pooling data from 12 studies of different designs; different durations of drug treatment; different time intervals of patient visits; different inclusion and exclusion criteria; and different patient characteristics. This inevitably raises an issue of data integrity and heterogeneity that can potentially cause biases. The sponsor followed the suggestion from the Agency and used Independent Committee to pick the eligible subjects. However, such approach can only limit biases, but can not eliminate biases.

By the entry criteria all patients must have received some prior chelation therapy (deferoxamine or deferasirox). The reviewer found that among the 264 patients, there are 23 patients who have no record about whether they took the prior chelation therapy. However, excluding these 23 patients will not alter the conclusion for primary efficacy. These 23 patients are:
LA_01: 66, 67;

LA_12: 171, 31, 36;
LA_0206: 222, 264, 7, 9
LA_04: 2, 40, 99
LA_11: 102, 104,106,107,108,109, 112,113,118,121,122, 124

These population overlapped but were not superimposable for the three endpoints

During the review process, the reviewer found there were some discrepancies between the Agency's results and the sponsor's results regarding the number of subject for the different studies for serum ferritin based on dataset that the sponsor sent to the Independent Committee for the selection of eligible subject (Table 4). The sponsor clarified as following:

1. *With the exception of the 2 patients not included by the Agency in the total number of patients (744 vs 746), the discrepancies between the Agency's numbers of subjects and ApoPharma's numbers can be demonstrated to be a consequence of steps taken by the company to ensure that patients enrolled in more than one study were counted only once when the dataset was integrated*
2. *Data on all patients with serum ferritin values were sent by ApoPharma to the Independent Committee for consideration. The Independent Committee did not include 2 of the 746 patients (BP_283 and LA_12_26) in the la36cohort dataset returned to ApoPharma. The sponsor surmise that there 2 patients were not included because they had neither a baseline value nor a value within the 1-year cut-off period.*

Table 4 Summary of eligible petients between the Agency's results and the sponsor's results

Study ID	Total N		N for eligible patients	
	FDA	Sponsor	**FDA**	Sponsor
LA_01	**32**	35	**8**	8
LA_0206	**151**	151	**59**	65
LA_03	**22**	24	**7**	8
LA_04	**165**	157	**58**	56
LA_08	**28**	25	**7**	7
LA_11	**23**	23	**12**	12
LA_12	**61**	69	**22**	19
LA_15	**29**	29	**18**	18
LA_16	**29**	29	**5**	5
LA_28	**8 3**	8	**24**	3
LA_30	**25**	100	**15**	36
BP	**86**	96	**26**	27
LA_10	**10**		**3**	
Total	**744**	746	**264**	264

Table 5shows reviewer's summary of race for serum ferritin using ITT population.

Table 5 Reviewer's summary of race by study for serum ferritin (ITT population)

Study	ASIAN	BLACK	MULTI-RACIAL	UNKNOWN	WHITE	Total
BP	0 0.00	0 0.00	0 0.00	0 0.00	27 10.23	27 10.23
LA_01	5 1.89	0 0.00	0 0.00	0 0.00	3 1.14	8 3.03
LA_0206	1 0.38	0 0.00	0 0.00	0 0.00	64 24.24	65 24.62
LA_03	4 1.52	0 0.00	0 0.00	0 0.00	4 1.52	8 3.03
LA_04	6 2.27	2 0.76	1 0.38	21 7.95	26 9.85	56 21.21
LA_08	0 0.00	0 0.00	0 0.00	0 0.00	7 2.65	7 2.65
LA_11	12 4.55	0 0.00	0 0.00	0 0.00	0 0.00	12 4.55
LA_12	0 0.00	0 0.00	0 0.00	0 0.00	19 7.20	19 7.20
LA_15	0 0.00	0 0.00	0 0.00	0 0.00	18 6.82	18 6.82
LA_16	0 0.00	0 0.00	0 0.00	0 0.00	5 1.89	5 1.89
LA_28	3 1.14	0 0.00	0 0.00	0 0.00	0 0.00	3 1.14
LA_30	15 5.68	0 0.00	0 0.00	0 0.00	21 7.95	36 13.64
Total	46 17.42	2 0.76	1 0.38	21 7.95	194 73.48	264 100.00

Reviewer's Comments:
There were only 2 black patients in the study LA36-0310. Such patient proportion did not represent US patient population.

Table 6 shows Reviewer's summary of demographic for serum ferritin using ITT population

Table 6 Reviewer's summary of demographic for serum ferritin using ITT population

COUNTRY	Frequency	Percent
CANADA	23	8.71
EGYPT	21	7.95
GREECE	8	3.03
INDONESIA	12	4.55
IRAN	18	6.82
ITALY	128	48.48
MALAYSIA	4	1.52
SINGAPORE	2	0.76
THAILAND	12	4.55
USA	36	13.64

Reviewer's Comment
A majority of patients were from Italy (48.48%). A total of 36 (13.64%) subjects are from USA.

Table 7 below shows the reviewer's summary of demographic for serum ferritin, LIC, and MRI T2*, respectively.

Table 7 Reviewer's summary of demographic for serum ferritin, LIC, and MRI T2*

	Serum ferritin	LIC	MRI T2*
Age			
Mean ± SD	20.1±12.3	19.0±6.3	24.3 ±4.7
(Minimum, Maximum)	(2, 76)	(6, 40)	(12, 33)
Sex			
Female	145 (55%)	54 (47%)	18 (46%)
Male	119 (45)	60 (53%)	21 (54%)
Race			
White	194 (73%)	93 (82%)	31 (79%)
Asian	46 (17%)	18 (16)	5 (13%)
Black	2 (1%)	3 (3%0	3 (8%)
Multi-racial	1 (0.4%)		
Unknow	21 (8%)		

Reviewer's Comment:
For the primary efficacy endpoint of serum ferritin, the subjects ranged from 2 to 76 years in age, and mean age were 20.1 years. There were more female patients than male patients (55% vs 45%). The majority of patients were white in race (73%).

Table 8, Figure 1 and Figure 2 below show the reviewer's summary of number of subjects over time and drop out rate over time

Table 8 Reviewer's summary of number of subjects and drop out rate over time

MONTH	Number of Subjects	Drop Out Rate
0	264	0
3	244	0.075758
6	185	0.299242
9	170	0.356061
12	137	0.481061
15	122	0.537879
18	108	0.590909
21	96	0.636364
24	96	0.636364
27	86	0.674242
30	69	0.738636

Figure 1 Reviewer's summary of number of subjects over time

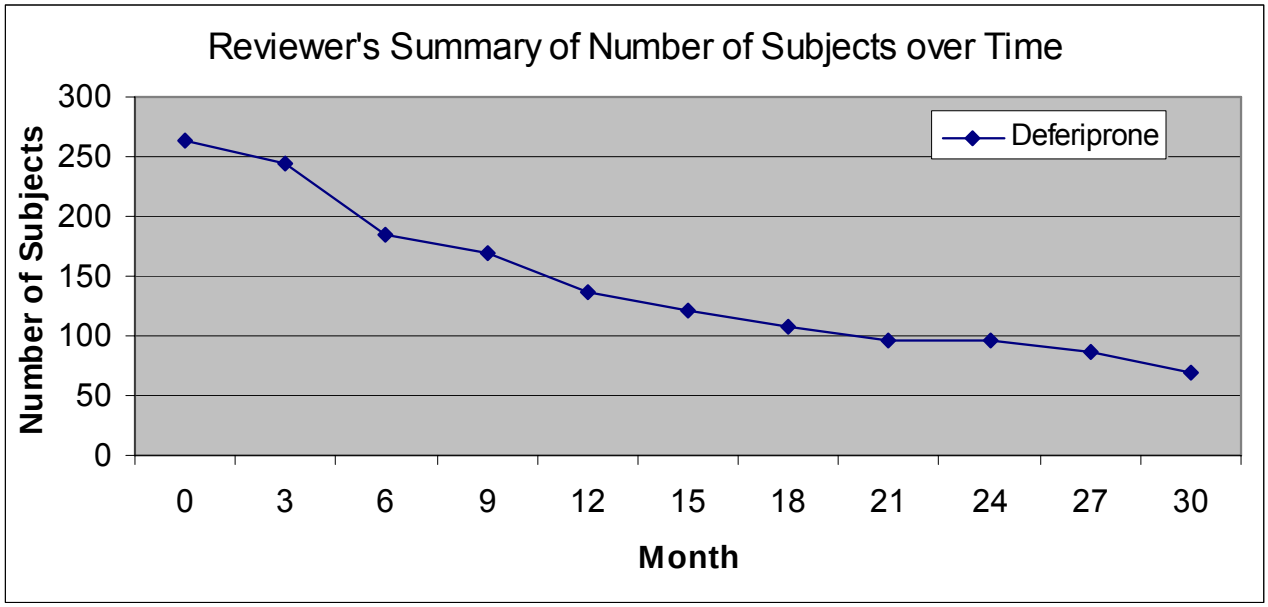
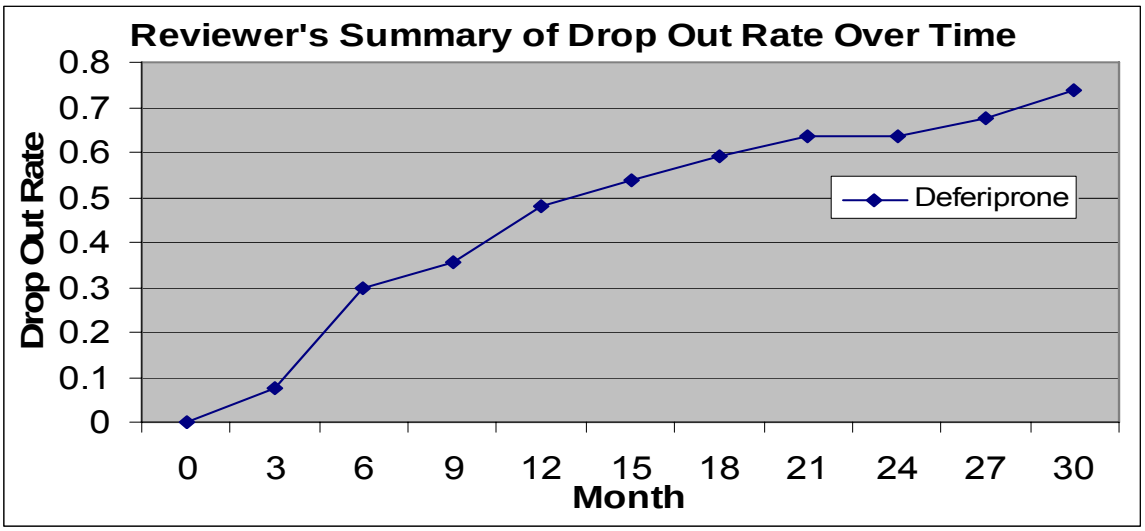


Figure 2 Reviewer's summary of drop out rate over time



Reviewer's Comments:

A total of 20 (7.58%), 79 (29.92%), 94 (35.61%) and 127 (48.11%) subjects dropped out by month 3, 6, 9 and 12, respectively. High drop out rate could make the analyses results unreliable and misleading. The reviewer conducted primary efficacy analysis using worst-case-scenario missing data imputation by treating all missing values as non-response. The result is given in the section of 3.5.2.1.

3.2.4 Statistical Methodologies

All statistical tests were two-sided, and a p-value ≤ 0.05 was used for the determination of statistical significance.

Analysis of primary efficacy endpoint –Serum Ferritin

For the primary efficacy endpoint, the success rate was calculated as the proportion of patients with a reduction in serum ferrin by ≥ 20% (serum ferritin baseline value >2, 500 ug/L) within one year of Ferriprox treatment. The success rate by study and the overall success rate and its 95% confidence interval (CI) were calculated based on Clopper-Pearson exact confidence interval.

Analysis of secondary efficacy endpoint-LIC

Analysis of LIC data was conducted only in those patients for whom baseline and post-Ferriprox LIC values were assessed by the same measurement technique (liver biopsy, SQUID or MRI). The success rate was calculated as the proportion of patients with a reduction in LIC by ≥ 20% (LIC baseline value > 7 mg Fe/g dw) within one year of Ferriprox treatment. The success rate by study and the overall success rate and its 95% confidence interval (CI) were calculated based on Clopper-Pearson exact confidence

Analysis of secondary efficacy endpoint-MRI T2*
The success rate was calculated as the proportion of patients with a increase in cardiac MRI T2* by ≥ 20% (MRI T2* baseline value < 20 ms) within one year of Ferriprox treatment. The success rate by study and the overall success rate and its 95% confidence interval (CI) were calculated based on Clopper-Pearson exact confidence

If the lower limit of the 95% CI for any efficacy measure was greater than the pre-defined criterion of treatment success (20%), the therapy was considered to be a success for that particular measure.

The last observation carry forward (LOCF) method was used for data imputation for patients who had not completed one year on Ferriprox therapy.

Reviewer's Comments:
This study is a retrospective study designed to analyze existing data from studies conducted previously to evaluate the efficacy of Ferrriprox. Efficacy data in LA36-0310 were derived from 12 of 17 studies; the study involves pooling data from 12 studies of different designs; different durations of drug treatment; different time intervals of patient visits; different inclusion and exclusion criteria; and different patient characteristics. This inevitably raises an issue of data integrity and heterogeneity that can potentially cause biases. There are variations among all studies in the success rates based on the serum ferritin point estimate. For instance, the success rate for study LA12 was as low as 26%, while that of study LA15 was as high as 100%. To investigate this issue, the reviewer performed meta-analysis for these 12 studies, and presented the results in Section 3.2.5

3.2.5 Results and Conclusions

3.2.5.1 Analysis results for serum ferritin

For serum ferritin, success rates and its 95% CIs by studies are summarized in the Table 9. The overall success rate was 52%, 95% CI was (45% to 58%).

Table 9 Reviewer's summary of success rate by study for serum ferritin (ITT)

Frequency Row Pct	Success N %	Number of Patients	Clopper-Pearson 95% CI
BP	12 44.44	27	(25.48, 64.67)
LA_01	4 50.00	8	(15.70, 84.30)
LA_0206	26 40.00	65	(28.04, 52.90)
LA_03	5 62.50	8	(24.49, 91.48)
LA_04	29 51.79	56	(38.03, 65.34)
LA_08	4 57.14	7	(18.41, 90.10)
LA_11	10 83.33	12	(51.59, 97.91)
LA_12	5 26.32	19	(9.15, 51.20)
LA_15	18 100.00	18	(81.47, 100)
LA_16	4 80.00	5	(28.36, 99.49)
LA_28	2 66.67	3	(9.43, 99.16)
LA_30	17 47.22	36	(30.41, 64.51)
Total	136 51.5	264	(45.31, 57.69)

Reviewer's Comments:

The overall success rate was 52% with its 95% CI of (45.3%, 57.7%). These results are consistent with the sponsor's results. As the lower limit of the 95% CI is greater than 20% for this subset patient, the results support that Ferriprox is an effective treatment in reducing serum in patients who failed standard cheleation therapy. However, there is a large amount of variation among these 12 studies. For example: the success rate for LA 12 was 26% with 95% CI

of (9.2%, 51.2%); and the success rate for LA15 was 100% with 95% CI of (81.5%, 100%). Therefore, simply pooling these 12 studies together could make the results unreliable.

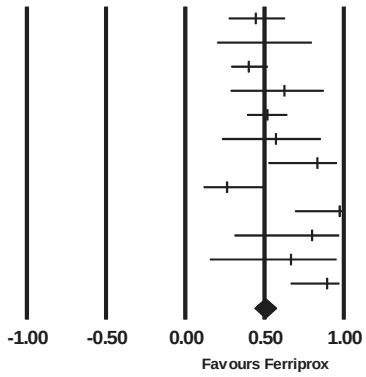
The lower bounds of the 95% confidence interval for 4 studies (LA-01, LA-12, LA-18, LA-28) are less than 20%.

Table 10 shows the reviewer's meta-analysis of serum ferritin

Table 10 Reviewer's meta-analysis for serum ferritin

Reviewer's Meta Analysis for Serum Ferritin

Study name	Statistics for each study					Event rate and 95%CI
	Event rate	Lower limit	Upper limit	Z-Value	p-Value	
BP	0.444	0.272	0.631	-0.576	0.565	
LA_01	0.500	0.200	0 800	0.000	1.000	
LA_0206	0.400	0.289	0.523	-1.601	0.109	
LA_03	0.625	0.285	0 875	0.699	0.484	
LA_04	0.518	0.389	0.645	0.267	0.789	
LA_08	0.571	0.230	0 856	0.377	0.706	
LA_11	0.833	0.523	0 958	2.078	0.038	
LA_12	0.263	0.114	0.498	-1.976	0.048	
LA_15	0.974	0.690	0 998	2.519	0.012	
LA_16	0.800	0.309	0 973	1.240	0.215	
LA_28	0.667	0.154	0 957	0.566	0.571	
LA_30	0.895	0.663	0 974	2.863	0.004	
	0.503	0.435	0 572	0.096	0.924	



Meta Analysis

Reviewer's Comments:

The meta-analysis using the Meta-Analysis software showed that the overall success rate was 50% with 95% CI of (43.5%, 57.2%). These values were lower than the analyses results from simple pooling. However, meta- analyses can not address heterogeneity among studies, such as different study design, different patient characteristics and so on.

This reviewer did sensitivity analysis by using worst-case-scenario missing data imputation. The 137 subjects who have missing values for the serum ferritin by the end of 12 month are imputed as non-response. The analysis showed the overall success rate is 26.89% with 95% CI of 21.64 to 32.67, with the lower bound being still slightly above 20%.

The reviewer also did sensitivity analysis for serum ferritin by using per-protocol population; the results showed the overall success rate is 51% with 95% CI of (45% to 58%).

Descriptive statistics for serum ferritin at baseline, last observation within 1 year + 3 months, and change from baseline to last observation within 1 year + 3 months, and change from baseline to last observation within 1 year + 3 months are presented in Table 10. Mean serum ferritin decreased by 962 ug/L within one year of therapy, from 4416 ug/L at baseline to 3453 ug/L at the last observation. These results are consistent with sponsor's results.

Table 11 Reviewer's summary of descriptive statistics for serum ferritin

Label	N	Mean	Std Dev	Minimum	Maximum
Baseline value	264	4415.49	2288.43	2505.00	16550.00
Last obs within 1 year + 3 months	264	3453.12	2098.65	184.0000000	16139.00
Change in serum ferritin from Baseline	264	-962.3708333	1907.36	-10385.00	10002.00

The influences of covariates for age, sex, baseline serum ferritn using logistic regression were summarized in the Table 12.

Table 12 Reviewer's Estimate for covariates using logistic regression for serum ferritin

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
AGE		1	0.0104	0.0106	0.9602	0.3271
SEX	F	1	-0.0597	0.1259	0.2246	0.6355
BASELINE		1	0.000142	0.000062	5.2106	0.0224

Reviewer's Comments:

Logistic regression indicates that age and sex were not significant covariates in predicting serum ferritin change from baseline, with p-values of 0.33 and 0.64, respectively. In contrast, the baseline of serum ferritin was a significant factor, with p-value of 0.02. Higher baseline of serum ferritin values could increase the probability of success rate with the treatment.

Table 13 and Figure 3 show reviewer's summary of mean serum ferritin over time. The figure shows decline in serum ferritin over time, continuing beyond one year of Ferriprox therapy. There was notable variation of the mean serum ferritin values after about 72 months, this because low numbers of patients beyond 72 months (N≤4) do not provide mean serum ferritin values representative of the full cohort of patiens.

Table 13 Reviewer's summary of mean serum ferritin over time within 24 month

MONTH	N	MEAN	STD	MINIMUM	MAXIMUM
0	264	4415.49	2288.43	2505.0	16550.0
3	244	3817.63	2198.88	853.0	16139.0
6	185	3634.46	1844.29	1102.0	11878.0
9	170	3501.67	1887.72	462.0	10836.8
12	137	3250.62	1790.47	184.0	9955.0
15	122	3254.46	1866.04	249.0	11839.8
18	108	3085.87	1659.57	169.5	8214.5
21	96	3333.80	1741.66	388.0	9073.7
24	96	3312.44	1834.56	407.0	9982.0

Figure 3 Reviewer's summary of mean serum ferritin over time

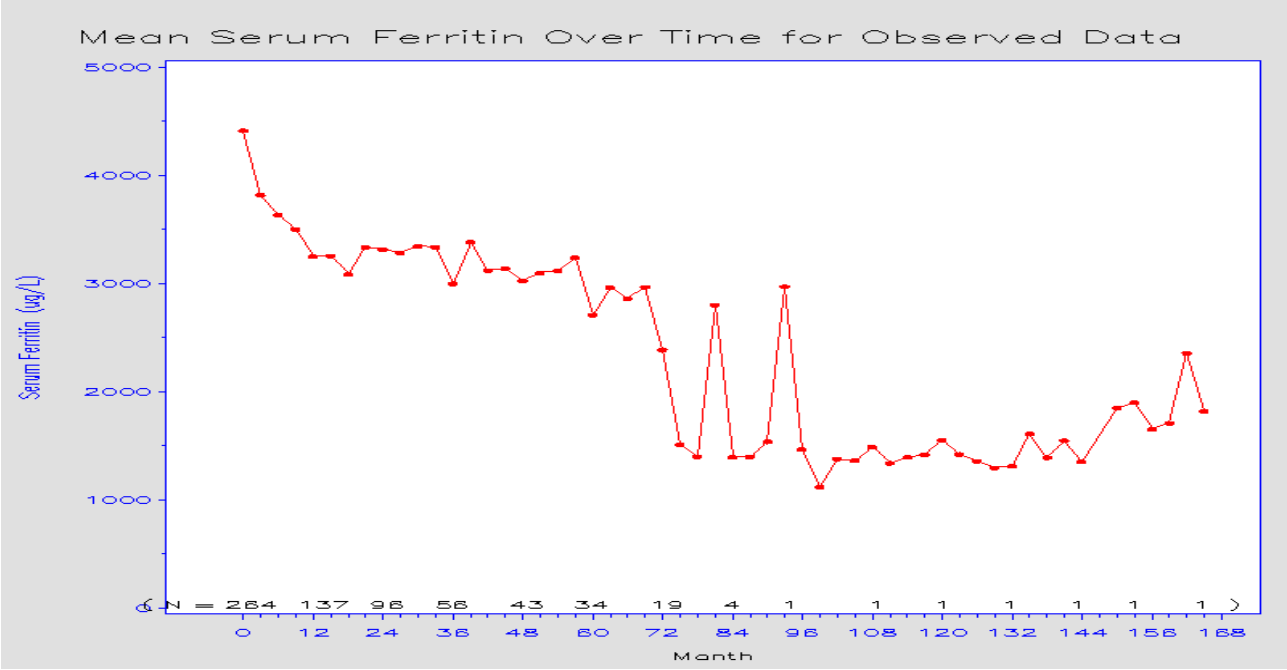
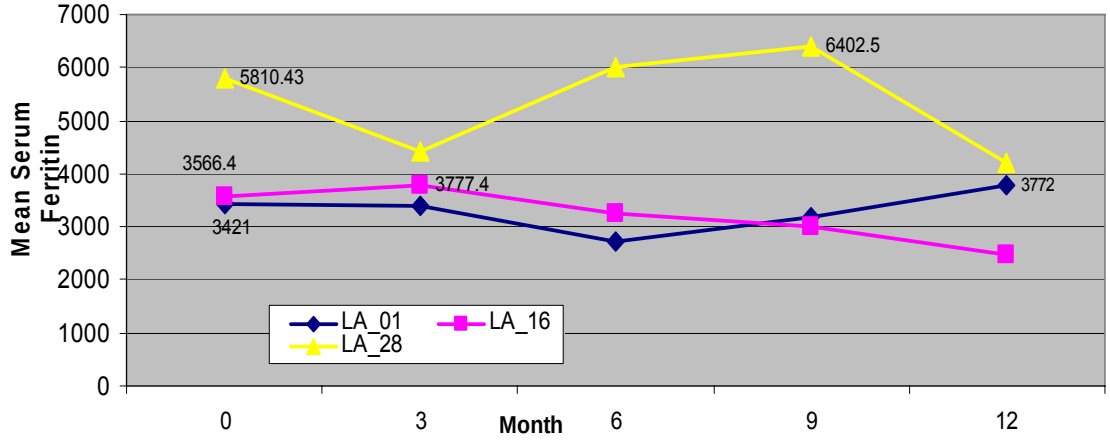


Figure 4 shows the reviewer's analysis of change in mean serum ferritin over time for study LA-01, LA-16 and LA-28.

Figure 4 Reviewer's analysis of mean serum ferritin change over time (study LA-01, LA-**16, LA-28)**



Reviewer's comments:
The mean serum ferritin change from baseline over time for study LA-01, LA-16 and LA-28 did not show decrease overtime.

For LIC, success rates and its 95% CIs by studies were summarized in the Table 14. The overall success rate was 41%, 95% CI was (32% to 51%).

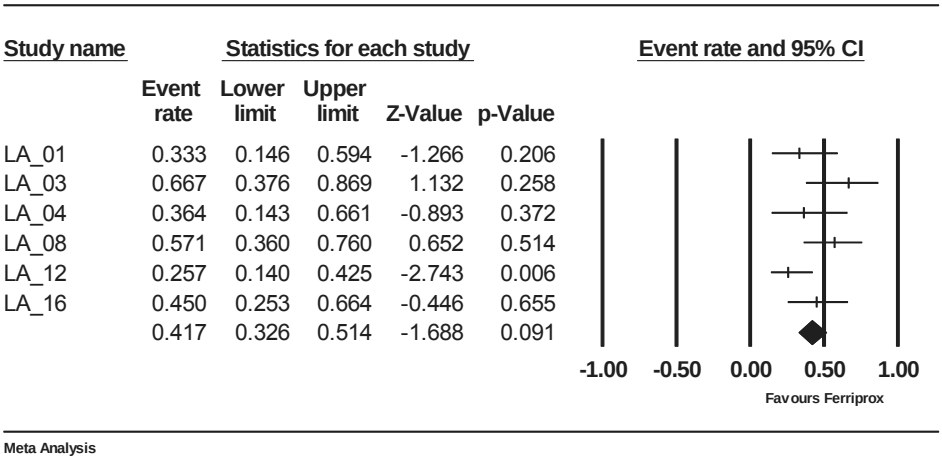
Table 14 Reviewer's summary of success rate by study for LIC

Study	Success	Number of Patients	Clopper-Pearson 95% CI
LA_01	5 (33.33%)	15	(11.82, 61.62)
LA_03	8 (66.67%)	12	(34.89, 90.08)
LA_04	4 (36.36%)	11	(10.93, 69.21)
LA_08	12 (57.14%)	21	(34.02, 78.18)
LA_12	9 (25.71%)	35	(12.49, 43.26)
LA_16	9 (45.00%)	20	(23.06, 68.47)
Total	47 (41.23%)	114	(32.09, 50.83)

Reviewer's Comments:
The overall success rate was 41.23% with its 95% CI of (32.09%, 50.83%). These results are consistent with the sponsor's results. As the lower limit of the 95% CI is greater than 20% for this subset of patients, the results support that ferriprox is an effective treatment in reducing LIC in patients who failed standard cheleation therapy. However, there were significant variations among these 6 studies. For example: the success rate for LA 04 was 25.71% with 95% CI of (12.49%, 43.26%); and the success rate for LA15 was 66.67% with 95% CI of (34.89%, 90.08%). Therefore, simply pooling these 6 studies could make the results unreliable.

Table 15 shows the reviewer's meta- analysis of LIC

Table 15 Reviewer's meta-analysis of LIC



Reviewer's Comments:
The meta-analysis using the Meta-Analysis software showed that that the overall success rate was 41.7% with 95% CI of (32.6%, 51.4%). These values were slightly different from the analysis results without using meta- analyses. However, meta- analyses can not address the heterogeneity among studies, such as differences in study designs, differences in patient characteristics, and so on.

Descriptive statistics for LIC at baseline, last observation within 1 year + 3 months, and change from baseline to last observation within 1 year + 3 months, and change from baseline to last observation within 1 year + 3 months are presented in Table 15. Mean serum ferritin decreased by 962 ug/L within one year of therapy, from 4416 ug/L at baseline to 3453 ug/L at the last observation. These results are consistent with sponsor's results.

Table 16 Reviewer's summary of descriptive statistics for LIC

Label	N	Mean	Std Dev	Minimum	Maximum
Baseline value	114	15.9137026	10.0999578	7.1000000	66.6000000
Numeric Result/Finding in Standard Units	114	14.5686832	9.1410746	1.6000000	54.2000000
Change in LIC from Baseline	114	-1.3450195	6.9271135	-27.9400000	14.5000000

The influences of covariates for age, sex, baseline serum ferritn using logistic regression were summarized in the Table 17.

Table 17 Reviewer's Summary of estimate for covariates using logistic regression for LIC

Analysis of Maximum Likelihood Estimates						
Parameter		**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Baseline		1	0.0225	0.0194	1.3448	0.2462
SEX	**F**	1	0.1667	0.1964	0.7208	0.3959
AGE		1	0.0586	0.0319	3.3743	0.0662

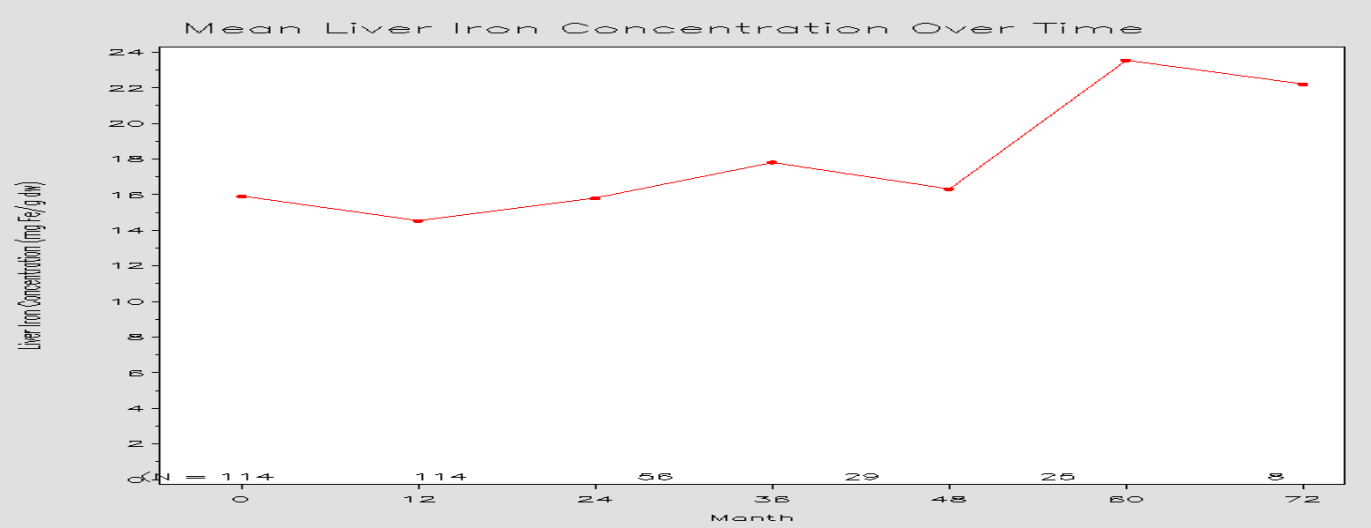
Reviewer's Comments:
Logistic regression indicates that age, sex and baseline LIC were not significant covariates in predicting LIC change from baseline, with p-values of 0.07, 0.40 and 0.25, respectively.

Table 18 and Figure 4 below show reviewer's summary of mean serum ferritin over time. The figure shows decline in serum ferritin over time, continuing beyond one year of Ferriprox therapy. There was notable variation of the mean serum ferritin values after about 72 months, this may be because low numbers of patients beyond 72 months (N≤4) do not provide mean serum ferritin values representative of the full cohort of patients.

Table 18 Reviewer's summary of mean serum ferritin over time within 24 month

MONTH	N	MEAN	STD	MINIMUM	MAXIMUM
0	114	15.91	10.10	7.1	66.6
12	114	14.54	9.22	1.6	54.2
24	56	15.80	7.97	2.3	35.5

Figure 5 Reviewer's summary of mean LIC over time



3.2.5.3 Analysis results for Secondary Endpoint: MRI T2*

For MRI T2*, success rates and its 95% CIs by studies were summarized in the Table 19. The overall success rate was 61.54%, 95% CI was (44.62% to 76.64%).

Table 19 Reviewer's summary of success rate by study for MRI T2*

Study	Success	Number of Patients	Clopper-Pearson 95% CI
LA_04	6 (60.00%)	10	(26.00 88.09%)
LA_16	18 (62.07%)	29	(42.03, 79.17)
Total	24 (61.54%)	39	(44.62, 76.64)

Reviewer's Comments:

The success rate for LA 04 was 60% with 95% CI of (26%, 88%); the success rate for LA16 was 62% with 95% CI of (42%, 79%). The overall success rate was 61.54% with its 95% CI of (44.62%, 76.64%). These results are consistent with the sponsor's results. As the lower limit of the 95% CI is greater than 20% for this subset patient, the results support that ferriprox is an effective treatment in reducing MRI T2 in patients who failed standard cheleation therapy.*

Descriptive statistics for MRI T2* at baseline, last observation within 1 year + 3 months, and change from baseline to last observation within 1 year + 3 months, and change from baseline to last observation within 1 year + 3 months are presented in the Table 20. These results are consistent with sponsor's results.

Table 20 Reviewer's summary of descriptive statistics for MRI T2*

Label	N	Mean	Std Dev	Minimum	Maximum
Baseline value	39	11.7558974	4.8670496	4.0000000	19.5000000
Numeric Result/Finding in Standard Units	39	15.0579487	7.0363953	3.4000000	28.0000000
Change in MRI T2* from Baseline	39	3.3020513	3.3611865	-2.0000000	12.7000000

The influence of covariates for age, sex, baseline MRI T2* using logistic regression were summarized in the Table 21.

Table 21 Reviewer's Summary of estimate for covariates using logistic regression for MRI T2*

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	1.2545	1.9058	0.4333	0.5104
LBSTBASE	1	0.0842	0.0742	1.2852	0.2569
SEX	1	-0.0953	0.3438	0.0768	0.7817
AGE	1	-0.0722	0.0789	0.8385	0.3598

Reviewer's Comments:

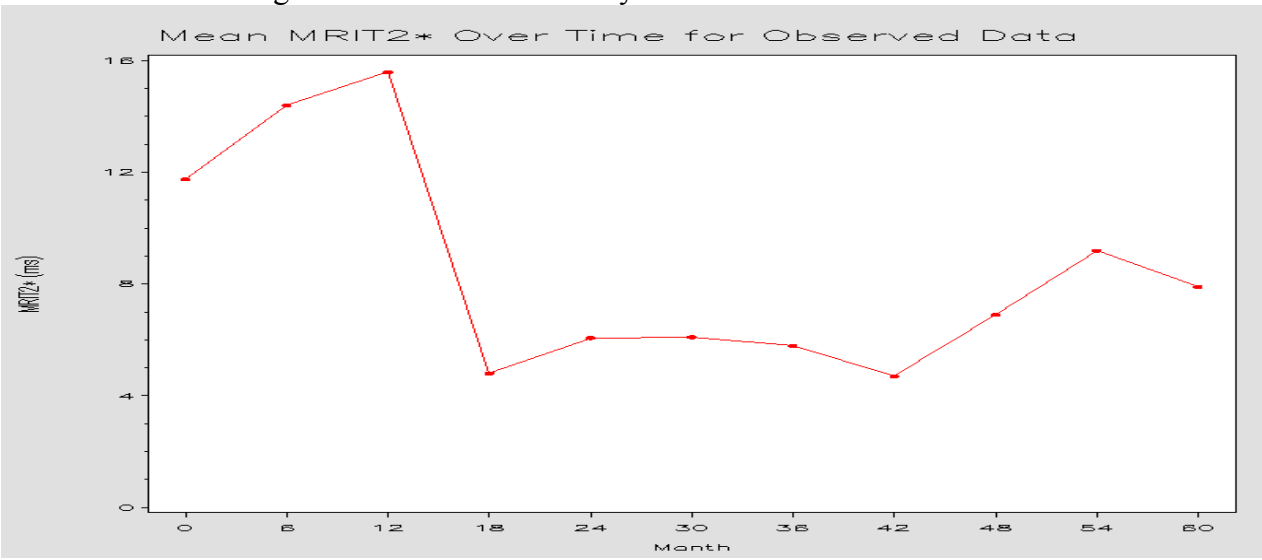
*Logistic regression indicates that age, sex and baseline MRI T2*were not significant covariates in predicting MRI T2* change from baseline, with p-values of 0.36, 0.87 and 0.26, respectively.*

Table 22 and Figure 5 below show reviewer's summary of mean MRI T2*over time. The figure shows decline in serum ferritin over time, continuing beyond one year of Ferriprox therapy. There was notable variation of the mean serum ferritin values after about 72 months, this maybe because low numbers of patients beyond 72 months (N≤4) do not provide mean serum ferritin values representative of the full cohort of patiens.

Table 22 Reviewer's summary of mean MRI T2* over time within 24 month

Obs	MONTH	N	MEAN	STD	MINIMUM	MAXIMUM
1	0	39	11.76	4.87	4.0	19.5
2	6	37	14.40	6.89	2.8	30.2
3	12	32	15.60	6.82	3.4	28.0
4	18	2	4.80	0.99	4.1	5.5
5	24	2	6.07	1.80	4.8	7.3

Figure 6 Reviewer's summary of mean MRI T2* over time



4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race, Age, and Geographic Region

Efficacy Endpoint of Serum Ferritin

Reviewer's subgroup analysis of gender, race, age and geographic region were summarized in the Table 23.

Table 23 Reviewer's subgroup analysis of gender, race, age and region for serum ferritin

Subgroup Category	Number of Patients	Success Rate (N, %) (95% CI)	P-value
Age			
<16	83	38 (46%) (35, 57)	0.23
≥16	181	98 (54%) (47, 62)	
Gender			
Female	145	73 (50%) (42, 59)	0.71
Male	119	63 (53%) (44, 62)	
Race			
White	194	95 (49%) (42, 56)	
Asian	46	27 (59) (43, 73)	
Unknown	21	12 (62) (38, 82)	
Country			
US	36	22 (61%) (43, 77)	0.28
Non-US	228	114 (50%) (43, 57)	
Region			
European Countries	136	54 (40%) (31, 48)	0.0001
Non-European	128	82 (64%) (55, 72)	

Note: All the nominal p-values of testing difference are provided for information only and based on two-sided Fisher's exact test

Reviewer's Comments:
There was no statistically significant difference in success rate by age, gender, or race groups. The success rates are significantly different among regions. Namely, there was a statistically significant difference in success rate between European countries and non-European countries, with p-value of 0.0001. However, the lower limit of the 95% CI was greater than 20% for all of the subsets of patients involved in the subgroup analyses.

4.2 Other Special/Subgroup Populations

Efficacy Endpoint of Serum Ferritin

Table 24 shows the reviewer's subgroup analysis by mean of baseline serum ferritin (>4415 vs ≤ 4415) and thalassemia (yes vs no).

Table 24 Reviewer's summary of subgroup analysis for serum ferritin by baseline of serum ferritin and thalassemia

Subgroup Category	Number of Patients	Success Rate (N, %) (95% CI)	P-value
Baseline Serum Ferritin			
>4415	90	54 (60%) (49, 70)	0.05

≤ 4415	174	82 (47%) (40, 55)	
Thalassemia			
Yes	228	115 (50%) (44, 57)	0.47
No	36	21 (58%) (41, 74)	

Note: All the nominal p-value are provided for information only and based on two-sided Fisher's exact test

Reviewer's Comments:
There were more responses (54) in the group with baseline serum ferritin >4415 compared to the group with baseline serum ferritin ≤4415. There was no statistically significant difference in success rates between patients with Thalaessmia and those without Thalassemia.

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

Statistical Issues

1. This study is a retrospective study designed to analyze existing data from studies conducted previously to evaluate the efficacy of Ferrriprox. Efficacy data in LA36-0310 were derived from 12 of 17 studies; the study involves pooling data from 12 studies of different designs; different durations of drug treatment; different time intervals of patient visits; different inclusion and exclusion criteria; and different patient characteristics. This inevitably raises an issue of data integrity and heterogeneity that can potentially cause biases. The sponsor followed the suggestion from the Agency and used Independent Committee to pick the eligible subjects. However, such approach can only limit biases, but can not eliminate biases.

2. There were only 2 black patients in the study LA36-0310. Such patient proportion did not represent US patient population.

3. There were more female patients than male patients (55% vs 45%). The majority of patients were white in race (73%)

4. A total of 20 (7.58%), 79 (29.92%), 94 (35.61%) and 127 (48.11%) subjects dropped out by month 3, 6, 9 and 12, respectively. High drop out rate could make the analyses results unreliable and misleading.

5. There are variations among all studies in the success rates based on the serum ferritin point estimate. For instance, the success rate for study LA12 was as low as 26%, while that of study LA15 was as high as 100%. The primary efficacy analysis pooling all 12 studies together ignoring the variation between studies could introduce unrealizable results.

6. Change in liver iron concentration was available for only a minority of patients and showed the weakest efficacy results

7. Change in cardiac MRI T2* was small

8. The subgroup analysis did not show internal consistency. The lower bound of the 95% confidence interval for 4 studies (LA-01, LA-12, LA-18, LA-28) were all below 20%.

9. By the entry criteria all patients must have received some prior chelation therapy (deferoxamine or deferasirox). The reviewer found that among the 264 patients, there are 23 patients who have no record about whether they took the prior chelation therapy. However, excluding these 23 patients will not alter the conclusion for primary efficacy. These 23 patients are:
   LA_01: 66, 67;
   LA_12: 171, 31, 36;
   LA_0206: 222, 264, 7, 9
   LA_04: 2, 40, 99
   LA_11: 102, 104,106,107,108,109, 112,113,118,121,122, 124

Findings

1. For serum ferritin, the overall success rate was 52%, and the lower limit of the 95% CI for the overall success rate was 45%. As the lower limit of the 95% CI is greater than 20% for this subset of patients, the results support that Ferriprox is an effective treatment in reducing serum ferritin in patients who failed standard cheleation therapy. However, there are significant variations among these 12 studies. Simply pooling these 12 studies could cause make the results unreliable. The reviewer's meta-analysis showed that the lower limit of the 95% CI for the overall success rate was 43.5%, slightly lower than the result from simple pooling analysis. The sensitivity analyses using per-protocol analysis and using worst-case-scenario missing data imputation did not change the conclusion from the primary efficacy analysis.

2. For LIC, the overall success rate was 41%, and the lower limit of the 95% CI for the overall success rate was 32%.

3. For MRI T2*, the overall success rate was 62%, and the lower limit of the 95% CI for the overall success rate was 45%.

4. Logistic regression indicates that age and sex were not significant covariates in predicting serum ferritin change from baseline, with p-values of 0.33 and 0.64, respectively. In contrast, the baseline of serum ferritin was a significant factor, with p-value of 0.02. Higher basline value of serum ferritin had higher probability of success rate in the treatment.

5. For subgroup analyses of serum ferritin, there was no statistically significant difference in success rate by age, gender, or race groups. The success rates are significantly different among regions. Namely, there was a statistically significant difference in success rate between European countries and non-European countries, with p-value of 0.0001. However, the lower limit of the 95% CI was greater than 20% for all of the subsets of patients involved in the subgroup analyses.

5.2 Conclusions and Recommendations

Deferiprone (Ferriprox) is an orally administered iron chelator that is being developed for the indication of the treatment of persons who had developed transfusion related hemosiderosis because of a chronic underlying anemia. NDA 21825 for Ferriprox (deferiprone) was submitted on January 29, 2009 and a Complete Response letter was sent to the sponsor on November 30, 2009. This resubmission of NDA 21825/SE0056 including study LA36-0310 was designed as an analysis of existing data from studies previously conducted to evaluate the efficacy of Ferriprox. No new data were collected and the original purpose of collecting the data and their application did not change. Efficacy data for study LA 36-0310 were derived from 12 of 17 studies. The primary efficacy endpoint was the change in serum ferritin concentration from baseline within one year of Ferriprox therapy. Ferriprox therapy was considered successful in individual patients who experienced a ≥20% decline in serum ferritin concentration within one year of therapy.

The sponsor's efficacy analysis for serum ferritin by pooling 12 studies showed that the overall success rate was 52% with 95% CI of (45%, 58%). As the lower limit of the 95% CI is larger than 20%, the protocol defined endpoint was met for this trial. However, this study has several serious limitations including lack of randomization, lack of control group, high rate of missing data and ignoring the variation between studies by simple pooling, all of which can introduce biases to the primary outcome. Therefore, it is unclear whether the efficacy shown in the study is solely due to the Ferriprox therapy, and the interpretation of these analysis results should be taken cautiously.

Appendix

The following analyses results are based on the AC member's questions during the Advisory Committee Meeting,

Q1. If the thresholds for response are changed from 20% to 30%, 40% and 50%, then what would be the response rates?

Answer: The following table gives the reviewer's summary of response by different threshold for serum ferritin.

Reviewer's summary of response rate by different threshold

Serum Ferritin Decline from Baseline	Frequency	Percent	Cumulative Frequency
<20%	128	48.48	128
>=20%	32	12.12	160
>30%	30	11.36	190
>40%	23	8.71	213
>50%	51	19.32	264

Q2. How many patients had serum ferritin value <2500 at 1 year?

Answer: Among 264 patients, there are 96 (36%) patient's serum ferritin value<2500 at 1 year.

CHECK LIST

Number of Pivotal Studies: 1

Protocol Number (s): LA 36-0310
Phase: 3
Control: None
Blinding: Retrospective study
Number of Centers: Multiple Centers
Region(s) (Country): Multiple Regions
Duration: 1 year
Treatment Arms: Ferriprox
Randomization: Retrospective study
Primary Endpoint: The change in serum ferritin concentration from baseline within one year of Ferriprox therapy
Primary Analysis Population: ITT

Statistical Design: designed as an analysis of existing data from studies previously conducted to evaluate the efficacy of Ferriprox. No new data were collected and the original purpose of collecting the data and their application did not change. Efficacy data for study LA 36-0310 were derived from 12 of 17 studies.

Primary Statistical Methodology: 95% CI

Interim Analysis: No

Sample Size: 264

- **Sample Size Determination**: Independent Committee selected subjects from the entire integrated dataset following approved inclusion and exclusion criteria.
- Were the **Covariates** pre-specified in the protocol? Yes
- Did the Applicant perform **Sensitivity Analyses**? Yes
- How were the **Missing Data** handled? Worst-case scenario conducted by reviwer
- Was there a **Multiplicity** involved? No
- **Multiple Secondary Endpoints**: Are they being included in the label? No

Were Subgroup Analyses Performed: Yes

- Were there any **Discrepancies** between the protocol/statistical analysis plan vs. the study report: No
- Overall, was the study positive :Yes

SIGNATURES/DISTRIBUTION LIST

Primary Statistical Reviewer: Qing Xu, Ph.D
Date: November 14, 2011

Statistical Team Leader: Rothmann Mark, Ph.D

Biometrics Division Director: Rajeshwari Sridhara, Ph.D

cc:
Project Manager: Mara Bauman Miller
Medical Officer: George Shashaty, M.D
Medical Team Leader: Kathy M Robie Sub, M.D
Medical Division Director: Ann T Farrel, M.D
Primary Statistical Reviewer: Qing Xu, Ph.D
Statistical Team Leader: Rothmann Mark, Ph.D
Biometrics Division Director: Rajeshwari Sridhara, Ph.D
Lillian Patrician

c:\NDA\statreview.doc

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

/s/

QING XU
09/16/2011

MARK D ROTHMANN
09/16/2011

RAJESHWARI SRIDHARA
09/16/2011

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

Office of Translational Science
Office of Biostatistics
Division of Biometrics V

Statistical IND Review and Evaluation

NDA#:	**21,825**	**(New Protocol Review)**
Date Received:	**12/02/2010**	
Product:	Ferriprox (Deferiprone)	
Indication:	Iron Chelator	
Sponsor:	ApoPharm. Inc	
Statistical Reviewer:	Qing Xu, Ph.D.	
Medical Reviewer:	George Shashaty, M.D.	

BACKGROUND:

The NDA was submitted to the Agency on January 29, 2009. The submission of the NDA included a single randomized controlled trial. The primary endpoint in that trial was the change in cardiac iron as measured by cardiac MRI T2* assessment after one year of treatment with deferiprone. The comparator drug was deferoxamine, which at the time of the study was the only approved drug for the indication. The Agency was concerned that the primary endpoint measured was not a validated surrogate for clinical utility. The sponsor also submitted several supportive studies and a large number of references of the use of deferiprone in patients with iron overload, again almost all of whom had thalassemia as the cause of anemia for which transfusions were required. At the conclusion of the review, the Agency sent a Complete Response (CR) letter to the sponsor describing the deficiencies of the data submitted. In the CR letter dated November 30, 2009, the Agency recommended that the sponsor perform "adequate and well-controlled" trials with deferiprone to support the application for approval of the NDA.

On September 24, 2009, the FDA communicated with the sponsor that it could not approve the deferiprone in its current form and requests additional data, which included at least one additional prospective, randomized, controlled clinical study and data establishing the minimum millisecond increase in cardiac T2* indicative of a clinical benefit. Subsequently, the FDA communicated to ApoPharma that it would consider accelerated approval for the deferiprone NDA with a single arm study in patients intolerant of or not responding to existing therapy, based on pre-existing data from the deferiprone clinical program.

A draft of this protocol LA36-0310, dated September 13, 2010 was submitted and reviewed, and a list of comments was forwarded to the sponsor from the Agency on November 9, 2010. The current submission includes the revision for draft protocol LA36-0310 with the sponsor's responses to the Agency's comments.

The primary objective:
To evaluate the efficacy of oral administration of deferiprone in the treatment of iron overload in patients in whom standard chelation has failed. Chelation failure is defined as iron accumulation above a boundary level, defined by high serum ferritin or LIC, or low cardiac MRI T2* levels. Success rate will be evaluated by compiling data from clinical trials that recorded one or more of the following effects during treatment with deferiprone for up to 1 year, and where data on serum ferritin, LIC, or cardiac MRIT2* are available prior to and after starting therapy with deferiprone. To allow assessment of patients whose efficacy assessments were conducted at about the 1 year anniversary of initiation of treatment, the up to 1 year window will include data obtained within ±3 months of the anniversary date.

Study Design:
This is a retrospective analysis of data derived from multiple studies that were conducted to support NDA 21-825, some controlled and some uncontrolled, in patients who mostly had an underlying diagnosis of beta thalassemia and transfusional iron overload. Patients to be enrolled into the current study will be programmatically selected from database submitted to the Agency as part of the NDA and integrated up to 11 May 2010 for ongoing clinical studies. The sponsor will subsequently assess the serum ferritin, LIC and cardiac MRI T2* values captured during treatment with deferiprione for up to 1 year for analysis of its efficacy in the cohort of patients selected by Clinical Data Management. The number of deferiprone-treated patients who meet the defined criteria for successful treatment outcomes will be determined and a success rate will be calculated. The lists of studies are summarized in the Appendix 1.

Study Population:
Study population will be composed of patients who show evidence of failed responses to previous chelation therapy.

Efficacy Endpoints:
The primary efficacy endpoint: will be the change in serum ferritin concentration from baseline within one year of deferiprone therapy. The efficacy of deferiprone in each patient will be established by determining whether treatment with deferiprone for up to 1 year (including therapy interruptions and patients who have discontinued therapy during that period due to adverse events or other reasons) has succeeded or failed in promoting no less than 20% decrease in serum ferritin concentration in a clinically acceptable proportion of patients. The current study will conclude that deferiprone is effective if a successful outcome is observed in at least 20% of patients who failed previous chelation therapy.

Secondary endpoints of: will be the changes in LIC and in cardiac MRI T2*.

Statistical Methods:
Analysis Population:
Two analysis populations will be used in the study:

- **Intent-to-treat (ITT) population** will be the primary population for this study. For each efficacy measure, the ITT population will be composed of those patients who had taken at least one dose of deferiprone and had at least one post-baseline measurement of that efficacy measure
- **Per Protocol population** will be the secondary population for this study. For each efficacy endpoint, the PP population will be composed of those patients who had completed the study that they originated from or at least one year of deferiprone therapy for long-term studies, and had no missing data for the end-of-study measurement or the last scheduled measurement at the end of the first year, respectively, for that efficacy measure.

Sample Size Estimation:
There is no sample size calculation in this study. The sponsor intends to have data from as many eligible patients as possible to determine the efficacy of the drug. The following table proposed by sponsor shows the sample size required to meet the criterion that the lower limit of the 95% confidence interval for the success rate is >20% for different expected success rates.

Sample size calculation for which the lower limit of 95% CI is at least 20%

Sample Size	Expected Success Rate	95% CI
1,648	22%	(20%, 24%)
438	24%	(20%, 28%)
206	26%	(20%, 32%)
122	28%	(20%, 36%)
81	30%	(20%, 40%)
59	32%	(20%, 44%)
44	34%	(20%, 48%)
40	36%	(20%, 52%)

Analysis of efficacy variables:
All statistical tests will be two-sided and p-value≤0.05 will be used for the determination of statistical significance in the statistical tests.

The primary efficacy endpoint will be based on the change in serum ferritin concentration from baseline to the end-of-study data or the last observation within 1 year of deferiprone therapy for studies of >1 year in duration. For deferiprone to be a successful treatment in patients who failed standard chelation therapy, there must be at least 20% of patients who demonstrate a favorable response on serum ferritin. The success rate will be calculated. A patient will be determined to be a successful responder if the patient's serum ferritin was reduced by at least 20% within 1 year of deferiprone treatment. The success rate by study and overall success rate and its 95% C.I. will be calculated based on normal approximation.

Secondary efficacy endpoints will be based on the changes in cardiac MRI T2* and LIC

from baseline to the end-of-study data or the last observation within 1 year of deferiprone therapy for studies of >1 year in duration. For deferiprone to be a successful treatment in patients who failed standard chelation therapy based on cardiac MRI T2*, there must be at least 20% of patients who demonstrate a favorable response on MRI T2*. The success rate will be calculated. A patient will be determined to be a successful responder if the patient's MRI T2* was increased by at least 20% 1 year of deferiprone treatment. The success rate by study and overall success rate and its 95% C.I. will be calculated based on normal approximation. For deferiprone to be a successful treatment in patients who failed standard chelation therapy based on LIC, there must be at least 20% of patients who demonstrate a favorable response on LIC.The proposed ≥20% reduction in serum ferritin or LIC is based on the understanding that it would mean a drop of a minimum of 500 µg/L or 1.4 mg Fe/g dw, respectively.

Overall Treatment Success
If the lower limit of the 95% confidence interval for any efficacy measure is greater than the pre-defined criterion of treatment success (20%), the therapy will be considered to be a success for that particular measure. The final assessment of treatment success for the purpose of this study will be based on the success rate for serum ferritin, the primary efficacy measure.

Missing Data Imputation
The sponsor intended to treat those subjects who are not known to have had a decrease in serum ferritn of at least 20% as a non-responder.

Subgroup Analyses
The sponsor will calculate success rates for the subgroup of patients with 2 or more serum deferiprone values, of which a majority of the values were greater than 2,500 µg/L, before starting deferiprone in order to determine if those patients exhibited the same degree of response to deferiprone treatment as did the patients who had failed on standard chelation therapy, but had only a single baseline ferritin concentration to designate them as a failed patient.

Sensitivity Analysis:
The sponsor proposed sensitivity analyses which include:

- Primary efficacy analysis using Pre Protocol population.
- LOCF method for imputation of missing efficacy endpoint data.

Statistical Conclusion and Recommendations:
In the statistical analysis plan (SAP), the sponsor has responded to the issues raised by the Agency in regard to the missing data imputation using LOCF, and selection bias from this retrospective study. The sponsor provided justification of LOCF imputation and justification of selection bias in the statistical analysis plan.

LOCF Missing Data Imputation
The sponsor believes that there is no subjective assessment involved by using LOCF missing data imputation for the efficacy data and the approach is easy to understand. They also believe, with the

proposed treatment success criteria (≥20% improvement in the efficacy endpoint) being applied to patients who had shown a poor response to previous chelation therapy, these patients would be considered treatment failure regardless of the LOCF data, thereby not favouring deferiprion.

Justification of Patient Selection Bias
The sponsor believes that patient selection bias may occur in a situation where precise inclusion and exclusion criteria have not been defined a priori before the study. In this study, the inclusion and exclusion criteria are clearly defined in the protocol prior to the patient selection process.

FDA's comments to the sponsor's responses:

FDA Comment #6 (November 9 2010):
The precise inclusion and exclusion criteria could control the selection bias by a certain level, but cannot eliminate the selection bias. To minimize such bias, we suggest using an Independent Committee to select subjects from the entire integrated dataset following approved inclusion and exclusion criteria by the Agency.

Sponsor's Response
ApoPharma agrees to use an independent committee to select subjects from the entire integrated dataset, following the approval by the Agency of the inclusion and exclusion criteria for study LA36-0310.

FDA's Response
No further comments

FDA Comment #7 (November 9 2010):
The subgroup analysis should also include age, gender, region and some other important baseline characteristics. The consistency or lack of consistency of the results should be evaluated across subgroups.

Sponsor's Response
ApoPharma concurs with analyzing the data based on age (pediatric vs. adult), gender, and primary baseline disease (thalassemia major vs. non-thalassemia major) for the subgroup analysis. Given that the ApoPharma Ferriprox studies were conducted largely in Italy and Greece, a subgroup analysis of data from European and non-European countries will be conducted. In each subgroup analysis, the result of the primary outcome will be compared between subgroups to check for consistency of the results.

FDA's Response:
No further comments

FDA Comment #8 (November 9 2010):

This evaluation involves pooling data from many studies of different designs; different durations of drug treatment; different time intervals of patient visits; different inclusion/exclusion criteria; and different patient characteristics. These differences should be addressed.

Sponsor's Response
The design of the study was planned to accommodate restrictions pertinent to analysis of pre-existing data, which were collected from studies that had diverse inclusion/exclusion criteria, diverse objectives and diverse treatment regimens and duration, but that nonetheless used relevant and interpretable endpoints. This emphasizes the strength of longitudinal comparisons of values of the selected parameters in individual subjects. Patients will be included or excluded on the basis of the criteria selected for establishing previously failed therapy, as described in the protocol, and thus, only those patients who meet the criteria will be included in the study. The main differences among patients enrolled in study LA36-0310 will be addressed in subgroup analyses (see response to FDA comment #7).

FDA's Response
No further comments. The Agency acknowledges that Patients will be included or excluded on the basis of the criteria selected for establishing previously failed therapy, as described in the protocol, and thus, only those patients who meet the criteria will be included in the study. The main differences among patients enrolled in study LA36-0310 will be addressed in subgroup analyses.

FDA Comment #9:
We agree with treating those subjects who are not known to have had a decrease in serum ferritin of at least 20% as a non-responder.

The Sponsor's Response
ApoPharma understands that for the primary analysis, the FDA's agreement with treating those subjects who are not known to have had a decrease in serum ferritin of at least 20% as a nonresponder implies that a ≥20% decline in serum ferritin will be an acceptable endpoint.

FDA's Response:
No further comments

Additional Statistical Comments

1. Based on ICH E9, the intention-to-treat population should include all randomized subjects.
2. The overall success rate and its 95% CI should be calculated based on Pearson-Clopper exact confidence interval.

Appendix 1
List of studies

Study	Title	Number of Subjects:	Efficacy Measurements		
		Entered/Completed	Serum Ferritin	LIC	MRI T2*
LA-01	Randomized Trial of L1 and Deferoxamine in Thalassemia Major	DFP: 35/27 DFO: 36/30	√	√	
LA-02/06	Trial of APO-66 (Deferiprone) in Thalassemia/Maintenance Study Protocol of Deferrum™ (Ferriprox) for Subjects with Thalassemia who Complete Apotex Protocol LA-02	DFP: 187/70	√		
LA-03	The Long Term Efficacy and Safety of Deferiprone in Subjects with Thalassemia	DFP: 29/25	√	√	
LA-04/06B	Compassionate-use of Deferiprone (L1) in Subjects with Iron Overload	DFP:168/23	√	√	√
LA-08-9701	Safety and Efficacy of Alternating Deferoxamine and Deferiprone compared with Deferoxamine Alone in the Treatment of Iron Overload in Thalassemia Subjects	Alternating DFP and DFO: 29/29 DFO: 30/30	√	√	
LA-11	Efficacy and Safety of Deferiprone (L1) in β-Thalassemia/Hemoglobin E Diseases Subjects in Thailand	DFP: 24/16	√		

LA-12-9907	Retrospective Assessment of Heart Failure and Survival During Iron Chelation with Deferiprone or deferoxamine in Subjects with Transfusion-Dependent B-Thalassemia	DFP: 54 DFO: 75 The study assessed existing data so study completion or withdrawal was not defined.	√	√	
LA-15-0002	Safety and Efficacy of Ferriprox™ for the Treatment of Iron Overload in Subjects with Transfusion-dependent Thalassemia in Iran	DFP: 29/26	√		
LA-16-0102	Randomized Trial Comparing the Relative Efficacy of Deferiprone to that of Deferoxamine in Removing Excess Cardiac Iron in Thalassemia Major Subjects	DFP: 29/27 DFO: 32/29	√	√	√
Borgna-Pignatti et al. (2006)	Cardiac Morbidity and Mortality in Deferoxamine- or Deferiprone-treated Patients with Thalassemia Major	DFP: 157 DFO: 359	√		
LA-28-CMP	The compassionate use/named patient program of Ferriprox oral solution in iron-overloaded pediatric patients with transfusion-dependent anemias	DFP: 83/62 (72 subjects previously enrolled in LA-30)	√		
LA-30	24-week open label, uncontrolled study of the safety and efficacy of Ferriprox (deferiprone) oral solution in iron-overloaded pediatric patients with transfusion-dependent anemia	DFP: 100/95	√		

Data cut off: 11 MAY 2010

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

/s/

QING XU
12/21/2010

MARK D ROTHMANN
12/21/2010

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

Office of Translational Science
Office of Biostatistics
Division of Biometrics V

Statistical IND Review and Evaluation

NDA#:	**21,825**	**(New Protocol Review)**
Date Received:	**09/23/2010**	
Product:	Ferriprox (Deferiprone)	
Indication:	Iron Chelator	
Sponsor:	ApoPharm. Inc	
Statistical Reviewer:	Qing Xu, Ph.D.	
Medical Reviewer:	George Shashaty, M.D.	

BACKGROUND:

The NDA was submitted to the Agency on January 29, 2009. The submission of the NDA included a single randomized controlled trial. The primary endpoint in that trial was the change in cardiac iron as measured by cardiac MRI T2* assessment after one year of treatment with deferiprone. The comparator drug was deferoxamine, which at the time of the study was the only approved drug for the indication. The Agency was concerned that the primary endpoint measured was not a validated surrogate for clinical utility. The sponsor also submitted several supportive studies and a large number of references of the use of deferiprone in patients with iron overload, again almost all of whom had thalassemia as the cause of anemia for which transfusions were required. At the conclusion of the review, the Agency sent a Complete Response (CR) letter to the sponsor describing the deficiencies of the data submitted. In the CR letter dated November 30, 2009, the Agency recommended that the sponsor perform "adequate and well-controlled" trials with deferiprone to support the application for approval of the NDA.

On September 24, 2009, the FDA communicated with the sponsor that it could not approve the deferiprone in its current form and requests additional data, which included at least one additional prospective, randomized, controlled clinical study and data establishing the minimum millisecond increase in cardiac T2* indicative of a clinical benefit. Subsequently, the FDA communicated to ApoPharma that it would consider accelerated approval for the deferiprone NDA with a single arm study in patients intolerant of or not responding to existing therapy, based on pre-existing data from the deferiprone clinical program.

This submission is the sponsor's response to the discussion. The sponsor intends to review already available data on serum ferritin, cardiac MRI T2* and LIC as measures of body iron burden before and after 1 year of treatment with deferiprone in patients who could not be adequately treated with other approved chelator therapy. The source of the data will be from the studies previously submitted in the

NDA with the addition of data from completed studies of a liquid formulation of deferiprone and a continuing compassionate use program.

This submission includes:

- Responses to comments made by Division in its letter to the sponsor dated August 3, 2010
- Protocol for study LA36-0310
- Statistical analysis plan for study LA36-010
- Six literature reference

The primary objective:
To evaluate the efficacy of oral administration of deferiprone in the treatment of iron overload in patients in whom standard chelation has failed. Chelation failure is defined as iron accumulation above a boundary level, defined by high serum ferritin or LIC, or low cardiac MRI T2* levels. Success rate will be evaluated by compiling data from clinical trials that recorded one or more of the following effects during treatment with deferiprone for up to 1 year, and where data on serum ferritin, LIC, or cardiac MRIT2* are available prior to and after starting therapy with deferiprone. To allow assessment of patients whose efficacy assessments were conducted at about the 1 year anniversary of initiation of treatment, the up to 1 year window will include data obtained within ±3 months of the anniversary date.

Study Design:
This is a retrospective analysis of data derived from multiple studies that were conducted to support NDA 21-825, some controlled and some uncontrolled, in patients who mostly had an underlying diagnosis of beta thalassemia and transfusional iron overload. Patients to be enrolled into the current study will be programmatically selected from database submitted to the Agency as part of the NDA and integrated up to 11 May 2010 for ongoing clinical studies. The sponsor will subsequently assess the serum ferritin, LIC and cardiac MRI T2* values captured during treatment with deferiprione for up to 1 year for analysis of its efficacy in the cohort of patients selected by Clinical Data Management. The number of deferiprone-treated patients who meet the defined criteria for successful treatment outcomes will determined and a success rate will be calculated. The lists of studies are summarized in the Appendix 1.

Study Population:
Study population will be composed of patients who show evidence of failed responses to previous chelation therapy.

Efficacy Endpoints:
The primary efficacy endpoint: will be the change in serum ferritin concentration from baseline within one year of deferiprone therapy. The efficacy of deferiprone in each patient will be established by determining whether treatment with deferiprone for up to 1 year (including therapy interruptions and patients who have discontinued therapy during that period due to adverse events or other reasons) has succeeded or failed in promoting no less than 20% decrease in serum ferritin

concentration in a clinically acceptable proportion of patients. The current study will conclude that deferiprone is effective if a successful outcome is observed in at least 20% of patients who failed previous chelation therapy.

Secondary endpoints of: will be the changes in LIC and in cardiac MRI T2*.

Statistical Methods:
Analysis Population:
Two analysis populations will be used in the study:

- **Intent-to-treat (ITT) population** will be the primary population for this study. For each efficacy measure, the ITT population will be composed of those patients who had taken at least one dose of deferiprone and had at least one post-baseline measurement of that efficacy measure
- **Per Protocol population** will be the secondary population for this study. For each efficacy, the PP population will be composed of those patients who had completed the study that they originated from or at least one year of deferiprone therapy for long-term studies, and had no missing data for the end-of-study measurement or the last scheduled measurement at the end of the first year, respectively, for that efficacy measure.

Sample Size Estimation:
There is no sample size calculation in this study. The sponsor intends to have data from as many eligible patients as possible to determine the efficacy of the drug. The following table proposed by sponsor shows the sample size required to meet the criterion that the lower limit of the 95% confidence interval for the success rate is >20% for different expected success rates.

Sample size calculation for which the lower limit of 95% CI is at least 20%

Sample Size	Expected Success Rate	95% CI
1,648	22%	(20%, 24%)
438	24%	(20%, 28%)
206	26%	(20%, 32%)
122	28%	(20%, 36%)
81	30%	(20%, 40%)
59	32%	(20%, 44%)
44	34%	(20%, 48%)
40	36%	(20%, 52%)

Analysis of efficacy variables:
All statistical tests will be two-sided and p-value≤0.05 will be used for the determination of statistical significance in the statistical tests.

The primary efficacy endpoint will be based on the change in serum ferritin concentration from baseline to the end-of-study data or the last observation within 1 year of deferiprone therapy for studies of >1 year in duration. For deferiprone to be a successful treatment in patients who failed standard chelation therapy, there must be at least 20% of patients who demonstrate a favorable response on serum ferritin. The success rate will be calculated. A patient will be determined to be a successful responder if the patient's serum ferritin was reduced by at least 20% within 1 year of deferiprone treatment. The success rate by study and overall success rate and its 95% C.I. will be calculated based on normal approximation.

Secondary efficacy endpoints will be based on the changes in cardiac MRI T2* and LIC from baseline to the end-of-study data or the last observation within 1 year of deferiprone therapy for studies of >1 year in duration. For deferiprone to be a successful treatment in patients who failed standard chelation therapy based on cardiac MRI T2*, there must be at least 20% of patients who demonstrate a favorable response on MRI T2*. The success rate will be calculated. A patient will be determined to be a successful responder if the patient's MRI T2* was increased by at least 20% 1 year of deferiprone treatment. The success rate by study and overall success rate and its 95% C.I. will be calculated based on normal approximation. For deferiprone to be a successful treatment in patients who failed standard chelation therapy based on LIC, there must be at least 20% of patients who demonstrate a favorable response on LIC.The proposed ≥20% reduction in serum ferritin or LIC is based on the understanding that it would mean a drop of a minimum of 500 µg/L or 1.4 mg Fe/g dw, respectively.

Overall Treatment Success
If the lower limit of the 95% confidence interval for any efficacy measure is greater than the pre-defined criterion of treatment success (20%), the therapy will be considered to be a success for that particular measure. The final assessment of treatment success for the purpose of this study will be based on the success rate for serum ferritin, the primary efficacy measure.

Missing Data Imputation
The sponsor proposed to use LOCF method for the missing data imputation to the efficacy data.

Subgroup Analyses
The sponsor will calculate success rates for the subgroup of patients with 2 or more serum deferiprone values, of which a majority of the values were greater than 2,500 µg/L, before starting deferiprone in order to determine if those patients exhibited the same degree of response to deferiprone treatment as did the patients who had failed on standard chelation therapy, but had only a single baseline ferritin concentration to designate them as a failed patient.

Sensitivity Analysis:
The sponsor proposed sensitivity analyses which include:

- Primary efficacy analysis using Pre Protocol population.
- LOCF method for imputation of missing efficacy endpoint data.

Statistical Conclusion and Recommendations:
In the statistical analysis plan (SAP), the sponsor has responded to the issues raised by the Agency in regard to the missing data imputation using LOCF, and selection bias from this retrospective study. The sponsor provided justification of LOCF imputation and justification of selection bias in the statistical analysis plan.

LOCF Missing Data Imputation
The sponsor believes that there is no subjective assessment involved by using LOCF missing data imputation for the efficacy data and the approach is easy to understand. They also believe, with the proposed treatment success criteria (≥20% improvement in the efficacy endpoint) being applied to patients who had shown a poor response to previous chelation therapy, these patients would be considered treatment failure regardless of the LOCF data, thereby not favouring deferiprion.

Justification of Patient Selection Bias
The sponsor believes that patient selection bias may occur in a situation where precise inclusion and exclusion criteria have not been defined a priori before the study. In this study, the inclusion and exclusion criteria are clearly defined in the protocol prior to the patient selection process.

Some additional statistical comments:

1. The precise inclusion and exclusion criteria could control the selection bias by a certain level, but can not eliminate the selection bias. To minimize such bias, we suggest using an Independent Committee to select subjects from the entire integrated dataset following approved inclusion and exclusion criteria by the Agency.

2. The subgroup analysis should also include age, gender, US vs Non-Us and some other important baseline characteristics. The consistency or lack of consistently of the results should be evaluated across subgroup.

3. This evaluation involves pooling data from many studies of different designs; different durations of drug treatment; different time intervals of patient visits; different inclusion/exclusion criteria; and different patient characteristics. These different should be addressed.

4. Provide justification on the clinical meaningfulness of the 20% success rate

5. We strongly suggest the sponsor to conduct prospective, randomized, controlled clinical study with a clear hypothesis, which shows a highly significant effect.

6. We agree with treating those subjects who are not known to have had a decrease in serum ferritin of at least 20% as a non-responder.

Appendix 1
List of studies

Study	Title	Number of Subjects:	Efficacy Measurements		
		Entered/Completed	Serum Ferritin	LIC	MRI T2*
LA-01	Randomized Trial of L1 and Deferoxamine in Thalassemia Major	DFP: 35/27 DFO: 36/30	√	√	
LA-02/06	Trial of APO-66 (Deferiprone) in Thalassemia/Maintenance Study Protocol of Deferrum™ (Ferriprox) for Subjects with Thalassemia who Complete Apotex Protocol LA-02	DFP: 187/70	√		
LA-03	The Long Term Efficacy and Safety of Deferiprone in Subjects with Thalassemia	DFP: 29/25	√	√	
LA-04/06B	Compassionate-use of Deferiprone (L1) in Subjects with Iron Overload	DFP:168/23	√	√	√
LA-08-9701	Safety and Efficacy of Alternating Deferoxamine and Deferiprone compared with Deferoxamine Alone in the Treatment of Iron Overload in Thalassemia Subjects	Alternating DFP and DFO: 29/29 DFO: 30/30	√	√	
LA-11	Efficacy and Safety of Deferiprone (L1) in β-Thalassemia/Hemoglobin E Diseases Subjects in Thailand	DFP: 24/16	√		

LA-12-9907	Retrospective Assessment of Heart Failure and Survival During Iron Chelation with Deferiprone or deferoxamine in Subjects with Transfusion-Dependent B-Thalassemia	DFP: 54 DFO: 75 The study assessed existing data so study completion or withdrawal was not defined.	√	√	
LA-15-0002	Safety and Efficacy of Ferriprox™ for the Treatment of Iron Overload in Subjects with Transfusion-dependent Thalassemia in Iran	DFP: 29/26	√		
LA-16-0102	Randomized Trial Comparing the Relative Efficacy of Deferiprone to that of Deferoxamine in Removing Excess Cardiac Iron in Thalassemia Major Subjects	DFP: 29/27 DFO: 32/29	√	√	√
Borgna-Pignatti et al. (2006)	Cardiac Morbidity and Mortality in Deferoxamine- or Deferiprone-treated Patients with Thalassemia Major	DFP: 157 DFO: 359	√		
LA-28-CMP	The compassionate use/named patient program of Ferriprox oral solution in iron-overloaded pediatric patients with transfusion-dependent anemias	DFP: 83/62 (72 subjects previously enrolled in LA-30)	√		
LA-30	24-week open label, uncontrolled study of the safety and efficacy of Ferriprox (deferiprone) oral solution in iron-overloaded pediatric patients with transfusion-dependent anemia	DFP: 100/95	√		

Data cut off: 11 MAY 2010

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

/s/

QING XU
10/21/2010

MARK D ROTHMANN
10/22/2010

Statistical Review and Evaluation

NDA#21825 **Date Received**:July 12 2009
Product Name:Deferiprone **Indication**:Iron Chelator
Statistical Reviewer: Qing Xu **Medical reviewer**: George Shashaty
Sponsor:ApoPharma, Inc **Meeting Schedule**:
Protocol Number and Title: LA36-0310
A clinical trial to evaluate the efficacy of deferiprone in patients with iron overload whom previous chelation therapy has been inadequate by analysis of data from clinical studies of deferiprone

Study Background:
The NDA was submitted to the Agency on January 29, 2009. The submission of the NDA included a single randomized controlled trial. The primary endpoint in that trial was the change in cardiac iron as measured by cardiac MRI T2* assessment after one year of treatment with deferiprone. The comparator drug was deferoxamine, which at the time of the study was the only approved drug for the indication. The Agency was concerned that the primary endpoint measured was not a validated surrogate for clinical utility. At the conclusion of the review, the Agency sent a Complete Response (CR) letter to the sponsor describing the deficiencies of the data submitted. In the CR letter dated November 30, 2009, the Agency recommended that the sponsor perform "adequate and well-controlled" trials with deferiprone to support the application for approval of the NDA.

This submission is the sponsor's response to the discussion. The sponsor intends to review already available data on serum ferritin, cardiac MRI T2* and LIC as measures of body iron burden before and after 1 year of treatment with deferiprone in patients who could not be adequately treated with other approved chelator therapy. The source of the data will be from the studies previously submitted in the NDA with the addition of data from completed studies of a liquid formulation of deferiprone and a continuing compassionate use program.

Primary Objective:
To evaluate the efficacy of oral administration of deferiprone in the treatment of iron overload in patients in whom standard chelation has failed.

Primary Endpoints: Change from baseline in serum ferritin

Secondary Endpoints:

- Change from baseline in cardiac MRIT2*
- Change from baseline in liver iron concentrations

Efficacy Analysis:
The primary endpoint will be based on changes in serum ferritin. For Ferriprox to be a successful in patients who failed standard chelation therapy there must be at least 20% of patients who demonstrate a favorable response. LOCF method will be used for data imputation.

Statistical Issues and Comments to be conveyed to the Sponsor:
The draft protocol contains only limited statistical information. Please submit a protocol with detailed statistical analysis plan which includes detailed sample size justification and detailed subgroup analysis plans. Please also provide detailed strategy and justifications for reviewing data from multiple retrospective studies to reduce bias.

Missing data should be kept to a minimum. Missing data for the primary analysis should be addressed. Please provide a justification for your choice of LOCF imputation or any other intended method of imputation. Sensitivity analyses should evaluate the limitations of the data.

Qing Xu, Ph. D.
Mathematical Statistician

Concur: Dr. Rothmann
Cc:
HFD-107/ Dr. Lee
HFD-107/Dr.George
HFD-107/Dr. Robie Suh
HFD-107/Dr. Kamiskas
HFD-107/Dr. Farrell
HFD-711/Dr. Rothmann
HFD-711/Dr. Sridhara
HFD-700/Ms. Patrician

This review consists of 2 pages of text
Review.doc

Application Type/Number	Submission Type/Number	Submitter Name	Product Name
NDA-21825	ORIG-1	APOPHARMA INC	FERRIPROX (DEFERIPRONE)

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

/s/

QING XU
08/02/2010

MARK D ROTHMANN
08/03/2010

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Science (OTS)
Office of Biostatistics

SECONDARY STATISTICAL REVIEW

CLINICAL STUDIES

NDA/Serial Number:	21-825 SE000
Drug Name:	Deferiprone (Ferriprox) Oral film coated tablets (500mg)
Indication(s):	Treatment of iron overload in patients with transfusion-dependent thalassemia and for whom the use of other iron chelators has been considered inappropriate.
Applicant:	ApoPharma Inc
Date(s):	Date of Submission:
	Letter Date January 29, 2009
	Stamp Date February 2, 2009
	PDUFA Goal Date November 30, 2009
Review Priority:	Standard
Biometrics Division:	Division of Biometrics V
Secondary Statistical Reviewer:	Jyoti Zalkikar, Ph. D., Statistical Team Leader
Concurring Reviewers:	Rajeshwari Sridhara, Ph.D., Acting Director, DBV
Medical Division:	Division of Medical Imaging and Hematology Products
Clinical Team:	George Shashaty, M.D., Clinical Reviewer
Project Manager:	Hyon-Zu Lee

Keywords: Confidence Intervals, Open Label, Randomized

Brief Overview of Clinical program for Deferiprone

Deferiprone (Ferriprox) Oral film coated tablets (500mg) is an orally active iron chelator. The proposed indications are (1) the treatment of iron overload in patients with transfusion-dependent thalassemia, and (2) the treatment of iron overload in patients with other transfusion-dependent anemia for whom the use of other iron chelators has been considered inappropriate.

The efficacy and safety of Deferiprone in this New Drug Application (NDA) was based on one main phase 3 (LA16-0102) and one supportive (retrospective) study (LA12-9907) study. These two efficacy studies in this NDA were not conducted under US IND and there was no FDA review of protocols, study designs and statistical analysis during the planning stage of this clinical program. The results from these studies were reviewed in detail in the primary statistical review and are subject of this secondary review.

This NDA submission included several other supportive and safety studies based on publications that were not reviewed in this as well as the primary statistical review. These include trials/studies that varied in nature, with most being single arm, non-comparative, retrospective, investigator initiated, safety, extension, compassionate use or registry studies, and published literature related to deferiprone.

Study LA16-0102:

Design: This was a two arm comparative, multicenter, randomized, open-label, deferoxamine controlled clinical trial. The patient population consisted of patients with Thalassemia major previously treated with deferoxamine. The study used 1:1 randomization to either continue deferoxamine or switch to deferiprone.

The primary endpoint measured was the change in MRI T2* from baseline at Month 12. This was a superiority study designed to show that mean change in MRI T2* from baseline at month 12 is greater for deferiprone than that for deferoxamine. The sample size calculation was based on the analysis method of two-sample t-test with 5% level of significance and 80% power, assuming that MRI T2* values were normally distributed and there was true mean difference of 2.3 ms between deferiprone and deferoxamine with a standard deviation of 2.5 ms . There was no justification in the study report for these assumptions. Non-parametric methods (that do not assume normality) were prospectively planned as alternative analysis methods in the case of violation of normality assumption by the company in their statistical analysis plan (SAP).

Results: MRI T2* change from baseline at month 12 (mean (milliseconds) ± SD) for the Study LA16-0102 was 3.9±3.6 for deferiprone (n = 29), and 2.3±4.1 for deferoxamine (n=31) (p=0.093). The protocol defined endpoint was not met for this trial. This analysis was based on the assumption of normal distribution for the primary endpoint. Further analyses revealed that these data were highly skewed and not normally distributed. The Shapiro-Wilk test of normality gave a p-value < 0.001 providing statistically significant evidence of departure from normality for these data. Therefore the use of two-sample t- test analysis method lacked power and compromised the strength of evidence in the data to reject null hypothesis of no difference between deferiprone and deferoxamine.

The company noted this situation and instead of following pre-specified SAP, they used log transformation on MRI T2* in order to "linearize the scale, potentially resulting in normalization of the MRI T2* data". This post-hoc analysis based on log transformation showed a statistically significant difference in favor of Deferiprone (p=0.0228). In addition to being post-hoc and not confirmatory, these

results are subject to the same limitations as the raw MRI T2* data due to lack of normality (p-value < 0.001 using Shapiro-Wilk test of normality).

The primary reviewer at the FDA followed the company's SAP and used non-parametric methods to evaluate these data which were prospectively planned in the case of violation of normality. These methods are not based on the assumption of normality and provide valid analyses of these data. The median difference from baseline at 12 months was 3.7 for deferiprone (n = 29), and 1.0 for deferoxamine (n=31). The primary reviewer used three appropriate non-parametric methods. The results for all three tests showed a statistically significant difference in favor of deferiprone in terms of MRI T2*(median test p-value=0.0048, Wilcoxon Rank Sum Test p-value = 0.0124 and Kolmogorov-Smirnov test asymptotic p-value= 0.0168 (12 months) indicating rejection of the null hypothesis that the distributions are identical for the two groups).

Analyses of the secondary endpoints, however, do not consistently support superiority of deferiprone in these patients (nominal p-values are not significant) compared to deferoxamine. Additionally, there is no evidence that small differences in the secondary endpoints (change in MRI T2* at 6 months, CMR LVEF, ECHO LVEF and LVSF, serum ferritin concentrations and liver iron concentrations) in patients treated with deferiprone and deferoxamine translate into a clinically meaningful benefit on either mortality or important morbidity.

Reviewer's Comments: Although the data from this study provided statistically significant evidence that treatment with deferiprone leads to bigger change in MRI T2* at 12 months compared to deferoxamine, it is not clear as to how this result translates into clinically important efficacy benefit. This study was not designed to and therefore does not provide evidence that change in MRI T2* is reasonably likely to predict clinical benefit due to lack of long-term follow-up of these patients. In the absence of convincing data showing the validity of MRI T2* as a surrogate for clinical outcome, the usefulness of this study for the evaluation of efficacy of deferiprone is questionable.

Study (LA12-9907):

Design: This was a non-randomized, observational study based on retrospective assessment on medical records from one center in Italy. The patient population consisted of subjects with transfusion dependent ß-thalassemia. The amended protocol for this study included all patients with less than three serum ferritin concentration determinations in last 2 years. The external comparator used was deferoxamine. Note that although comparative, this study does not benefit from randomization and therefore balance between the two treatment arms in terms of unmeasured baseline variables is questionable. This leads to potential for introduction of bias that can not be resolved. This is a serious limitation of this study.

The primary endpoint was incidence of cardiac disease using NYHA classification and physicians assessment of CHF, LVEF, and LVSF during the study period of approximately 5 years. Primary statistical analysis was to compare the long-term (approximately 5 years) efficacy of deferiprone (DFP) to that of deferoxamine (DFO). Changes (worsening, no change, or improvement) in cardiac disease were based on cardiac status using NYHA class. Although this endpoint appears objective, it is based on criteria that may not be appropriate indicators of cardiac disease.

Results:

Out of 168 patients screened, 129 patients (54 deferiprone, 75 deferoxamine) qualified. Age at the baseline and age at the start of the chelation therapy were significantly different in two groups for the main population. At baseline, patients receiving deferiprone were significantly younger and had started

chelation therapy earlier than patients in the deferoxamine group. Thus the two arms were not comparable for the main population.

Age-matched population was generated by matching the subjects in the two arms for age at the start of chelation therapy. This population had 94 patients (47 deferiprone, 47 deferoxamine), and was balanced with respect to most measured baseline variables.

Assessment of cardiac disease using NYHA classification for the age-matched population with n=47 in each treatment group was: cardiac disease- last assessment (baseline) 14.9% versus 23.4% (p=0.4323), worsening of NYHA from first to last assessment 4.4% versus 19.2% (p=0.0502), with cardiac disease during the study among patients who were initially cardiac disease-free 5.0% versus 19.1% (p=0.0888) for the deferiprone versus deferoxamine therapy respectively. All these nominal p-values (based on two-sided (conservative) Fisher's exact test) were not significant at 5% level.

Reviewer's Comments: This study has several serious limitations including lack of randomization, no information regarding some important baseline variables such as splenectomy status, limited information (a lot of missing values) at baseline on variables such as hepatic iron concentration, questionable assessment of cardiac disease and comparisons that are not statistically significant.

ICH E10 guidance recognizes the drawbacks of observational (with external control) trials and asks to use caution when using p-values. The guidance calls for several considerations when using observational studies and dramatic treatment difference with highly significant p-value should be seen. Further the guidance states that the external controls may be accepted when (i) usual course of the disease is highly predictable, (ii) endpoints are objective, (iii) impact of baseline variables on the endpoints is well characterized (iv) much more extreme levels of statistical significance are observed. It is clear that all these criteria are NOT met in the current setting. Therefore, this observational study which is subject to several limitations does not provide independent corroboration of efficacy of deferiprone.

Conclusions and Recommendations

Deferiprone (Ferriprox) Oral film coated tablets (500mg) is an orally active iron chelator developed for the treatment of iron overload in patients with transfusion-dependent thalassemia.

The proposed indication is the treatment of iron overload in patients with transfusion-dependent thalassemia, and for the treatment of iron overload associated with other transfusion-dependent anemia in patients for whom the use of other iron chelators has been considered inappropriate.

The efficacy and safety of Deferiprone in this New Drug Application (NDA) was based on one main phase 3 (LA16-0102) and one supportive (retrospective) study (LA12-9907) study.

The single randomized trial LA16-0102 has serious limitations including imaging endpoint of MRI T2* of questionable clinical meaningfulness, no "within study" evidence that MRI T2* is reasonably likely to predict clinical outcome, lack of consistent demonstration benefit based on secondary endpoints and inadequate safety database of only 29 patients exposed to deferiprone. The observational study LA 12-9907 did not provide independent corroboration due to serious drawbacks such as lack of randomization, and missing of limited information regarding some important baseline variables such as splenectomy status, and hepatic iron concentration.

The submitted data does not provide robust and meaningful statistical evidence to support the efficacy of deferiprone for the proposed indications.

Application Type/Number	Submission Type/Number	Submitter Name	Product Name
NDA-21825	ORIG-1	APOPHARMA INC	FERRIPROX (DEFERIPRONE)

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

/s/

JYOTI ZALKIKAR
11/20/2009

RAJESHWARI SRIDHARA
11/22/2009

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Science (OTS)
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/Serial Number:	21-825 SE000
Drug Name:	Deferiprone (Ferriprox) Oral film coated tablets (500mg)
Indication(s):	Treatment of iron overload in patients with transfusion-dependent thalassemia and for whom the use of other iron chelators has been considered inappropriate.
Applicant:	ApoPharma Inc
Date(s):	Date of Submission:
	Letter Date January 29, 2009
	Stamp Date February 2, 2009
	PDUFA Goal Date November 30, 2009
Review Priority:	Standard
Biometrics Division:	Division of Biometrics V
Statistical Reviewer:	Satish C. Misra, Ph. D.
Concurring Reviewers:	Jyoti Zalkikar, Ph. D., Statistical Team Leader
	Rajeshwari Sridhara, Ph.D., Acting Director, DBV
Medical Division:	Division of Medical Imaging and Hematology Products
Clinical Team:	George Shashaty, M.D., Clinical Reviewer
Project Manager:	Hyon-Zu Lee

Keywords: Confidence Intervals, Open Label, Randomized

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1. EXECUTIVE SUMMARY

1.1 Conclusions and Recommendations

Deferiprone (Ferriprox) Oral film coated tablets (500mg) is an orally active iron chelator developed for the treatment of iron overload in patients with transfusion-dependent thalassemia.

The proposed indication is the treatment of iron overload in patients with transfusion-dependent thalassemia, and for the treatment of iron overload associated with other transfusion-dependent anemias for whom the use of other iron chelators has been considered inappropriate.

The efficacy and safety of Deferiprone in this New Drug Application (NDA) was based on one main phase 3 (LA16-0102) and one supportive (retrospective) study (LA12-9907) study. This submission included several other supportive and safety studies based on publications.

The primary results of the single randomized trials LA16-0102 are summarized in the Table 1 below:

Table 1: Change in MRI T2* at Month 12 from Baseline - ITT Population

MRI T2* (milliseconds)	Randomized Treatment Groups -- Difference from Baseline at 12 Months		
	Deferiprone (n=29)		Deferoxamine (n=31†)
Analysis - Protocol & SAP identified[1]			
Mean (milliseconds) ± SD	3.9 ± 3.6		2.3 ± 4.1
95% CI	(2.6, 5.3)		(0.8, 3.9)
p-value (unequal variance)		0.0993	
Sponsor's post-hoc analysis-based on log transformation [2]			
Diff (log 12 (or 6) – log base)	0.236703		0.122555
Nominal p-value		0.0228	
Analysis– Non-parametric tests[3]			
Median	3.7		1.0
(10^{th} percentile, 90^{th} percentile)	(-0.1, 8.7		(-1.2, 6.8)
Range (Minimum, Maximum)	(-2.0, 12.7)		(-0.1, 8.7)
p-value – Median Test		0.0048	
p-value -Wilcoxon (Rank Sum Test)		0.0124	
p-value - Kolmogrov-Smirnov Test		0.0168	

[1] The MRI T2* between the Deferiprone and Deferoxamine treatment groups was compared by the two sample t-test assuming unequal variances. The assumptions of normality behind these tests are violated

[2] The Log (MRI T2*) between the Deferiprone and Deferoxamine treatment groups was compared by the two sample t-test by the sponsor.

[3] Since the data are non-normal, non-parametric statistics (Wilcoxon/Kruskal-Wallis Rank Sum tests and the Median Test (Number of Points Above Median), Kolmogorov-Smirnov test) provide valid analyses. The asymptotic p-value for the Kolmogorov-Smirnov test is 0.0168 (12 months). This indicates rejection of the null hypothesis that the distributions are identical for the two groups.

† Subject C1-40 had baseline MRI T2* level value only and was not eligible to be included in the ITT population.

This single randomized trial LA16-0102 has serious limitations including imaging endpoint of MRI T2*, no "within study" evidence that MRI T2* is reasonably likely to predict meaningful clinical outcome, "primary" analysis questionable due to lack of normality, inadequate safety database of only 29 patients

exposed to deferiprone, and the observational study LA 12-9907 not providing independent corroboration due to serious limitations including lack of randomization, no information regarding some important baseline variables such as splenectomy status, and limited information (a lot of missing values) at baseline such as hepatic iron concentration. The submitted data does not support the proposed indications.

1.2 Brief Overview of Clinical Studies

Deferiprone (Ferriprox) Oral film coated tablets (500mg) is an orally active iron chelator. The proposed indications are (1) the treatment of iron overload in patients with transfusion-dependent thalassemia, and (2) the treatment of iron overload in patients with other transfusion-dependent anemias for whom the use of other iron chelators has been considered inappropriate.

There were two efficacy studies in this NDA.. These two efficacy studies were was not conducted under US IND and there was no FDA review of protocols, study designs and statistical analysis plans of the two pivotal studies of this oral iron chelator.

The first efficacy study (LA16-0102) is a two arm comparative, multicenter, randomized, open-label, active controlled clinical trial comparing the relative efficacy of deferiprone (n=29) to that of deferoxamine (n=31) in removing excess cardiac iron in thalassemia major patients. The patient population in the this study consisted of patients with Thalassemia major previously treated with the comparator drug, deferoxamine. The study used 1:1 randomization to either continue deferoxamine or switch to deferiprone.

The inclusion criteria included diagnosis of thalassemia major with transfusion dependency, maintenance of a mean pre-transfusion hemoglobin (Hgb) of ≥9 g/dL, age between 18 and 36 years, receipt of deferoxamine for at least the last 5 years (if subject had previously received deferiprone, he/she must not have received it for at least the previous 2 years), MRI T2* ≥8 milliseconds and <20 milliseconds, use of adequate contraception and not breastfeeding, and written informed consent. The exclusion criteria included anemia other than thalassemia, HIV positive, evidence of cardiomyopathy as demonstrated by a diminished left ventricular ejection fraction (LVEF) or a diminished left ventricular shortening fraction (LVSF), presence of a significant arrhythmia or treatment for same, previous discontinuation of deferoxamine or deferiprone because of an adverse reaction to either chelator, abnormal liver function tests (>3x ULN), disorders associated with neutropenia or thrombocytopenia in the previous 12 months, use of other investigational products, presence of medical conditions that make it unwise to enter the trial, pregnancy or breastfeeding, metallic objects in the body, and history of malignancy.

Patients randomized to deferiprone arm were to receive oral deferiprone at a dose of 33.3 mg/kg three times daily (total daily dose of 100 mg/kg). Therapy was to be initiated at a dose of 75 mg/kg/d in three divided doses with a minimal interval of 4 hours between doses. At week 4, the daily dose was to be escalated to 85 mg/kg/d; then, at week 8, the dose was to be escalated to 100 mg/kg/d. Lower doses could be prescribed in the event of the development of adverse reactions. The sponsor selected the 100 mg/kg/d dose because it is the maximum recommended dose where deferiprone has been approved. The sponsor has not performed any formal dose-response studies.

Patients randomized to deferoxamine arm were to receive subcutaneous infusions of deferoxamine (manufactured by Novartis) over an interval of up to 12 hours duration, 5–7 days per week at a dose of 50 mg/kg body weight. A lower dose could be prescribed in the event of the development of adverse reactions. The dose used is the maximal recommended dose of the drug.

The primary endpoint measured was the change in MRI T2* from baseline at Month 12. Note that the patients in this study were not followed for clinical outcomes and therefore this study was not designed to obtain internal validation of MRI T2* change as a surrogate for any clinical outcome indicative of reduced cardiac iron. The secondary endpoint included Change in MRI T2* at Month 6 from Baseline, Measures of cardiac function, such as Left Ventricular Ejection Fraction (LVEF) assessed by both Cardiovascular Magnetic Resonance (CMR) and echocardiogram (ECHO), Left Ventricular Shortening Fraction (LVSF) assessed by echocardiogram.

This was a superiority study designed to show that mean change in MRI T2* from baseline at month 12 is greater for deferiprone than that for deferoxamine.

The second efficacy study (LA12-9907) is non-randomized, observational study based on retrospective assessment on medical records from one center in Italy. The patient population consisted of subjects with transfusion dependent ß-thalassemia and the external comparator used was deferoxamine. The primary endpoint was incidence of cardiac disease using NYHA classification and physicians assessment of CHF, LVEF, and LVSF during the study period of approximately 5 years. Primary statistical analysis was to compare the long-term (approximately 5 years) efficacy of Deferiprone (DFP) to that of Deferoxamine (DFO). Changes (worsening, no change, or improvement) in cardiac disease were based on cardiac status using NYHA class. There were 54 subjects in the deferiprone arm and 75 subjects in the deferoxamine arm.

These two identified pivotal trials are reviewed in detail in this report. Studies not reviewed in this report include twelve trials/studies that varied in nature, with most being single arm, non-comparative, retrospective, investigator initiated, safety, extension, compassionate use or registry studies, and published literature related to deferiprone.

1.3 Statistical Issues and Findings

This reviewer evaluated the evidence in support of the efficacy of Deferiprone (Ferriprox) Oral film coated tablets (500mg) from the results of two key pivotal trials. Study LA16-0102 was a superiority study and designed to show that mean change in MRI T2* from baseline at month 12 is greater for deferiprone than that for deferoxamine.

The sample size calculation was based on the analysis method of two-sample t-test with 5% level of significance and 80% power, assuming that MRI T2* values were normally distributed and there was true mean difference of 2.3 ms between deferiprone and deferoxamine with a standard deviation of 2.5 ms . There was no justification in the study report for these assumptions.

Also, in case of assumption violation, alternative analysis methods were not clearly outlined in the prospective statistical analysis plan. The violation of normality assumption is of concern due to small sample size of 30 patients in each arm and its potential impact on the interpretation of outcome.

MRI T2* change from baseline at month 12 (mean (milliseconds) ± SD) for the Study LA16-0102 was 3.9±3.6 for deferiprone (n = 29), and 2.3±4.1 for deferoxamine (n=31) (p=0.093). The protocol defined endpoint was not met for this trial. This analysis was based on the assumption of normal distribution for the primary endpoint. Further analyses revealed that these data were highly skewed and not normally distributed. The Shapiro-Wilk test of normality gave a p-value < 0.001 providing statistically significant evidence of departure from normality for these data. Therefore the use of prospective two-sample t- test analysis method lacked power.

The company noted this situation and used log transformation on MRI T2* in order to "linearize the scale, potentially resulting in normalization of the MRI T2* data". The company's post-hoc analysis based on log transformation showed a statistically significant difference in favor of Deferiprone (p=0.0228). These results are also subject to the same limitations as the raw MRI T2* data as FDA's analysis shows lack of normality for log transformed data as well (p-value < 0.001 using Shapiro-Wilk test of normality).

This reviewer used non-parametric methods to evaluate these data which were prospectively planned in the case of violation of normality assumption by the company in their statistical analysis plan. These methods are not based on the assumption of normality and provide valid analyses of these data. The results for both tests showed a statistically significant difference in favor of deferiprone in terms of MRI T2*. The median difference from baseline at 12 months was 3.7 for deferiprone (n = 29), and 1.0 for deferoxamine (n=31) (median test p-value=0.0048, and Wilcoxon Rank Sum Test p-value = 0.0124).

The primary endpoint analyzed in the trial (change in MRI T2*) cannot yet be considered to be a clinically important benefit, nor is it established as a surrogate endpoint for the treatment of transfusional hemosiderosis in thalassemia patients. This study does not provide evidence that MRI T2* is reasonably likely to predict clinical outcome due to lack of long-term follow-up of these patients. In case of absence of convincing data showing the validity of MRI T2* as a surrogate for clinical outcome, the usefulness of this study to the evaluation of efficacy of deferiprone is questionable.

Analyses of the secondary endpoints do not add additional support for the efficacy of the use of deferiprone in these patients. There is no evidence that the differences in secondary endpoints (MRI T2*, CMR LVEF, ECHO LVEF and LVSF) in patients treated with deferiprone translate into a clinically meaningful benefit on either mortality or important morbidity.

Study LA12-9907 is an observational study based on the retrospective assessment of medical records at a single center in Italy. The patient population consisted of subjects with transfusion dependent ß-thalassemia and the external comparator used was deferoxamine. Note that although comparative, this study has a serious limitation of lack of randomization. This study does not benefit from randomization and therefore balance between the two treatment arms in terms of unmeasured baseline variables is questionable. This leads to potential for introduction of bias that can not be resolved. This is a serious limitation of this study.

The primary endpoint was incidence of cardiac disease during the study using NYHA classification and physicians assessment of CHF, LVEF, and LVSF. Changes (worsening, no change, or improvement) in cardiac disease were based on cardiac status using NYHA class. Although this endpoint appears objective, it is based on criteria that may not be appropriate indicators of cardiac disease.

The amended protocol included all patients with less than three serum ferritin concentration determinations in last 2 yrs . Out of 168 patients screened, 129 patients (54 deferiprone, 75 deferoxamine) qualified. Age at the baseline and age at the start of the chelation therapy were significantly different in two groups for the main population. At baseline, patients receiving deferiprone were significantly younger and had started chelation therapy earlier than patients in the deferoxamine group. Thus the two arms were not comparable for the main population resulting from the amended protocol.

Age-matched population was generated by matching the subjects in the two arms for Age at the start of chelation therapy. This population had 94 patients (47 deferiprone, 47 deferoxamine), and was balanced with respect to most measured baseline variables.

Assessment of cardiac disease using NYHA classification for the age-matched group with n=47 in each group was: cardiac disease- last assessment (baseline) 14.9% versus 23.4% (p=0.4323), worsening of NYHA from first to last assessment 4.4% versus 19.2% (p=0.0502), with cardiac disease during the study among patients who were initially cardiac disease-free 5.0% versus 19.1% (p=0.0888) for the deferiprone versus deferoxamine therapy respectively. All these nominal p-values (based on two-sided (conservative) Fisher's exact test) were not significant even at 5% level.

This study has several serious limitations including lack of randomization, amendment to protocol, no information regarding some important baseline variables such as splenectomy status, limited information (a lot of missing values) at baseline such as hepatic iron concentration.

ICH E10 guidance recognizes the drawbacks of observational (with external control) trials and asks to use caution when using p-values. The guidance calls for several considerations when using observational studies and dramatic treatment difference with highly significant p-value should be seen. Further the guidance states that the external controls may be accepted when (i) usual course of the disease is highly predictable, (ii) endpoints are objective, (iii) impact of baseline variables on the endpoints is well characterized (iv) much more extreme levels of statistical significance are observed. It is not clear that all these criteria are met in the current setting.

Therefore, this observational study is subject to serious limitations and does not provide independent corroboration.

2. INTRODUCTION

2.1 Overview

Deferiprone is a small bidentate, 3-hydroxypyridin-4-one iron chelating agent. Its low molecular weight, its neutral charge, and its lipid solubility favor access of deferiprone to intracellular iron pools. The amount of iron eliminated from the body has been associated with the dose of Deferiprone.

Deferiprone (Ferriprox) is an orally administered drug proposed for marketing with the following specific indication:

"for the treatment of iron overload in patients with transfusion-dependent thalassemia. Deferiprone is also indicated for the treatment of iron overload in patients with other transfusion-dependent anemias for whom the use of other iron chelators has been considered inappropriate."

The recommended dosage are 25 to 33 mg/kg body weight, orally, three times a day, for a total daily dose of 75 to (b) (4) mg/kg body weight. The drug is supplied in the form of 500 mg film-coated tablets.

Deferiprone was initially developed by independent clinical investigators and first in human use was in 1987 (UK). Ciba-Geigy was the initial commercial developer. ApoPharma assumed development in 1993. It was approved in EU in 1999, and currently in 60 EU countries. The approved indication in EU is “for the treatment of iron overload in patients with thalassaemia major when deferoxamine therapy is contraindicated or inadequate” Deferoxamine (Desferal) and Deferasirox (Exjade) are also iron chelators. Other treatment of iron overload include Phlebotomy and lessen intake.

The sponsor submitted an IND (IND 45724) for the use of deferiprone to treat iron overload to FDA on July 15, 1994.

A pre-NDA meeting with FDA was held on October 9, 2001. Several concerns were raised by FDA during this meeting including lack of adequate and well-controlled trials,, a need for well defined clinical endpoints, safety considerations including agranulocytosis, neutropenia, frequency of withdrawals, termination of studies, need for non-thalassemic population, absence of nonclinical studies. In response to the sponsor’s question regarding the sufficiency of the then-available data to support approval of the drug, the Division stated that the current efficacy database did not appear to be strong and did not satisfy the regulatory requirements for adequate and well-controlled trials (AWCT), and that if what had been submitted in the background package represented all available data, the sponsor should submit efficacy and safety data from historical populations. The Division stated that efficacy endpoints would have to be linked to clinical improvement and that serum ferritin would have to be a validated surrogate for hepatic iron concentration. The sponsor stated that the estimated time of the NDA submission would be July, 2002.

The Agency granted Orphan drug designation to Deferiprone® on December 12, 2001 and Fast Track status on January 26, 2004. A pre-NDA meeting was held on July 9, 2004. The sponsor proposed application to be based on meta-analysis of studies (none AWCT); approval sought for non-thalassemia populations without studies of efficacy or safety. A telecon was held on December 14, 2004. The sponsor proposed changing primary end point (EP) for all trials to serum ferritin from Liver Iron Concentration **(**LIC). A pre-NDA meeting was held on May 15, 2006. The sponsor was advised that, “Pivotal” studies appear insufficient (small, retrospective), due to absence of non-thalassemic population, serum ferritin EP, validation of MRI T2* EP, restriction of indication to studied population, termination

of studies, need for AWCTs, use of retrospective studies, and inadequacy of Statistical Analysis Plan (SAP).

Referring to the Pre-NDA meeting between the agency and the sponsor on May 15, 2006, the sponsor submitted an NDA on January 29, 2009 as a last submission of a rolling NDA in the electronic Common Technical Document (eCTD) format . This submission included one main phase 3 study (LA 16-0102) and one supportive pivotal (observational) study (LA 12-9907). The results from several other studies (LA 01, LA 02, LA 02/06, LA 03, LA 04, LA 08-9701, LA 10-9902, LA 11, LA 15-0002, Borgna-Pignatti, LA 17-9701) were also submiited. All these supportive studies generally were uncontrolled, retrospective, had poor data quality, and had small number of subjects. Several publications from the literature were also cited and included.

The focus of this review is two identified pivotal trials LA 16-0102 and LA 12-9907.

Study LA-16-0102 is a randomized, open label, active comparator clinical study, and submitted as confirmatory evidence of safety and efficacy. The primary endpoint is putative changes in cardiac iron content as measured by magnetic resonance imaging (MRI) T2* performed at baseline and after one year of study drug therapy. The study enrolled transfusion-dependent patients with thalassemia who had previously been receiving deferoxamine. Patients were randomized to either continue deferoxamine or initiate therapy with deferiprone. Secondary endpoints included change in MRI T2* at Month 6 from baseline, cardiac ejection fractions, Left Ventricular Ejection Fraction (LVEF) assessed by both Cardiovascular Magnetic Resonance (CMR) and echocardiogram (ECHO), and Left Ventricular Shortening Fraction (LVSF) assessed by echocardiogram.

Study LA-12-9907 is a non-randomized, retrospective clinical study submitted as the main supportive clinical study. This study consisted of one clinical site's multi-year experience with deferiprone and deferoxamine therapy among 129 patients protocol-selected to undergo retrospective medical chart review. Baseline data were collected and follow-up information extracted from medical records over the five year data extraction period. The primary endpoint is incidence of cardiac disease using NYHA classification and physicians assessment of CHF, LVEF, and LVSF during the study period of approximately 5 years.

2.2 Data Sources

The applicant submitted this NDA including the data to the FDA CDER Electronic Document Room (EDR). The data sets were well documented and included definition files. The analysis dataset was not adequate and required data management, programming and information request. The clinical study reports and datasets are located at the following location:

\\CDSESUB1\EVSPROD\NDA021825\021825.ENX

3. STATISTICAL EVALUATION

There are two efficacy studies in this NDA and supportive epidemiology and safety studies based on publications.

The first efficacy study (LA16-0102) is a two arm comparative, multicenter, randomized, open-label, active controlled clinical trial comparing the relative efficacy of deferiprone (n=29) to that of deferoxamine (n=31) in removing excess cardiac iron in thalassemia major patients. This study focused predominantly upon putative changes in cardiac iron content as measured by magnetic resonance imaging (MRI) T2* performed at baseline and after one year of study drug therapy. The study enrolled transfusion-dependent patients with thalassemia who had previously been receiving deferoxamine. Patients were randomized to either continue deferoxamine or initiate therapy with deferiprone. The objective is to assess the efficacy of deferiprone versus deferoxamine to remove excess cardiac iron by measuring change in MRI T2* over 12 months of treatment

The second non-randomized, retrospective clinical study (LA-12-9907) was submitted as the main supportive clinical study. This study consisted of one clinical site's multi-year experience with deferiprone and deferoxamine therapy among 129 patients protocol-selected to undergo retrospective medical chart review. Baseline data were collected and follow-up information extracted from medical records over the five year data extraction period. The retrospective assessment included incidence of heart failure and survival during iron chelation with Deferiprone (n =54) or deferoxamine (n =75) in subjects with transfusion dependent ß-thalassemia.

All the analyses were performed by this reviewer and the tables and graphs presented here are based on this reviewer's analyses. The nominal p-values on secondary endpoints, and subgroup analyses are exploratory and given for information only. There was no formal protocol defined α-allocation for these analyses.

3.1 Evaluation of Efficacy – Study LA16-0102

3.1.1 Subject disposition

A total of 160 subjects with transfusion dependent β- thalassemia were screened for inclusion/exclusion criteria and all were assessed for MRI T2* and LVEF. Subjects were selected if the MRI T2* was between 8 and 20 ms and the LVEF was > 56%. Ninety-nine were excluded and a total of 61 patients were enrolled in the trial at the 4 investigational sites. They were stratified into 2 groups based on baseline MRI T2* (8 to 14 ms, 14 to 20 ms) and then randomized between deferiprone and deferoxamine treatment groups. While equal numbers were randomized into the 2 arms for the lower MRI T2* group (n=16 assigned to each treatment arm), there were a few more subjects (n=16) in the deferoxamine treated arm than in the deferiprone arm (n=13) in the higher T2* group.

Fifty-six (56) subjects completed the study and 5 subjects discontinued prematurely, 2 in the deferiprone arm (1 because of elevated liver enzymes and 1 who developed cytomegalovirus hepatitis and did not wish to continue) and 3 in the deferoxamine arm (2 for personal reasons and 1 because of deteriorating heart function).

3.1.2 Subject demographic and baseline characteristics

- **Baseline Demographic Characteristics**

Demographic characteristics for the 61 subjects showed that there was similar number of male and female subjects in the two treatment groups and that all the subjects were Caucasian. The range in age was from 18 to 35 years, and the mean age was 25.1 and 26.2 years of age for the Deferiprone and Deferoxamine treatment groups, respectively. The two groups were balanced at baseline with respect to age, weight, gender and ethnicity. The results are summarized in Table 2 below:

Table 2: Summary of Demographic Characteristics by Treatment Group

Characteristics	Randomized Treatment Groups Deferiprone (n=29)	 Deferoxamine (n=32)
Gender		
Male	15 (52%)	16 (50%)
Female	14 (48%)	16 (50%)
Age (years) at Baseline		
Mean ± SD	25.2 ± 3.8	26.2 ± 4.7
Median	25.0	27.0
Min, Max	18, 32	18, 35
Race (100% Caucasian)		
Ethnicity		
Greek	16 (55%)	18 (56%)
Italian	13 (45%)	14 (44%)
Weight (kg)		
Mean ± SD	57.7 ± 7.9	60.6 ± 13.2
Min, Max	43.8, 72.3	41.1, 91.0

- **Baseline Medical Characteristics**

Patients with thalassemia major are among those worldwide who receive substantial amounts of red cell products. The life-long need for transfusion renders patients vulnerable to transfusion transmitted viral infections such as hepatitis C virus and HIV. In LA16-0102, subjects' splenectomy status, hepatitis C, and HIV status was evaluated and the distribution of the status of subject regarding splenectomy, hepatitis C, and HIV tests at baseline is presented in Table 3. By design, HIV positive were excluded from the trial. All subjects enrolled in the trial were HIV negative.

Table 3: Summary of Splenectomy, and Hepatitis C at Baseline

Characteristics	Randomized Treatment Groups	
	Deferiprone (n=29)	Deferoxamine (n=32)
Splenectomy		
Yes	4 (14%)	11 (34%)
No	25 (86%)	21 (66%)
Hepatitis C		
Yes	18 (62%)	16 (50%)
No	11 (38%)	16 (50%)

The above table shows that there were more subjects who had splenectomy in the deferoxamine group than in the deferiprone group and more patients in the deferiprone treatment group tested positive to hepatitis C virus at baseline than in the deferoxamine treatment group.

- **Baseline Serum Ferritin**

Baseline serum ferritin concentrations (≤ 2,500 μg/L or > 2,500 μg/L) are summarized in Table 4 below:

Table 4: Baseline Serum Ferritin Levels

Serum Ferritin	Randomized Treatment Groups		
	Deferiprone (n=29)	Deferoxamine (n=32)	Total (n=61)
≤ 2,500 μg/L	24 (83%)	19 (59%)	43
> 2,500 μg/L	5 (17%)	13 (41%)	18

Fewer patients assigned to the deferiprone arm had a baseline serum ferritin greater than 2500 μg/L (17%) compared to those assigned to the deferoxamine arm (41%). Mean baseline serum ferritin concentrations were lower in the deferiprone treatment group than in the deferoxamine treatment group (1791 ± 1029 μg/L versus 2795 ± 2441 μg/L,). There were more subjects in the deferiprone treatment group than in the deferoxamine treatment group who had baseline serum ferritin concentrations below 2,500 μg/L (83% versus 59%).

- **Concurrent Medications and Treatment Compliance:**

Concurrent medication use was universal and differed somewhat between the 2 arms of the trial, particularly with respect to medications for respiratory symptoms (deferiprone, 28%; deferoxamine, 69%).

Compliance with oral and subcutaneous therapies was evaluated and compared between the two treatment groups. Compliance to oral therapy was reviewed monthly. The compliance was calculated as the percent of the number of openings (within interval higher than 4 hours recorded) divided by the number of doses prescribed in the deferiprone treatment group. In the deferoxamine treatment group, compliance was calculated as the percentage of completed infusions (as determined by the Crono™ infusion pump) divided by the number of infusions prescribed. The overall compliance with treatment was similar between the 2 arms (deferiprone, 93.7 ± 5.3%; deferoxamine, 93.2 ± 9.7%).

3.1.3 Analysis population

For the statistical evaluation of efficacy data, the intent-to-treat (ITT) and per protocol (PP) populations were used.

Intent-to-Treat (ITT) Population: Data from subjects who had received at least one dose of the drug, and who had at least two measurements of which one measurement was made after baseline were included in the ITT analysis. When there were no data available at a particular visit, last observation carried forward (LOCF) was used to fill in the missing data.

Per Protocol (PP) Population: Data from all randomized subjects who had completed the study were included in the PP analysis.

For the statistical evaluation of safety data, **Observed Cases (OC)** approach was used.

OC Population: Data from all randomized subjects were used at each visit. When there were no data available at a particular visit, LOCF was not used to fill in the missing data.

In this study, the assessment of efficacy was primarily based on the ITT population because this is a superiority study.

3.1.4 Efficacy variables

The primary efficacy analysis evaluated MRI T2* data at baseline, and at 6 and 12 months after commencing treatment.

Primary objective was to determine if deferiprone® exhibits superior efficacy in removing excess iron from the heart compared to deferoxamine®, as reflected by MRI T2* assessments. The primary endpoint was change in MRI T2* from baseline at month 12. The sponsor has proposed T2* as a reliable measure of cardiac iron content, with the implication that lessening of cardiac iron content is predictive of clinical benefit.

A sample size of 60 participants (30 in each treatment arm) was based on the analysis method of two-sample t-test with 5% level of significance, expected dropout rate of up to 20% (sponsor's study LA02) and 80% power, assuming that MRI T2* values were normally distributed and there was true mean difference of 2.3 ms between deferiprone and deferoxamine with a standard deviation of 2.5 ms. There was no justification in the study report for these assumptions. In case of assumption violation, alternative analysis methods were not clearly outlined in the prospective statistical analysis plan. We are concerned about the normality assumption due to small sample size of 30 patients in each arm and its potential impact on the interpretation of outcome. Actual sample size was 29 subjects in deferiprone arm and 31 subjects in deferoxamine arm.

The secondary efficacy measurements were the assessment of serum ferritin concentration and LIC. The serum ferritin concentration was assessed by Microparticle Enzyme Immunoassay (MEIA) and evaluated at baseline, quarterly and at month 12 or at early withdrawal. LIC was performed at baseline, and at 12 months (± 1 month), or at early withdrawal. LIC was assessed by the Superconducting Quantum-Interference Device (SQUID) BioSusceptometer.

Measures of cardiac function (LVEF and LVSF) were also evaluated through CMR and echocardiogram (ECHO). LVEF was assessed by both CMR and echocardiogram and LVSF was assessed by

echocardiogram. CMR was measured at baseline and was repeated at 6 months (± 1 month) and 12 months (± 1 month) or at the time of early withdrawal (± 1 month) and ECHO was measured at baseline and at 12 months or upon early withdrawal.

Additional analyses included the effects of covariates (splenectomy, baseline serum ferritin level).

3.1.5 Protocol defined primary efficacy analysis

Magnetic Resonance Imaging T2-Star (MRI T2*) was measured at baseline, 6 months, and 12 months. The protocol defined primary endpoint was the change in MRI T2* from baseline at month 12 (detect a difference of 2.3 ± 2.5 ms between the two treatments). This was to be analyzed by the analysis of variance (t-test) to determine the statistical significance of the difference between the two treatments.

The results for the protocol identified primary endpoint (changes in MRI T2* from baseline at 12 months) are given in Table 5 below and show that there is no statistically significant difference in the MRIT2* mean change from baseline at 12 months between Deferiprone and Deferoxamine.

Table 5: Change in MRI T2* at Month 12 from Baseline - ITT Population

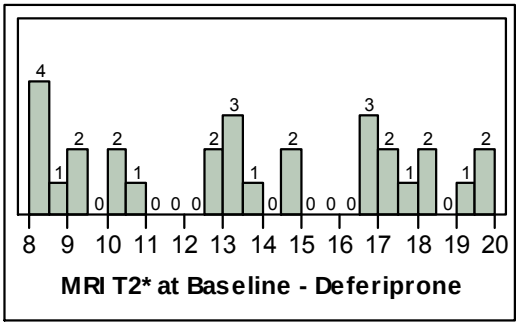
MRI T2* (milliseconds)	**Difference from Baseline at 12 Months**	
	Deferiprone (n=29)	**Deferoxamine (n=31)**
Mean (milliseconds) ± SD	3.9 ± 3.6	2.3 ± 4.1
95% CI	(2.6, 5.3)	(0.8, 3.9)
p-value (based on t-test with unequal variance)	0.0993	

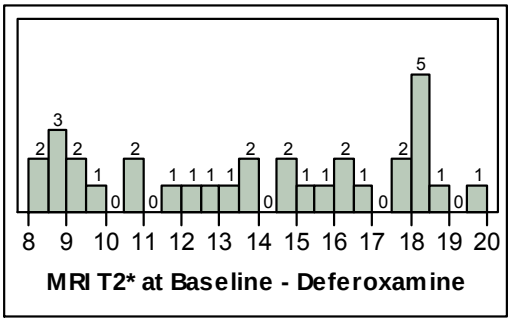
The observed difference between the two treatment arms for the mean change in MRI T2* from baseline at month 12 was 1.6 ms with a p-value of 0.0993 and was not statistically significant at 5% level. This analysis was based on the assumption of normal distribution for the primary endpoint.

3.1.6 Normality Issues

The distribution of the data were examined for departures from normality using various parametric and non-parametric tests. Figures 1, 2 and 3 show the distribution of MRI T2* at Baseline, at Month 12 and change in MRI T2* at Month 12 from Baseline respectively. These data are highly skewed and not normally distributed. The Shapiro-Wilk test of normality gives a p-value < 0.001 providing statistically significant evidence of departure from normality for these data. Therefore the use of prospective two-sample t- test analysis method may be invalid because the assumptions behind those tests are violated.

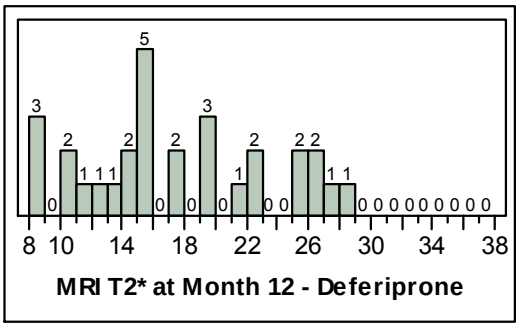
Figure 1: Histogram of MRI T2* at Baseline

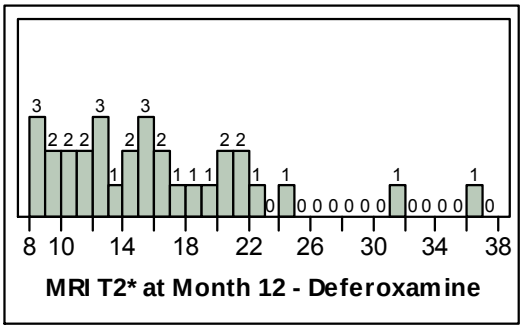




Test for normality $p < 0.001$ (Shapiro-Wilk test)

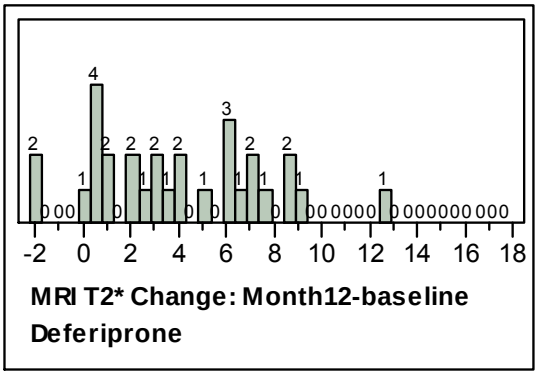
Figure 2: Histogram MRI T2* at Month 12

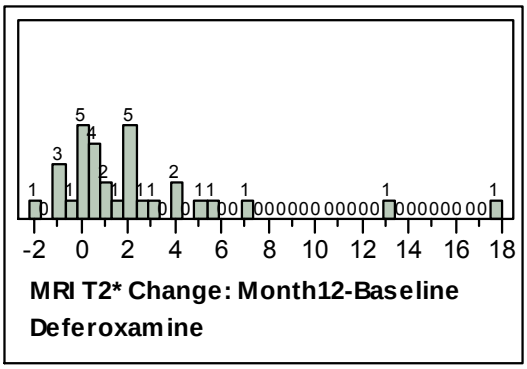




Test for normality $p < 0.001$ (Shapiro-Wilk test)

Figure 3: Histogram of Change in MRI T2* at Month 12 from the Baseline





Test for normality $p < 0.001$ (Shapiro-Wilk test)

The company noted this situation and used log transformation on MRI T2* in order to "linearize the scale, potentially resulting in normalization of the MRI T2* data". This log transformation was post-hoc, not prospectively planned in the protocol or the statistical analysis plan..

3.1.7 Sponsor's post-hoc primary efficacy analysis based on log(MRI T2*)

Sponsor stated that the degree of cardiac iron load is inversely related to T2*. Thus a difference of 2-3 units at a region of high T2* (e.g., over 20 ms) does not reflect the same degree of change in cardiac iron load as that for the same difference of 2-3 units at the region of low T2* (e.g., around 10 ms). Sponsor used log-transformation of the MRI T2* data to "linearize the scale, resulting in normalization of the MRI T2* data". The two-sample t-test was used to compare the mean changes in Log (MRI T2*) from baseline to 12 months between the two treatment groups. This log transformation was not prospectively planned in the protocol or the statistical analysis plan, the results for this post-hoc statistical analysis based on log transformation are given in Table 6 below:

Table 6: Change in Log (MRI T2*) at Month 12 from Baseline - ITT Population

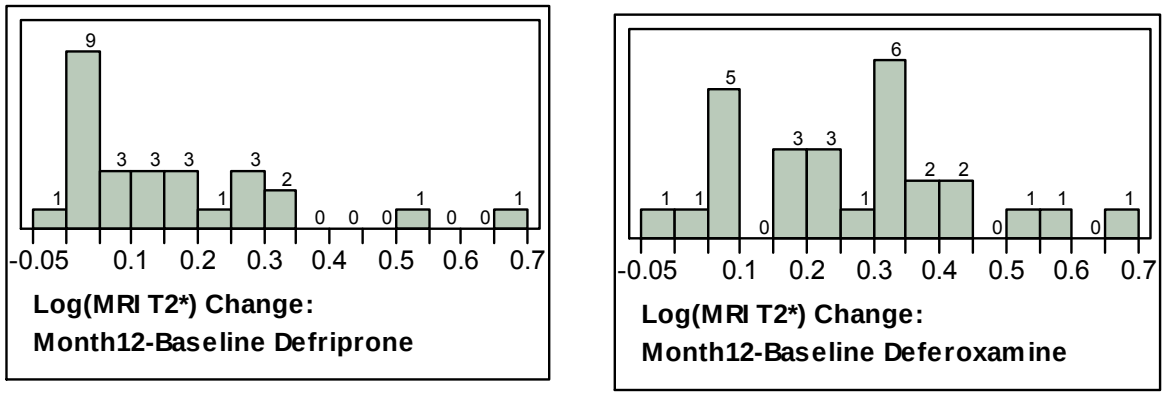
Log(MRI T2*) (milliseconds)	Difference from Baseline at 12 Months	
	Deferiprone (n=29)	Deferoxamine (n=31)
Mean Difference	0.237	0.123
p-value (based on t-test)*	0.0228	

*p-value is based on the post-hoc analysis and therefore nominal.

The sponsor stated that there is statistically significant difference in the log(MRIT2*) mean change from baseline at 12 months between Deferiprone and Deferoxamine.

These results are also subject to the same limitations as the raw MRI T2* data as our analysis shows lack of normality for log transformed data as well. The distribution of the log transformed data were examined for departures from normality using various parametric and non-parametric tests. Figure 4 shows the distribution of change in the log-transformed MRT T2* at Month 12 from Baseline. Again, there is a notable skewedness and log transformed MRT T2* data are not normally distributed (p-value < 0.001 - Shapiro-Wilk test of normality). Therefore the use of prospective two-sample t- test analysis method may be invalid because the assumptions behind those tests are violated.

Figure 4: Histogram of Change in Log(MRI T2*) at Month 12 from the Baseline



Test for normality p < 0.001 (Shapiro-Wilk test)

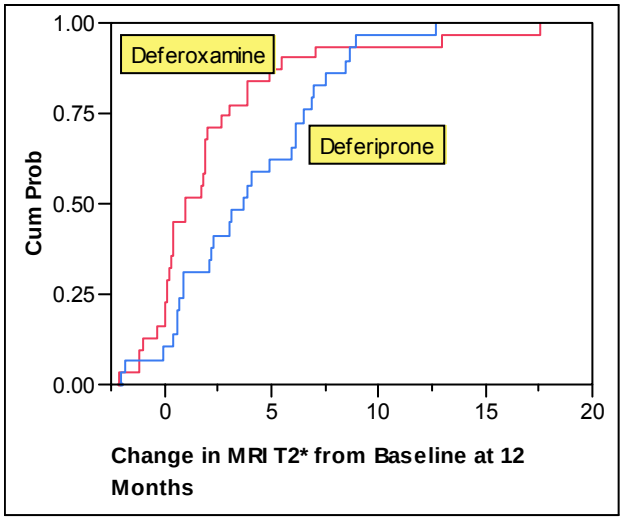
3.1.8 FDA's efficacy analysis based on non-parametric methods

Since the data are non-normal , non-parametric statistical methods that do not make the assumption of normal distributions provide valid analyses. Post-hoc Statistical Analysis using non-parametric methods was performed. Table 7 shows that the median change in the difference from baseline at 12 months for MRI T2* was 1.0 in the Deferoxamine group as compared to 3.7 in the Deferiprone group. This difference was statistically significant (p=0.0048 (Median Test)). The distributions of the change in the difference from baseline at 12 months for MRI T2* are shown in Figure 5. These two distributions are statistically significantly different (p=0.0168 (Kolmogrov-Smirnov Test) in favor of deferiprone.

Table 7: Median Change in MRI T2* at Month 12 from Baseline - ITT Population

MRI T2* (milliseconds)	**Difference from Baseline at 12 Months**	
	Deferiprone (n=29)	**Deferoxamine (n=31)**
Median	3.7	1.0
(10^{th} percentile, 90^{th} percentile)	(-0.1, 8.7)	(-1.2, 6.8)
(Minimum, Maximum)	(-2.0, 12.7)	(-0.1, 8.7)
p-value – Median Test	0.0048	
p-value - Wilcoxon (Rank Sum Test)	0.0124	
p-value - Kolmogrov-Smirnov Test	0.0168	

Figure 5: Distribution function of change in MRI T2* at Month 12 from the Baseline



This analyses shows that there is statistically significant shift in the distributions MRIT2* change from baseline at 12 months between Deferiprone and Deferoxamine in favor of Deferiprone.

3.1.9 Results of the statistical analysis of secondary efficacy endpoints

The secondary endpoint included Change in MRI T2* at Month 6 from Baseline, Measures of cardiac function, such as Left Ventricular Ejection Fraction (LVEF) assessed by both Cardiovascular Magnetic Resonance (CMR) and echocardiogram (ECHO), Left Ventricular Shortening Fraction (LVSF) assessed by echocardiogram.

The following were major secondary endpoints:

- Change in MRI T2* at Month 6 from Baseline
- Measure of cardiac function - Left Ventricular Ejection Fraction (LVEF) assessed by both Cardiovascular Magnetic Resonance (CMR) and echocardiogram (ECHO)
- Measure of cardiac function - Left Ventricular Shortening Fraction (LVSF) assessed by echocardiogram.
- Relative efficacy of deferiprone with respect to that of deferoxamine as assessed by serum ferritin concentration and the liver iron concentration (LIC). LIC was assessed by the use of a superconducting quantum interference device (SQUID) BioSusceptometer.
- **Serum Ferritin Concentrations**

The measurements provided by the sponsor (LVEF, LVSF, serum ferritin, LIC) are often used clinically to evaluate cardiac function and to determine the effect of a drug on body iron levels. However, these measurements may not be predictive of clinical benefit in altering morbidity or mortality in transfusion induced iron overload in patients with thalassemia.

The evaluation of major secondary endpoints is given below. The nominal p-values in each of these tables are provided for information only. These p-values are not intended for statistical inference because there was no protocol specified α-allocation for the control of type 1 error rate.

- **Change in MRI T2* at Month 6 from Baseline**

Secondary objective was to assess if Deferiprone (Ferriprox®) exhibits a trend in removing excess iron from the heart compared to Deferoxamine®, as reflected by MRI T2* assessments at 6 months. The primary endpoint was change in MRI T2* from baseline at month 6.

The distribution of the data were examined for departures from normality using various parametric and non-parametric tests. These tests showed that the difference in MRI T2* evaluations at month 6 or the log-transformed MRT T2* data from the baseline were not normally distributed. Therefore, non-parametric tests of hypotheses were also used for valid analyses. The results are summarized in the Table 8 below:

Table 8: Change in MRI T2* at Month 6 from Baseline - ITT Population

MRI T2* (milliseconds)	Randomized Treatment Groups -- Difference from Baseline at 6 Months		
	Deferiprone (n=29)		Deferoxamine (n=31†)
Analysis - Protocol & SAP identified[1]			
Mean (milliseconds) ± SD	2.8 ± 3.2		1.5 ± 3.0
95% CI	(1.6, 3.9)		(0.4, 2.6)
p-value (unequal variance)		0.1261	
Analysis-based on log transformation - (not indentified in protocol or SAP) [1]			
Diff (log 12 (or 6) – log base)	0.165423		0.083382
p-values[3]		0.0404	
Analysis– Non-parametric tests[2]			
Median	2.2		0.5
(10^{th} percentile, 90^{th} percentile)	(-0.3, 8.9)		(-0.8, 7.4)
Range (Minimum, Maximum)	(-1.8, 10.7)		(-1.6, 7.4)
p-value – Median Test*		0.0729	
p-value - Wilcoxon (Rank Sum Test)*		0.0482	
p-value - Kolmogrov-Smirnov Test*		0.0866	

* Nominal p-values are provided for information only

[1] The MRI T2* between the Deferiprone and Deferoxamine treatment groups was compared by the two sample t-test assuming unequal variances. The assumptions of normality behind these tests are violated

[2] Since the data are non-normal but other criteria are met, non-parametric statistics (Wilcoxon/Kruskal-Wallis Rank Sum tests and the Median Test (Number of Points Above Median), Kolmogorov-Smirnov test).provide valid analyses. The asymptotic p-value for the Kolmogorov-Smirnov test is 0.0168 (12 months). This indicates rejection of the null hypothesis that the distributions are identical for the two groups.

[3] The Log (MRI T2*) between the Deferiprone and Deferoxamine treatment groups was compared by the two sample t-test by the sponsor.

† Subject C1-40 had baseline MRI T2* level value only and was not eligible to be included in the ITT population.

In summary, a trend is seen in favor of deferiprone in terms of MRI T2* at 6 months. These data and analyses at 6 months have the same drawbacks at MRI T2* data and analyses at 12 months.

- **Measure of cardiac function - Left Ventricular Ejection Fraction (LVEF) assessed by both Cardiovascular Magnetic Resonance (CMR) and echocardiogram (ECHO)**

An analysis was performed on CMR LVEF (ITT population). As shown in the following Table 9, the baseline mean LVEF in the deferiprone arm was 69.7 ± 5.4% and increased by 3.1 ± 3.6% at 12 months after beginning treatment. The baseline mean LVEF in the deferoxamine arm was 68.4 ± 4.9% and increased by 0.3 ± 3.4% at 12 months after beginning treatment. The nominal p value for the difference between the treatment groups was 0.0034.

Table 9: CMR LVEF (%) Values at Baseline and after 6 and 12 Months of Therapy

	Randomized Treatment Groups					
	Baseline		Change from Baseline to 6 Months		Change from Baseline to 12 Months	
CMR LVEF (%)	Deferiprone (n=29)	Deferoxamine (n=31†)	Deferiprone (n=29)	Deferoxamine (n=31†)	Deferiprone (n=29)	Deferoxamine (n=31†)
Mean ± SD	69.7 ± 5.4	68.4 ± 4.9	2.0 ± 2.7	0.5 ± 3.5	3.1 ± 3.6	0.3 ± 3.4
95% CI	(67.6, 71.7)	(66.6, 70.2)	(1.0, 3.0)	(-0.8, 1.8)	(1.7, 4.4)	(-0.9, 1.6)
Range	(58, 80)	(60. 79)	(-3, 9)	(-9, 9)	(-3, 11)	(-8, 5)
p-value*	0.3408		0.0722		0.0035	

* Nominal p-values are provided for information only. Mean changes in CMR LVEF from baseline to 6 months and 12 months were compared between the two treatment groups by using the two-sample t-test (assuming unequal variances).
† Subject C1-40 had baseline CMR LVEF level value only and was not eligible to be included in the ITT population.

No patient had an ejection fraction of less than 56% at any time during the study. The sponsor indicates that this implies that there was no evidence of congestive heart failure in any of the patients at any time (congestive heart failure is usually defined as an LVEF of <50%).

An analysis was performed on Echocardiogram (ECHO) LVEF (ITT population). As shown in Table 10, the mean baseline LVEF of subjects treated with deferiprone was 64.7 ± 6.7% and over the course of 12 months of treatment, the LVEF increased by 2.5 ± 6.0%. The mean baseline LVEF of subjects treated with deferoxamine was 64.3 ± 6.9% and over the course of 12 months of treatment, the LVEF decreased by 0.6 ± 4.9%. The nominal p value for the difference between the treatment groups was 0.0358.

Table 10: ECHO LVEF (%) Values at Baseline and after 12 Months of Therapy

	Randomized Treatment Groups			
	Baseline		Change from Baseline to 12 Months	
ECHO LVEF (%)	**Deferiprone (n=29)**	**Deferoxamine (n=32)**	**Deferiprone (n=28†)**	**Deferoxamine (n=31†)**
Mean ± SD	64.7 ± 6.7	64.3 ± 6.9	2.5 ± 6.0	-0.6 ± 4.9
95% CI	(62.2, 67.2)	(61.9, 66.7)	(0.2, 4.8)	(-2.4, 1.2)
Range	(54, 79)	(50. 77)	(-9, 16)	(-8, 10)
p-value*	0.8086		0.0382	

* Nominal p-values are provided for information only. Mean changes in from baseline to 12 months were compared between the two treatment groups by using the two-sample t-test (assuming unequal variances).
† Subjects A1-47 and C1-40 did not have ECHO LVSF value at 12 months and were not eligible to be included in the ITT population.

An analysis was performed on Echocardiogram (ECHO) LVSF (ITT population). As shown in the following Table 11, the mean baseline LVSF of subjects treated with deferiprone was 36.3 ± 4.4% and over the course of 12 months of treatment, the LVSF increased by 2.6 ± 7.4%. The mean baseline LVSF of subjects treated with deferoxamine was 36.4 ± 4.3% and over the course of 12 months of treatment, the LVSF decreased by 1.1 ± 3.8%.

Table 11: ECHO LVSF (%) Values at Baseline and after 12 Months of Therapy

ECHO LVSF (%)	Randomized Treatment Groups			
	Baseline		Change from Baseline to 12 Months	
	Deferiprone (n=29)	Deferoxamine (n=32)	Deferiprone (n=28†)	Deferoxamine (n=31†)
Mean ± SD	36.3 ± 4.4	36.4 ± 4.3	1.9 ± 5.5	-1.1 ± 3.8
95% CI	(34.6, 38.0)	(34.9, 37.9)	(-0.2, 4.0)	(-2.5, 0.3)
Range	(31.2, 47.0)	(30.0, 44.0)	(-8.0, 12.0)	(-8.8, 8.0)
p-value*	0.9540		0.0202	

* Nominal p-values are provided for information only. Mean changes from baseline to 12 months were compared between the two treatment groups by using the two-sample t-test (assuming unequal .variances).
† Subjects A1-47 and C1-40 did not have ECHO LVSF value at 12 months and were not eligible to be included in the ITT population.

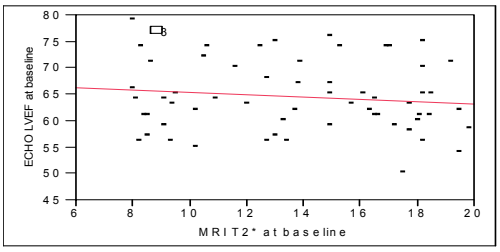
- **Relationship between MRI T2* and MRI LVEF, ECHO LVEF and ECHO LVSF**

In response to a request from the Medical officer, this reviewer analyzed the correlations between the MRI T2*, the CMR LVEF, the echocardiogram LVEF and the echocardiogram LVSF. LVEF and LVSF are considered meaningful measure of cardiac function.

There were no statistically significant correlations between the MRI T2* and the ECHO LVEF or the ECHO LVSF at baseline or at any follow-up evaluation, nor with the change in ECHO LVEF or the ECHO LVSF at the end of the study. There were modest correlations at 6 (nominal p values= 0.027) and 12 (nominal p values= 0.015) months of therapy between the MRI T2* and the MRI LVEF but not at baseline (nominal p value, 0.2329). The maximum correlation was 0.31. These data suggest that, at best, the change in MRI T2* explains approximately 10% of the variation in heart function as measured by MRI LVEF, ECHO LVEF or ECHO LVSF. The scatter plots and correlations are given below in Figures 6, 7 and 8:

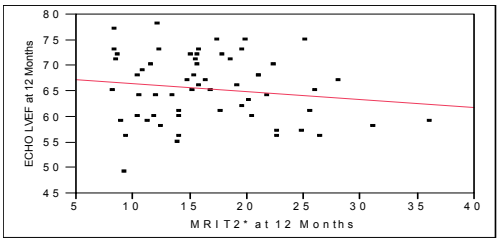
Figure 6: Scatter plot and correlation between MRI T2* and ECHO LVEF at

Baseline



Pearson's correlation = - 0.1298
Spearman's rank correlation Rho = - 0.1221

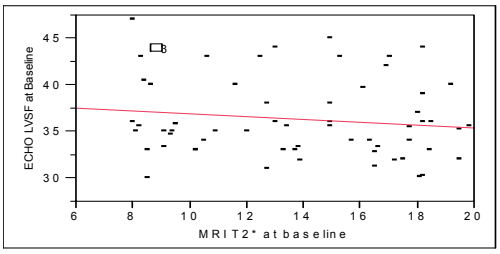
12 months



Pearson's correlation = - 0.1473
Spearman's rank correlation Rho = - 0.1034

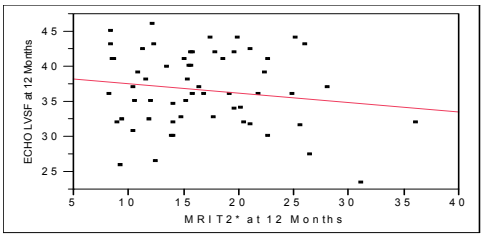
Figure 7: Scatter plot and correlation between MRI T2* and ECHO LVSF at

Baseline



Pearson's correlation = - 0.1388
Spearman's rank correlation Rho = - 0.1276

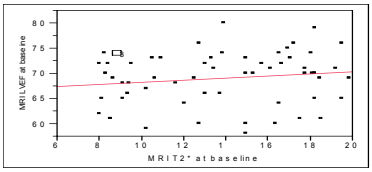
12 months



Pearson's correlation = - 0.1551
Spearman's rank correlation Rho = - 0.0738

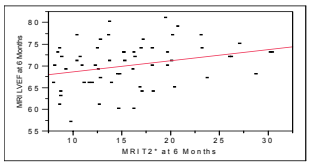
Figure 8: Scatter plot and correlation between MRI T2* and MRI LVEF at

Baseline



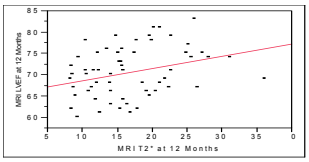
Pearson's correlation = 0.1554
Spearman's rank correlation Rho = 0.1491

6 months



Pearson's correlation = 0.2848
Spearman's rank correlation
Rho = 0.2932 nominal p=0.0230

12 months



Pearson's correlation = 0.3127
Spearman's rank correlation
Rho = 0.3051 nominal p=0.0178

- **Liver Iron Concentration**

As a secondary objective, this study also evaluated the relative efficacy of deferiprone with respect to that of deferoxamine as assessed by serum ferritin concentration and LIC. LIC was assessed by the use of a superconducting quantum interference device (SQUID) BioSusceptometer.

There were no significant differences in the changes of LIC and serum ferritin levels from baseline to 12 months between the two treatment groups (nominal p=0.3961 and 0.1598, respectively). The results are given in Tables 12 and 13. These results suggest that the two treatments have similar efficacy in controlling non-cardiac iron load at the doses employed. Covariates (splenectomy status, hepatitis C status, baseline serum ferritin) had no significant influence on the comparison of LIC between the two treatment groups.

Table 12: Liver Iron Concentration (LIC) at Baseline and at End of Treatment

LIC (mg Fe/g dry weight liver)	Randomized Treatment Groups			
	Baseline		Change from Baseline to 12 Months	
	Deferiprone (n=28[1])	Deferoxamine (n=32)	Deferiprone (n=27[2])	Deferoxamine (n=30[2])
Mean ± SD	6.16 ± 6.02	6.32 ± 5.77	-0.93 ± 2.93	-1.54 ± 2.49
Range	(1.5, 33.3)	(0.7, 26.4)	(-8.7, 5.2)	(-8.8, 1.6)
p-value*	0.9161		0.3961	

Fe=iron

* Nominal p-values are provided for information only. Mean changes from baseline to 12 months were compared between the two treatment groups by using the two-sample t-test.

[1]Subjects C1-44 did not have baseline LIC value only and was not eligible to be included in the ITT population.

[2] Subjects A1-20, C1-40, C1-44 and C1-52 did not have LIC value at 12 months and were not eligible to be included in the ITT population.

- **Serum Ferritin Concentrations**

The baseline mean serum ferritin level (1,791 µg/L) was lower in the deferiprone treated group compared to that in the deferoxamine treated group (2,795 µg/L). Mean serum ferritin levels tended to rise in the deferiprone treated patients at 3 months, but then there was a gradual fall until, at the end of study, the mean serum ferritin was 1,609 µg/L. In contrast, mean serum ferritin levels fell progressively in patients treated with deferoxamine, and plateaued at 12 months so that at the end of study, the mean level was 2.247 µg/L. These data are summarized in the following Table 13.

Table 13: Serum Ferritin Levels at Baseline and at End of Treatment

Serum Ferritin Levels	Randomized Treatment Groups			
	Baseline		Change from Baseline to 12 Months	
	Deferiprone (n=29)	Deferoxamine (n=32)	Deferiprone (n=29)	Deferoxamine (n=32)
Mean ± SD	1791 ± 1029	2795 ± 2441	-181 ± 826	-466 ± 739
Range	(289, 5345)	(280, 9300)	(-2179, 1990)	(-2208, 606)
p-value*	0.0391		0.1598	

Fe=iron

* Nominal p-values are provided for information only. Mean changes from baseline to 12 months were compared between the two treatment groups by using the two-sample t-test.

3.1.8 Sensitivity analyses

Table 14 describes the disposition of patients for the main study (LA16-0102). A total of 61 patients were randomized and dosed with 29 patients in the deferiprone arm and 32 patients in the deferoxamine arm. This was also safety population for this study. A total of 56 patients completed 12 months of treatment with 27 patients in the deferiprone arm and 29 patients in the deferoxamine arm (Per protocol population (PP)).

Table 14: Disposition of Patients

	Treatment		
Population	N (Deferiprone)	N (Deferoxamine)	Total
ITT population (dosed)	29	32	61
PP population† (completed)	27	29	56

† Per Protocol (PP) population -- All randomized subjects who had completed the study. Five Subjects (A1-20, A1-47, A1-48, C1-40 and C1-52) did not complete the study and were not eligible to be included in the PP population.

The results obtained from the Per Protocol (PP) population (data from all randomized subjects who had completed the study were included in the PP analysis) were similar to those obtained on the ITT population. The issues with the "primary" analysis were similar to the to the ITT population, i.e., lack of normality. Applicable Nonparametric analyses show difference in terms of MRI T2* for the PP. The results of analysis based on the PP population support the findings based on the ITT population.

3.2 Evaluation of Efficacy – Study LA12-9907

Study 12-9907 was an observational study based on the retrospective assessment of medical records at a single center in Italy. The patient population consisted of subjects with transfusion dependent ß-thalassemia and the external comparator used was deferoxamine. Deferiprone, (Apotex Research Inc., Canada) was administered orally three times a day, seven days per week, at the dose of 75 mg/kg/day. The dose was adjusted to the patients needs within the range from 35 to 100 mg/kg/day. Deferoxamine (Novartis Pharma, Switzerland) was prescribed at the dose of 20 to 60 mg/kg/day, as an 8 to 12 hours subcutaneous infusion via infusion pump, four to seven days a week. Batch numbers were not collected for this study. The duration of the study was 5 years (between 31 January 1995 and 29 March 2001)

This was an open label, non-randomized, single center, parallel longitudinal, retrospective assessment of medical records of patients with transfusion-dependent β-thalassemia treated for at least four years with deferiprone (Ferriprox®) or deferoxamine to compare incidence of cardiac disease and cardiac disease free survival.

3.2.1 Patient Population - Study LA12-9907

The clinical records of the 168 patients with β-thalassemia treated at the center were screened, and 129 met all of the inclusion and exclusion criteria (Main Population). Initially, the study was to be limited to those patients who had at least three serum ferritin concentration determinations in the two years preceding the initiation of the study period. This exclusion criterion could limit the number of otherwise qualified patients, sponsor amended the protocol to include all patients with less than three serum ferritin concentration determinations in the two years preceding the initiation of the study period. The decision was made after the completion of the review period and prior to the assessment of any data. To comply with the initial protocol, a subgroup analysis was also conducted on the 107 patients (50 patients treated with deferiprone (DFP), and 57 patients treated with deferoxamine (DFO)) with at least three serum ferritin values during the two years prior to the start of the study. (Per Protocol Population). To reduce any potential effect of time on chelation therapy on the occurrence and/ or progression of iron-induced cardiac disease, patients from either therapy arm were matched for age at the start of chelation therapy. Forty-seven patients treated with deferiprone were found to have a match with forty-seven patients from the deferoxamine group for the same age at start of chelation therapy. A second subgroup analysis was done on these 94 patients (Age Matched). Table 15 describes the analysis populations.

Table 15: Analysis of study populations - Study LA12-9907

Analysis Populations	Number of Patients DFP	DFO	Total	Characteristics
Main Population (extended)	54	75	129	Patients meeting all inclusion and exclusion criteria (as per protocol, amendment #1 and inclusion of patients with less than three serum ferritin assessments)
Per Protocol Population (Subgroup I)	50	57	107	Patients from the Main group with at least three serum ferritin values during the two years prior to the start of the study (as per protocol and amendment #1)
Age-matched (Subgroup II)	47	47	94	Age-matched -- Patients from the Main group matched for age at start of chelation therapy

3.2.2 Baseline characteristics

Table 16 displays the baseline (the start of the study) characteristics for the two therapy groups.

Age at the baseline and age at the start of chelation therapy were significantly different in two therapy groups for the main population resulting in the two arms not comparable for the main population from the amended protocol. Age matched (matched for age at start of chelation therapy) group is balanced (no significant difference) with respect to baseline variables and appears to be a valid group for a comparison of two therapies. Note that some important baseline variables such as splenectomy status were not measured in this study. Also limited information such as baseline hepatic iron concentration was available.

Table 16: Comparison of the two therapy groups at the start of the study (Baseline)

Baseline Variables	Main Population -- n= 129 DFP N = 54	DFO N = 75	p-value*	Per Protocol n = 107 DFP N = 50	DFO N = 57	p-value*	Age-matched -- n = 94 DFP N = 47	DFO N = 47	p-value*
Percentage of female	44% (24/54)	49% (37/75)	0.58	44% (22/50)	46% (26/570	0.867	43% (20/47)	49% (23/47)	0.535
Mean age ± SD [years]	17.1 ± 4.1	19.4 ± 6.9	0.02	16.7 ± 3.7	18.2 ± 7.2	0.172	17.0 ± 4.3	16.5 ± 5.8	0.600
Mean age ± SD at the start of chelation therapy [years] (Number of patients available)	4.5 ± 2.7 (54)	6.8 ± 4.7 (72)	0.001	4.2 ± 2.2 (50)	6.5 ± 4.7 (57)	0.002	4.4 ± 2.6 (47)	4.4 ± 2.7 (47)	1.00
Mean serum ferritin ± SD [µg/L] -- Serum Ferritin Concentration (Number of patients available)	2033 ± 919 (51)	1809 ± 1464 (60)	0.33	2054 ± 917 (50)	1779 ± 1482 (57)	0.249	2072 ± 950 (45)	1516 ± 725 (40)	0.004
Percentage of patients with more than 50% of their serum ferritin results > 2,500 µg/L	24% (12/51)	15% (9/60)	0.25	24% (12/50)	14% (8/57)	0.187	24% (11/45)	10% (4/40)	0.081
Percentage of patients positive for HCV antibodies - Hepatitis C Status	87% (45/52)	80% (52/65)	0.35	85% (41/48)	74% (37/50)	0.161	85% (39/46)	73% (32/44)	0.161
Mean transfusional iron input ± SD [mg Fe/kg body weight/day] (Number of patients available)	0.464 ± 0.085 (49)	0.432 ± 0.110 (61)	0.10	0.463 ± 0.085 (49)	0.437 ± 0.111 (57)	0.167	0.463 ± 0.089 (43)	0.482 ± 0.083 (40)	0.336
Mean urinary iron excretion ± SD [mg Fe/day] (Number of patients available)	14.7 ± 10.7 (49)	15.5 ± 12.2 (48)	0.73	14.7 ± 10.7 (49)	15.6 ± 12.3 (47)	0.711	14.4 ± 10.5 (43)	12.2 ± 6.7 (36)	0.269
Mean hepatic iron concentration ± SD - SQUID [mg Fe/g liver wet weight] (Number of patients available)	1.6 ± 0.7 (46)	15.5 ± 12.2 (48)	0.73	1.6 ± 0.7 (46)	0.9 ± 0.6 (16)	0.002	1.5 ± 0.7 (40)	1.1 ± 0.6 (11)	0.062
Mean hepatic iron concentration ± SD - Biopsy [mg Fe/g liver dry weight] (Number of patients available)	1.6 ± 0.7 (46)	0.9 ± 0.6 (16)	0.002						
Mean hepatic iron concentration ± SD - Biopsy [mg Fe/g liver dry weight] (Number of patients available)	8.5 ± 5.7 (34)	---	----	8.5 ± 5.7 (34)	---		8.8 ± 5.8 (32)	--	
Percentage of patients with **cardiac disease** at the first assessment	13% (7/54)	16% (12/75)	0.63	10% (5/50)	14% (8/57)	0.524	15% (7/47)	11% (5/47)	0.537

* nominal p-values are provided to asses balance in baseline variables.

3.2.3 Efficacy variables

Primary objective was to investigate the incidence of cardiac disease using NYHA classification and physicians assessment of CHF, LVEF, and LVSF, and the survival in patients treated with deferiprone and to compare the results with those patients treated with conventional therapy, daily subcutaneous infusion of deferoxamine, over the same period of time. A secondary objective was to evaluate the progression of cardiac disease in patients treated with either deferiprone or deferoxamine

3.2.4 Efficacy analysis

The primary endpoint was incidence of cardiac disease during the study using NYHA classification and physicians assessment of CHF, LVEF, and LVSF. Changes (worsening, no change, or improvement) in cardiac disease were based on cardiac status using NYHA class.

- **Assessment of Cardiac Disease using the NYHA classification**

The following Table 17 compares different aspects of cardiac disease between therapy groups, using the NYHA classification. All the nominal p-values are based on two-sided Fisher's exact test. The results are given in Table 17 below.Recall that the age-matched population is most appropriate for a comparison of two therapies. All the nominal p-values (based on two-sided Fisher's exact test) for this age-matched group are not significant even at 5% level of significance.

Table 17: Different aspects of cardiac disease using the NYHA classification

	Frequency of patients (%) with cardiac disease at the last assessment		
Therapy	Main Population n=129	Per protocol n=107	Age-matched n=94
Deferiprone	7/54 (13.0 %)	6/50 (12.0%)	7/47 (14.9%)
Deferoxamine	22/75 (29.3 %)	16/57 (28.1%)	11/47 (23.4%)
p-value (Exact)*	0.033	0.0549	0.4323
Therapy	**Frequency of patients (%) with a worsening of NYHA classification from the first to the last assessment**		
	Main Population n=129	Per protocol n=107	Age-matched n = 94
Deferiprone	2/54 (3.7%)	2/50 (4.0%)	2/47 (4.3%)
Deferoxamine	15/75 (20.0%)	11/57 (19.3%)	9/47 (19.2%)
p-value (Exact)*	0.007	0.0183	0.0502
Therapy	**Frequency of patients (%) with cardiac disease during the study among patients who were initially cardiac disease-free**		
	Main Population n=110	Per protocol n= 94	Age-matched n= 82
Deferiprone	2/47 (4.3 %)	2/45 (4.4%)	2/40 (5.0%)
Deferoxamine	13/63 (20.6%)	11/49 (22.5%)	8/42 (19.1%)
p-value (Exact)*	0.022	0.0155	0.0888
Therapy	**Frequency of patients (%) with a worsening of NYHA classification among patients with cardiac disease at the first assessment**		
	Main Population n=19	Per protocol n=13	Age-matched n= 12
Deferiprone	0/7 (0%)	0/5 (0%)	0/7 (0%)
Deferoxamine	4/12 (33.3%)	2/8 (25%)	2/5 (40%)
p-value (Exact)*	0.2451	0.4872	0.1515
Therapy	**Frequency of patients (%) with an improvement of NYHA classification among patients with cardiac disease at the first assessment**		
	Main Population n=19	Per protocol n= 13	Age-matched n= 12
Deferiprone	3/7 (42.9%)	1/5 (20%)	3/7 (42.9%)
Deferoxamine	3/12 (25.0%)	3/8 (37.5%)	1/5 (20%)
p-value (Exact)*	0.6169	1.00	0.5758

* All the nominal p-values are based on two-sided Fisher's exact test) are provided for information only.

- **Survival Analysis**

Table 18 provides Information on 3 out of 4 patients who died of cardiac disease. Four patients (Patients # 15, 21, 85 and 171) all treated with deferoxamine died during the study period. Three of these patients had cardiac disease at the first assessment of the study period and died because of irreversible worsening of their cardiac condition (Table 18). One death occurred during the second year, whereas the other two occurred during the last year of the review period. The forth death occurred in a patient with a history of drug addiction but no signs of cardiac disease. This patient died within a few hours of being admitted into a provincial hospital for acute abdominal pain. No cause of death was provided to sponsor. This death was not included among the deaths in the survival analysis.

Table 18: Information on 3 out of 4 patients who died of cardiac disease

Age at start of Review period	Gender	Age of start of first chelation (years)	Chelation therapy during study	Cardiac Disease at Baseline as per NYHA Class	Compliance with chelation (%)	HIC closest to time of death (mg/g wet weight)	% serum ferritin > 2,500 μg/L
26 years	Male	13	DFO	Yes	54	NA	89
23 years	Male	8	DFO	Yes	73	1.004	25
23 years	Female	NA	DFO	Yes	NA	6.100	NA

3.3 Evaluation of Safety

- **Study LA16-0102**

Safety was evaluated for all randomized patients (61 patients in total). Total exposure was 27 patient-years for the deferiprone treated patients and 30 patient-years for the deferoxamine treated patients. The mean dose of deferiprone was 92 mg/kg/d and the mean dose of deferoxamine was 43 mg/kg for 5.7 days/wk. Adverse reactions (AR) were summarized according to the Medical Dictionary for Regulatory Activities (MedDRA) version 7.0. The most frequent adverse reactions are summarized in the following Table 19.

Table 19: Most Frequent Adverse Reactions reported in Study LA 16-0102

Most Frequent Adverse Reactions reported in Study LA 16-0102	Deferiprone		Deferoxamine	
	Total Reports	279	Total Reports	131
	Total Subjects Reporting	29	Total Subjects Reporting	32
	Total Subjects	29	Total Subjects	32
	Total Exposure (subject-years)	27.34	Total Exposure (subject-years)	30.2
Body System/ Preferred term	#	%	#	%
Gastrointestinal Disorders	19	65.5	6	18.8
General Disorders & Administration Site	3	10.3	11	34.4
Investigations	19	65.5	13	40.6
Metabolism & Nutrition Disorders	9	31.0	0	0.0
Musculoskeletal & Connective Tissue Disorders	12	41.4	7	21.9
Nervous System Disorders	8	27.6	6	18.8
Skin & Subcutaneous Tissue Disorders	3	10.3	6	18.8

Gastrointestinal adverse reactions (AR) were more common in patients receiving deferiprone than in those receiving deferoxamine. Common symptoms included upper abdominal pain, nausea, vomiting, and diarrhea. Administration site reactions (erythema, induration, inflammation and pruritis) occurred only in patients receiving deferoxamine. Hypersensitivity reactions were reported in 1 patient in each arm of the trial. The neutrophil count was decreased in 1 patient receiving deferiprone and in 4 patients receiving deferoxamine. No patients in either arm experienced agranulocytosis. Elevations in hepatic enzymes occurred in more patients in the deferiprone (alanine aminotransferase, 11 patients; aspartate aminotransferase, 6 patients; gamma-glutamyltransferase, 4 patients) compared to the deferoxamine

treated patients (4; 1; and 0, respectively). There was also an upward trend in ALT at 3, 6 and 9 months in patients receiving deferiprone compared to those treated with deferoxamine, although the trend was no longer apparent at 12 months of treatment. T wave inversions occurred in 5 patients receiving deferiprone and none of those receiving deferoxamine. More patients treated with deferiprone had an increase in appetite and weight gain than those treated with deferoxamine. Arthralgia was reported in 8 patients receiving deferiprone compared to 3 patients receiving deferoxamine. Other adverse reactions were reported in approximately similar numbers of patients in each arm of the trial.

There were no deaths during the course of the trial. The safety database from this study is inadequate to establish that important toxicity has been properly defined and balanced against efficacy. The number of patients treated with deferiprone totaled only 29. In addition, the length of the study was only 12 months, and it would be expected that patients would be treated for a lifetime with deferiprone.

The safety database from randomized trial LA16-0102 is inadequate because there were only 29 patients exposed to deferiprone during its conduct. From previous studies and postmarketing reports, gastrointestinal ARs, arthralgia and increases in serum transaminases are known to be associated with the use of deferiprone, and they occurred in this study as well.

- **Study LA12-9907**

One hundred and forty nine (149) patients were evaluated in the study. There was no formal collection of safety data. The mean exposure time to deferiprone was 5.27 years (range, 3.52 to 6.1 years) and to deferoxamine was 5.91 years (range, 2.04 to 6.16 years). Total interruption for deferiprone therapy was 1392 days (average, 38.67; range, 3 to 411 days) and for deferoxamine was 865 days (average, 41.19; range 2 to 285 days).

There were 4 deaths observed during the study and did not appear to be due to adverse reactions to either medication.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race and Age

Data were analyzed by gender for the primary endpoint of Change in MRI T2* at Month 12 from Baseline between Deferiprone and Deferoxamine Treatment Groups - ITT Population. The race was 100% Caucasian and the age range was 18 to 35 years. Table 20 shows that the results are consistent with the primary analysis in both gender groups (Female and Male).

Table 20: Change in MRI T2* at Month 12 from Baseline for Gender

MRI T2* (milliseconds)	Difference in MRI T2* from Baseline at 12 Months			
	Male		Female	
	DFP	DFO	DFP	DFO
N	15	15	14	16
Analysis - Protocol & SAP identified				
Mean (milliseconds) ± SD	4.2 ± 3.8	2.9 ± 5.4	3.6 ± 3.4	1.7 ± 2.4
95% CI	(2.1, 6.4)	(-0.1, 5.9)	(1.7, 5.6)	(0.4, 3.0)
p-value (unequal variance)*	0.4358		0.0926	
Analysis-based on log transformation - (not indentified in protocol or SAP)				
Diff (log 12 – log base)	0.2526	0.1277	0.2197	0.1177
p-values*	0.1213		0.1152	
Analysis– Non-parametric tests				
Median	3.7	1.7	3.2	1.0
(10^{th} percentile, 90^{th} percentile)	(-0.6, 10.2)	(-1.2, 14.8)	(-1.0, 8.8)	(-0.8, 6.0)
Range (Minimum, Maximum)	(-2.0, 12.7)	(-1.2, 17.6)	(-1.8, 8.8)	(-2.1, 7.1)
p-value – Median Test*	0.0120		0.1501	

* Nminal p-values are provided for information only

4.2 Other Special/Subgroup Populations

Three special groups (Splenectomy Status, Hepatitis C Status, and Serum Ferritin Concentrations) of clinical relevance were identified by the clinical team. Data were analyzed for these three special groups using non-parametric methods for the primary endpoint of Change in MRI T2* at Month 12 from Baseline between Ferriprox and Deferoxamine Treatment Groups - ITT Population. The results are given in the following Tables 21, 22 and 23 for Splenectomy Status, Hepatitis C Status, and Serum Ferritin Concentrations respectively.

- **Efficacy by Splenectomy Status**

Table 21: Efficacy by Splenectomy Status

MRI T2* (milliseconds)	Difference from Baseline at 12 Months			
	Splenectomy - Yes		Splenectomy - No	
	Ferriprox n=4 (14%)	Deferoxamine n=11 (34%)	Ferriprox n=25 (86%)	Deferoxamine n=21 (66%)
Median	8.0	2.4	3.0	0.4
Inter-quartile range (Q1, Q3)	(6.8, 11.7)	(0.8, 5.5)	(0.7, 6.0)	(0.1, 1.9)
(Minimum, Maximum)	(6.5, 12.7)	(-0.3, 13.0)	(-2, 8.9)	(-2.1, 17.6)
p-value – Median Test*	0.0226		0.0404	
p-value - Wilcoxon (Rank Sum Test)*	0.0403		0.0112	
p-value - Kolmogrov-Smirnov Test*	0.0516		0.0199	

* Nominal p-values are provided for information only

- **Efficacy by Hepatitis C Status**

Table 22: Efficacy by Hepatitis C Status

MRI T2* (milliseconds)	Difference from Baseline at 12 Months			
	Hepatitis C - Positive		Hepatitis C - Negative	
	Ferriprox n=18 (62%)	Deferoxamine n=16 (50%)	Ferriprox n=11 (38%)	Deferoxamine n=16 (50%)
Median	3.8	1.9	3.1	0.4
Inter-quartile range (Q1, Q3)	(0.9, 7.1)	(0.2, 4.9)	(0.6, 5.9)	(-0.2, 1.8)
(Minimum, Maximum)	(-2, 12.7)	(-1.2, 17.6)	(0.4, 8.7)	(-2.1, 7.1)
p-value – Median Test*	0.3806		0.1900	
p-value - Wilcoxon (Rank Sum Test)*	0.4059		0.0071	
p-value - Kolmogrov-Smirnov Test*	0.4068		0.0485	

* Nominal p-values are provided for information only

- **Efficacy by Serum Ferritin Concentrations (≤ 2,500 µg/L or > 2,500 µg/L)**

Table 23: Efficacy by Serum Ferritin Concentrations (≤ 2,500 µg/L or > 2,500 µg/L)

MRI T2* (milliseconds)	Difference from Baseline at 12 Months			
	Serum Ferritin Concentrations ≤ 2,500 µg/L		Serum Ferritin Concentrations > 2,500 µg/L	
	Ferriprox n=24 (83%)	Deferoxamine n=19 (59%)	Ferriprox n=5 (17%)	Deferoxamine n=13 (41%)
Median	3.8	1.9	2.3	0.4
Inter-quartile range (Q1, Q3)	(0.9, 6.8)	(-0.1, 4.2)	(-0.8, 8.4)	(0.2, 1.9)
(Minimum, Maximum)	(-1.8, 8.9)	(-2.1, 13)	(-2, 12.7)	(-1, 17.6)
p-value – Median Test*	0.0645		0.6091	
p-value - Wilcoxon (Rank Sum Test)*	0.0583		0.4289	
p-value - Kolmogrov-Smirnov Test*	0.2447		0.4686	

* Nominal p-values are provided for information only

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

The statistical issues in this NDA submsission have been discussed in detail in Section 3 of this review. The discussion of collective evidence is not applicable to this application as it contained only one randomized, well-controlled study.

5.2 Conclusions and Recommendations

Deferiprone (Ferriprox) Oral film coated tablets (500mg) is an orally active iron chelator developed for the treatment of iron overload in patients with transfusion-dependent thalassemia.

The proposed indication is the treatment of iron overload in patients with transfusion-dependent thalassemia, and for the treatment of iron overload associated with other transfusion-dependent anemias for whom the use of other iron chelators has been considered inappropriate.

The efficacy and safety of Deferiprone in this New Drug Application (NDA) was based on one main phase 3 (LA16-0102) and one supportive (retrospective) study (LA12-9907) study. This submission included several other supportive and safety studies based on publications.

The primary results of the single randomized trials LA16-0102 are summarized in the Table 24 below:

Table 24: Change in MRI T2* at Month 12 from Baseline - ITT Population

MRI T2* (milliseconds)	Randomized Treatment Groups -- Difference from Baseline at 12 Months	
	Deferiprone (n=29)	Deferoxamine (n=31†)
Analysis - Protocol & SAP identified[1]		
Mean (milliseconds) ± SD	3.9 ± 3.6	2.3 ± 4.1
95% CI	(2.6, 5.3)	(0.8, 3.9)
p-value (unequal variance)	0.0993	
Sponsor's post-hoc analysis-based on log transformation [2]		
Diff (log 12 (or 6) – log base)	0.236703	0.122555
Nominal p-value	0.0228	
Analysis– Non-parametric tests[3]		
Median	3.7	1.0
(10^{th} percentile, 90^{th} percentile)	(-0.1, 8.7	(-1.2, 6.8)
Range (Minimum, Maximum)	(-2.0, 12.7)	(-0.1, 8.7)
p-value – Median Test	0.0048	
p-value -Wilcoxon (Rank Sum Test)	0.0124	
p-value – Kolmogrov-Smirnov Test	0.0168	

[1] The MRI T2* between the Deferiprone and Deferoxamine treatment groups was compared by the two sample t-test assuming unequal variances. The assumptions of normality behind these tests are violated

[2] The Log (MRI T2*) between the Deferiprone and Deferoxamine treatment groups was compared by the two sample t-test by the sponsor.

[3] Since the data are non-normal, non-parametric statistics (Wilcoxon/Kruskal-Wallis Rank Sum tests and the Median Test (Number of Points Above Median), Kolmogorov-Smirnov test) provide valid analyses. The asymptotic

p-value for the Kolmogorov-Smirnov test is 0.0168 (12 months). This indicates rejection of the null hypothesis that the distributions are identical for the two groups.
† Subject C1-40 had baseline MRI T2* level value only and was not eligible to be included in the ITT population.

This single randomized trial LA16-0102 has serious limitations including imaging endpoint of MRI T2*, no "within study" evidence that MRI T2* is reasonably likely to predict meaningful clinical outcome, "primary" analysis questionable due to lack of normality, inadequate safety database of only 29 patients exposed to deferiprone, and the observational study LA 12-9907 not providing independent corroboration due to serious limitations including lack of randomization, no information regarding some important baseline variables such as splenectomy status, and limited information (a lot of missing values) at baseline such as hepatic iron concentration. The submitted data does not support the proposed indications.

SIGNATURES/DISTRIBUTION LIST

Signature List:

Satish C. Misra, Ph. D., Statistical Reviewer

Jyoti Zalkikar, Ph. D., Statistical Team Leader

Rajeshwari Sridhara, Ph.D., Acting Director, Division of Biometrics V

Distribution List:

Hyon-Zu Lee, Project Manager
George Shashaty, M.D., Clinical Reviewer
Kathy M Robie Suh, M.D. -- Clinical Team Leader
Rafel Rieves, M.D. –Division Director, DMIHP
Lillian Patrician, OB

Application Type/Number	Submission Type/Number	Submitter Name	Product Name
NDA-21825	ORIG-1	APOPHARMA INC	FERRIPROX (DEFERIPRONE)

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

/s/

SATISH C MISRA
11/10/2009

JYOTI ZALKIKAR
11/10/2009

RAJESHWARI SRIDHARA
11/10/2009

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

NDA Number: 21-825	**Applicant: ApoPharma Inc**	**Stamp Date: 2-Feb-2009**
Drug Name: Ferriprox (deferiprone) Oral film coated tablets (500mg)	**NDA/BLA Type: eCTD NDA** \\CDSESUB1\EVSPROD\NDA021825\021825.ENX	**Letter Date: 29-Jan-2009**

On **initial** overview of the NDA/BLA application for RTF:

	Content Parameter	Yes	No	NA	Comments
1	Index is sufficient to locate necessary reports, tables, data, etc.	X			
2	ISS, ISE, and complete study reports are available (including original protocols, subsequent amendments, etc.)	X			
3	Safety and efficacy were investigated for gender, racial, and geriatric subgroups investigated.	X			
4	Data sets in EDR are accessible and conform to applicable guidances (e.g., existence of define.pdf file for data sets).	X			

IS THE STATISTICAL SECTION OF THE APPLICATION FILEABLE? YES

If the NDA/BLA is not fileable from the statistical perspective, state the reasons and provide comments to be sent to the Applicant.

Comments:

Able to locate the definition files for datasets, datasets, modules, table of contents, index, protocols, etc., though information is not optimally organized.

Our Guidance document for RTF (http://www.fda.gov/guidance/rtf.pdf) states, "CDER should not, in general, accept for full review applications that can be readily identified as not approvable or non-reviewable because of major flaws or omissions."

There is only one adequate, randomized and well controlled study. There are several supportive, retrospective, safety and efficacy studies based on publications and, data from a variety of types of studies, including clinical and epidemiological. Data may support one indication, but interpolative assessment for other indications is made.

Also, thalassaemia major is a rare disease world-wide. According to the sponsor, the safety profile of Ferriprox has been well characterized over more than 20 years of clinical experience, including almost a decade of post-marketing exposure in Europe and elsewhere.

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

Please identify and list any potential review issues to be forwarded to the Applicant for the 74-day letter.

Content Parameter (possible review concerns for 74-day letter)	Yes	No	NA	Comment
Designs utilized are appropriate for the indications requested.		X		No data for one indication – interpolative statement used
Endpoints and methods of analysis are specified in the protocols/statistical analysis plans.	X			Endpoints not directly related to evaluate the indications being sought
Interim analyses (if present) were pre-specified in the protocol and appropriate adjustments in significance level made. DSMB meeting minutes and data are available.			X	
Appropriate references for novel statistical methodology (if present) are included.			X	
Safety data organized to permit analyses across clinical trials in the NDA/BLA.		X		Most of the safety data comes from epidemiological studies
Investigation of effect of dropouts on statistical analyses as described by applicant appears adequate.		X		Only one prospectively designed randomized study with very small numbers

Comments to be conveyed to the sponsor in 74-day letter

The indication being proposed is for the treatment of iron overload in patients with transfusion-dependent thalassemia; and for the treatment of iron overload associated with other transfusion-dependent anemias for whom the treatment of other iron chelators has been considered inappropriate. Studies LA16-0102 and LA12-9907 enrolled only subjects with thalassemia. Virtually all the other data provided also related to patients with thalassemia. There are no efficacy and safety data provided that supports the indication for the treatment of iron overload associated with other transfusion-dependent anemias for whom the treatment of other iron chelators has been considered inappropriate. Please provide efficacy and safety data in appropriate format that directly relate to the indications being sought. You may also revise the indication to the population in which efficacy and safety have been demonstrated.

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

Brief summary of controlled clinical trials
The following table contains information on the relevant trials contained in the submission.

Study number	Design	Treatment arms/Sample size	Primary endpoint/Analysis	Sponsor's findings
LA16-0102	Prospective Two arm, Comparative, Multicenter, Randomized, Open-label, Active controlled	Ferriprox (Deferiprone) (n=29) vs. Deferoxamine (n = 31)	Myocardial T2* at month 12, consistent with a reduction in cardiac iron concentration,	significantly greater improvement in myocardial T2* at month 12, (27% (Ferriprox) vs. 13%, (Deferoxamine), p=0.023)
LA12-9907	Retrospective assessment (long term > 4 years) of heart failure and survival	Ferriprox (Deferiprone) (n=54) vs. Deferoxamine (n = 75)	heart failure and survival during iron chelation in subjects with transfusion dependent ß-thalassemia.	13% of Ferriprox-treated subjects recipients had New York Heart Association (NYHA)-classifiable cardiac disease, compared to 29% of deferoxamine recipients.
Supportive Epidemiology Study	Retrospective assessment of cardiac events - seven centers in Italy	Ferriprox (Deferiprone) (n=157) vs. Deferoxamine (n = 359)	cardiac events in patients with thalassemia major switched from deferoxamine to Ferriprox therapy.	Significantly lower cardiac events in Ferriprox as compared to Deferoxamine (0%, n=157 in Ferriprox versus 14.5% , n=359 in Deferoxamine $p < 0.001$)

Satish C. Misra, Ph. D. — 18 March 18, 2009
Reviewing Statistician — Date

Jyoti Zalkikar, Ph. D. — 18 March 18, 2009
Supervisor/Team Leader — Date

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

/s/

Satish Misra
3/18/2009 03:16:20 PM
BIOMETRICS

Jyoti Zalkikar
3/24/2009 01:07:20 PM
BIOMETRICS