U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Pharmacoepidemiology and Statistical Science
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/Serial Number:	20-986/SE5-033
Drug Name:	Novolog (Insulin aspart)
Indication(s):	Type 1 diabetes in pediatrics
Applicant:	Novo Nordisk
Date(s):	Submitted 3/15/05
	User Fee goal date 9/15/05
Review Priority:	Priority
Biometrics Division:	Biometrics 2 (HFD-715)
Statistical Reviewer:	J. Todd Sahlroot, Ph.D.
Concurring Reviewers:	S. Edward Nevius, Ph.D.
Medical Division:	Metabolic and Endocrine Drug Products (HFD-510)
Clinical Team:	Eddie Gabry, M.D.
	David Orloff, M.D.
Project Manager:	Julie Rhee

Keywords: clinical studies, NDA review, pediatric exclusivity

Table of Contents

1. EXECUTIVE SUMMARY

1.1 Conclusions and Recommendations

In this 24-week, randomized, multi-center, open-label, active-controlled, parallel-group study in patients aged 6-18 comparing Novolog (insulin aspart) to two other short-acting insulins, Novolin R and Lispro, each in combination with NPH, Novolog was shown to be statistically non-inferior to Novolin R but was not non-inferior to Lispro. Non-inferiority was evaluated using a primary endpoint of HbA1c change from baseline and pre-defined non-inferiority margin of 0.4 (Table 1).

Table 1. HbA1c change from baseline at 24 weeks -- ITT population

HbA1c (%)	Novolog (n=169)	Novolin R (n=81)	LIspro (n=89)
Baseline mean (SD)	8.31 (1.19)	8.38 (1.32)	8.35 (1.17)
Change from baseline			
Mean (SD)	+0.07 (1.02)	+0.15 (1.08)	-0.13 (1.01)
LS mean [1] (SE)	+0.22 (0.12)	+0.42 (0.15)	-0.02 (0.14)
Novolog minus Novolin R			
LS mean [1] (97.5% CI [2])	-0.19 (-0.51, +0.12)		
Novolog minus Lispro			
LS mean [1] (97.5% CI [2])	0.24 (-0.06, +0.54)		

[1] Least square mean based on ANCOVA model with terms for treatment, Hba1C at screening category (≤ 10%, >10%), age category ([6-12), [12-18]), center and baseline HbA1c as covariate
[2] 97.5% CI used instead of 95% CI to control overall Type 1 error for 2 primary comparisons.

The groups were similar in rates of hypoglycemia.

1.2 Brief Overview of Clinical Studies

The submission consisted of a single clinical study (ANA-2126) comparing Novolog (insulin aspart) to two other marketed short-acting insulins, Novolin R and Lispro. Table 2 shows major design characteristics of the study.

Table 2. Study characteristics for ANA-2126

Centers Dates	Patients	# randomized	Design Primary endpoint	Duration
64 US centers 4/02 - 6/04	M and F ages 6-18 with Type 1 diabetes also taking NPH insulin. Screening HbA1c ≤ 12%	Novolog n=187 Novolin R n=96 Lispro n= 95	Randomized [1] active-controlled open label Change from baseline in HbA1c	2 weeks Screening 24 weeks treatment

1 Randomization stratified by screening HbA1c (≤ 10%, >10%) and age category ([6-12), [12-18])

1.3 Statistical Issues and Findings

There were no outstanding statistical issues.

2. INTRODUCTION

2.1 Overview

Study 2126 was conducted in response to an FDA Pediatric Written Request issued on Dec.14, 1999 and (final amendment dated Oct. 5, 2004) which specified the following statistical information in the final amended version:

The analysis of the primary efficacy variable will use a statistical model with the change from baseline HgbA1c as the dependent variable, and treatment and randomization stratification factors as independent variables. Non-inferiority of the test drug compared to control will be assessed by constructing a 97.5% two-sided confidence interval for the between-group difference in change from baseline HgbA1c using the least square means. The test drug will be considered non-inferior to each control if the appropriate confidence bound falls within a non-inferiority margin of 0.4%.

The sponsor's proposed label addressing Study 2126 omits data from the LIspro treatment arm. This issue will be discussed internally at the labeling meeting:

A 24-week, parallel-group study of children and adolescents with type 1 diabetes (n = 283) age 6-18 years compared the following treatment regimens: NovoLog (n = 187) or Novolin R (n = 96). NPH insulin was administered as the basal insulin. NovoLog

achieved glycemic control comparable to Novolin R, as measured by change in HbA1c . The incidence of hypoglycemia was similar for both treatment groups.

2.2 Data Sources

The raw data was located in

\\Cdsesub1\n20986\S_033\2005-03-14\crt\datasets\2126

The Final Report was located in

\\Cdsesub1\n20986\S_033\2005-03-14\clinstat\2126

3. STATISTICAL EVALUATION

Design

2126 was a 24-week, randomized, multi-center, open-label, active-controlled, parallel-group study comparing Novolog to Novolin R or Lispro, in combination with NPH. Novolin NPH was used except in subjects randomized to Lispro who received Humulin NPH which is recognized as equivalent to Novolin NPH. Novolog or Lispro was administered immediately before meals. Novolin R was administered 20 to 30 minutes before meals. NPH was administered before dinner or at bedtime at the Investigator's discretion, and before breakfast mixed with Novolog, Novolin R, or Lispro as needed. Doses of Novolog, Novolin R or Lispro, and doses of NPH were adjusted to achieve specified blood glucose targets that depended on age category.

There were 8 clinic visits: Screening (Week – 2), Randomization (Week 0 (baseline)) and on-treatment visits at Weeks 4, 8, 12, 16, 20 and 24. HbA1c was measured at Weeks -2, 0, 12 and 24.

At 300 completers, the trial had 80% power to demonstrate the non-inferiority of Novolog to Novolin R or Lispro based on a standard deviation of 1.0, a non-inferiority margin of 0.4 and a one-sided alpha of 2.5%. The original sample size was projected at n=336 to include a 10% dropout rate but was increased to n=378 to accommodate the observed 20% dropout rate.

3.1 Evaluation of Efficacy

Demographic / baseline variables

Race, age, sex, weight, height, BMI, duration of diabetes and Tanner Stage were similar between groups for all randomized patients. The mean age was 11.6 years. 75% of patients were Caucasian.

Disposition

Table 3 shows patient disposition by Study visit. A slightly greater percentage of Lispro patients remained on study compared to the other two treatment groups.

Table 3. Number of patients by weeks on study [1]

Last Week on study	Novolog	Novolin R	Lispro	Total
Randomized	187 (100%)	96 (100%)	95 (100%)	378 (100%)
Week 0 (baseline)	186 (99%)	95 (99%)	95 (100%)	376 (99%)
Week 12	170 (91%)	81 (84%)	89 (94%)	340 (90%)
Week 24	150 (80%)	72 (75%)	81 (85%)	303 (80%)
Completers [2]	146 (78%)	74 (77%)	80 (84%)	300 (79%)
ITT	169 (90%)	81 (84%)	89 (94%)	339 (90%)

[1] On-study time defined by time of last HbA1c measurement.
[2] Sponsor's designation

Table 4 shows numbers of patients who discontinued and reasons for discontinuation. The number of Novolog discontinuations was greater due to the 2:1:1 randomization**.**

Table 4. Reasons for discontinuation

	Novolog	Novolin R	Lispro	Total
Non-compliance	26 (63%)	12 (55%)	11 (73%)	49 (63%)
AEs	2 (5%)	3 (14%)	0	5 (6%)
Ineffective therapy	4 (10%)	1 (5%)	1 (7%)	6 (7%)
Other	9 (22%)	6 (26%)	3 (20%)	18 (23%)
Total	41 (100%)	22 (100%)	15 (100%)	78 (100%)

HbA1c

Table 5 shows HbA1c results at 24 weeks for the ITT population. The FDA Written Request specified two primary comparisons, Novolog vs Novolin R and Novolog vs Lispro. Each comparison was to be assessed using a 2-sided 97.5% confidence interval (CI) based on a Bonferroni multiple comparison adjustment for the two comparisons. The Table also shows nominal 95% CIs without any adjustment for multiplicity. Based on a non-inferiority margin of 0.4, Novolog was shown to be statistically non-inferior to Novolin R but not non-inferior to Lispro. Although not required for the purpose of assessing the efficacy of Novolog, Lispro was shown to be statistically superior to Novolin R at the 5% level of significance.

Results were similar for the completers population.

Table 5. HbA1c change from baseline at 24 weeks -- ITT population

	Novolog (n=169)	Novolin R (n=81)	LIspro (n=89)
Baseline mean (SD)	8.31 (1.19)	8.38 (1.32)	8.35 (1.17)
Change from baseline			
Mean (SD)	+0.07 (1.02)	+0.15 (1.08)	-0.13 (1.01)
LS mean [2] (SE)	+0.22 (0.12)	+0.42 (0.15)	-0.02 (0.14)
Novolog minus Novolin R			
LS mean [1] (97.5% CI)	-0.19 (-0.51, +0.12)		
Nominal 95% CI	(-0.47, 0.08)		
Novolog minus Lispro			
LS mean [1] (97.5% CI)	+0.24 (-0.06, +0.54)		
Nominal 95% CI	(-0.02, 0.50)		
Lispro minus Novolin R			
LS mean [1]	-0.44		
Nominal 95% CI	(-0.74, -0.13)		

1 Least square mean based on ANCOVA model with terms for treatment, Hba1C at screening category (≤ 10%, >10%), age category ([6-12), [12-18]), center and baseline HbA1c as covariate

Basal Insulin Dose

Basal (NPH) insulin doses were comparable between treatment groups, suggesting any observed difference in HbA1c should be attributed to the short-acting (randomized) insulins and not the basal insulin. Table 6 shows basal insulin dose in terms of total units (IU) and by age category since NPH doses were adjusted to achieve specified blood glucose targets that depended on age category. The sponsor presented insulin dose (e.g., Table 8-2) normalized by weight (U/kg).

Table 6. Mean daily basal insulin dose (IU) [1]

Mean ± SD (n)	Novolog	Novolin R	Lispro	Test for difference in means (p-value) [2]
Age [6-12)				
Week 0	22.4 ± 10.8 (82)	24.6 ± 11.0 (46)	20.0 ± 11.1 (41)	.15
Week 4	22.4 ± 10.4 (88)	22.3 ± 11.0 (45)	19.9 ± 11.2 (43)	.44
Week 8	22.9 ± 10.9 (87)	23.3 ± 11.8 (40)	21.5 ± 13.2 (41)	.75
Week 12	24.1 ± 11.6 (85)	24.2 ± 12.1 (39)	23.6 ± 15.6 (41)	.98
Week 16	24.8 ± 11.8 (84)	25.4 ± 12.7 (38)	24.9 ± 16.8 (38)	.97
Week 20	24.6 ± 11.8 (82)	26.2 ± 14.0 (39)	25.2 ± 17.0 (36)	.83
Week 24	25.3 ± 11.9 (80)	27.2 ± 14.8 (39)	26.3 ± 18.1 (41)	.79
Age [12-18]				
Week 0	44.0 ± 17.2 (63)	44.5 ± 23.4 (31)	46.6 ± 21.4 (31)	.84
Week 4	43.7 ± 18.0 (79)	39.6 ± 20.1 (43)	42.7 ± 19.7 (46)	.52
Week 8	46.0 ± 18.2 (71)	41.5 ± 20.0 (43)	44.4 ± 20.8 (43)	.50
Week 12	48.0 ± 19.3 (70)	44.2 ± 23.1 (38)	44.2 ± 20.8 (41)	.54
Week 16	49.0 ± 20.1 (67)	46.3 ± 25.3 (37)	48.6 ± 22.1 (43)	.83
Week 20	50.4 ± 20.6 (61)	46.4 ± 23.8 (36)	48.3 ± 23.4 (40)	.70
Week 24	51.8 ± 22.2 (70)	46.5 ± 24.5 (39)	49.1 ± 23.4 (42)	.51

1 mean = average of the 2-day measurements
2 p-value from ANOVA of mean daily insulin dose with treatment as factor in the model

3.2 Evaluation of Safety

Hypoglycemia

Hypoglycemic episodes were categorized into mutually exclusive categories as (1) minor, (2) major or (3) symptomatic only. The sponsor's diagram below shows the criteria used to categorize a hypoglycemic episode. Major episodes required assistance by a third party.

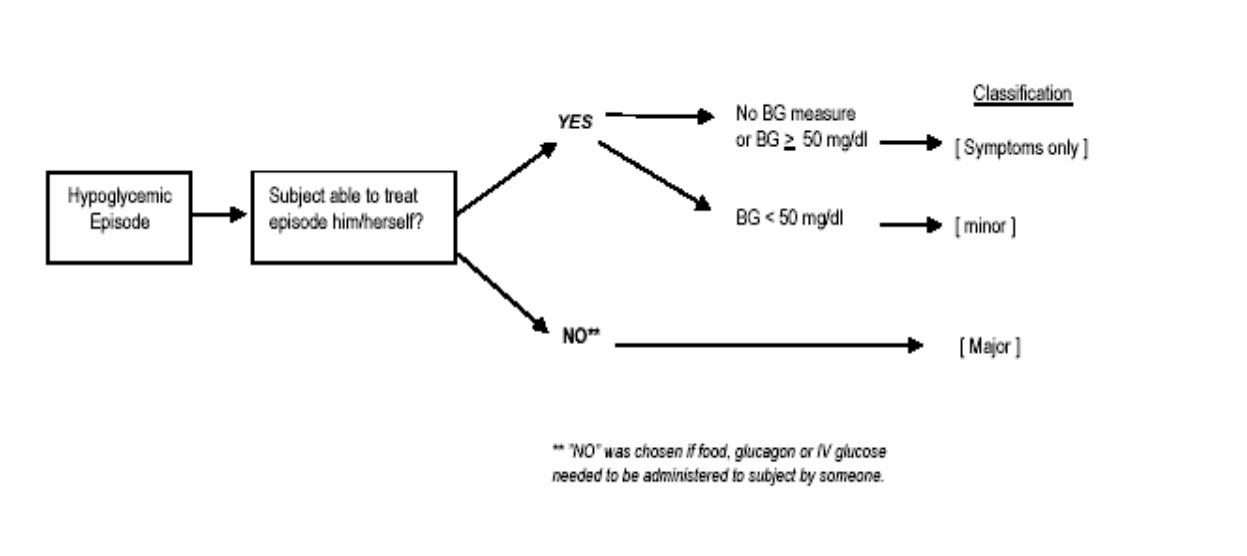

Table 7 shows the number and rate of the different types of hypoglycemia after randomization. The Table is similar to the sponsor's Final Report Table 10-7 except (1) median rates are reported instead of means due to the skewed distributions for the rates and (2) the method used to calculate individual patient annual rates is different. The sponsor calculated annualized rates as (12/ # months treated) x (hypo count). This reviewer calculated rates as 365.25/(# days treated) x (hypo count).

Frequencies and rates were compared between the groups for all hypoglycemia types. Only the Novolog vs Novolin R comparison of rates of symptomatic-only episodes was nominally significant (Wilcoxon, p=.044). These results are not definitive; reported data on hypoglycemia in this open-label trial may be subject to reporting bias that would be minimized in a double-blind design.

Table 7. Hypoglycemic episodes

Hypo type	Novolog N=187		Novolin R N=95 [1]		Lispro N=95	
	N (%)	Med rate	N (%)	Med rate	N (%)	Med rate
All	159 (85%)	28.3	84 (88%)	25.9	81 (85%)	30.1
BG<36 [2]	75 (40%)	0	41 (43%)	0	32 (34%)	0
Major	11 (6%)	0	9 (9%)	0	7 (7%)	0
Minor	147 (79%)	12.4	82 (86%)	17.2	75 (79%)	10.0
Sym only	120 (64%)	4.4	51 (54%)	2.1	61 (64%)	5.5

1 One Novolin R patient did not receive study drug
2 BG < 36 includes episodes with blood glucose < 36 mg/dL or that were categorized as major

Tanner Stage

The sponsor's assessment of Tanner stage data (Final Report, p.55) was:

In general, the percentages of subjects in Tanner Stage 1 at the end of the study tended to decrease from baseline percentages for all treatment groups and for both males and females. Similarly, percentages of subjects in Tanner Stage V at the end of the study tended to increase in all treatment groups for both males and females. No notable differences were observed between treatment groups.

This reviewer found the Tanner Stage (pubic hair) data presented a more complex picture than that described by the sponsor. Tables A1, A2 and A3 in the Appendix show the numbers of patients (males and females) with Tanner Stage transitions (i, j) where i = Baseline Stage and j = Week 24 Stage. Table 8 below shows the numbers of patients in each group with advances in Tanner Stage. Only patients with baseline Tanner Stage 4 or less were included in the analysis. An advance in Tanner Stage was defined as one or more increases in Tanner Stage over the baseline Tanner Stage, e.g., from Stage 2 at baseline to Stage 3 at Week 24 or Stage 3 at baseline to Stage 5 at Week 24. Based on the computation of nominal p-values, Novolog showed a borderline statistically significant greater increase in the number of patients with increases of one or more Tanner Stages (pubic hair) during the trial compared to Novolin R and Lispro (36% vs 25% and 21%, p=.057). Pair-wise comparisons of Novolog with Novolin R and Lispro yielded p-values of .11 and .032, respectively.

Table 8. Tanner Stage (pubic hair) at Week 24 (males and females)

	Novolog N=142	Novolin R N=65	Lispro N=70	Total N=277
Advance in Tanner stage	51 (36%)	16 (25%)	15 (21%)	82 (30%)
No advance in Tanner Stage	91 (64%)	49 (75%)	55 (79%)	195 (70%)
Chi-square (all 3 groups)	P=.057			
Chi-square (pair-wise with Novolog)		P=.11	P=.032	

There was no difference between the groups in Tanner Stage advancement for breast development (Table 9).

Table 9. Tanner Stage breast development at Week 24

	Novolog N=70	Novolin R N=28	Lispro N=35	Total N=133
Advance in Tanner stage	22 (31%)	8 (29%)	10 (29%)	40 (30%)
No advance in Tanner Stage	48 (69%)	20 (71%)	25 (71%)	93 (70%)
Chi-square comparing all 3 groups p=0.94				

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race and Age

There were no significant interactions (p>.10) between treatment and sex or age category (Table 10).

The treatment-by-race interaction p-value was significant at the 10% level (p = 0.065). The nominally significant result was driven by the Lispro responses in the two smallest subgroups, Asian and Other, which showed large increases from baseline relative to modest changes in the Novolog treatment groups. The result did not hold up when the 15 patients belonging to these two subgroups were removed from the analysis. The interaction p-value in this analysis was not significant (p=0.34).

Table 10. Descriptive statistics for HbA1c by required subgroup

	Novolog (n=169)	Novolin R (n=81)	LIspro (n=89)
Males	(n=94)	(n=36)	(n=39)
Baseline mean	8.37	8.87	8.14
Mean change from baseline	+0.14	+0.30	+0.15
Females	(n=75)	(n=45)	(n=50)
Baseline mean	8.24	8.39	8.52
Mean change from baseline	-0.01	+0.03	-0.34
Age 6-12	(n=88)	(n=40)	(n=43)
Baseline mean	8.07	8.26	8.26
Mean change from baseline	+0.03	+0.00	-0.32
Age 12-18	(n=81)	(n=41)	(n=46)
Baseline mean	8.55	8.50	8.44
Mean change from baseline	+0.12	+0.29	+0.06
Caucasian	(n=130)	(n=56)	(n=70)
Baseline mean	8.29	8.19	8.24
Mean change from baseline	+0.06	+0.11	-0.08
Black	(n=16)	(n=13)	(n=9)
Baseline mean	8.50	9.05	8.46
Mean change from baseline	+0.25	+0.54	-0.37
Hispanic	(n=15)	(n=7)	(n=8)
Baseline mean	8.34	8.52	9.14
Mean change from baseline	+0.07	-0.19	-0.69
Asian	(n=2)	(n=2)	(n=1)
Baseline mean	8.45	7.90	7.80
Mean change from baseline	-0.35	+0.45	+0.60
Other	(n=6)	(n=3)	(n=1)
Baseline mean	8.20	8.97	9.60
Mean change from baseline	+0.13	-0.40	+2.60

4.2 Other Special/Subgroup Populations

The interaction between treatment and screening HbA1c (<10%, ≥ 10%) was not statistically significant (p=0.78) (Table 11).

Table 11. Descriptive statistics for HbA1c by screening HbA1c

	Novolog (n=169)	Novolin R (n=81)	LIspro (n=89)
Screening HbA1c <10.0	(n=150)	(n=70)	(n=77)
Baseline mean (SD)	8.05 (0.94)	8.00 (0.94)	8.07 (0.95)
Mean change from baseline	+0.13	+0.25	-0.09
Screening HbA1c <10.0	(n=19)	(n=11)	(n=12)
Baseline mean (SD)	10.41 (0.85)	10.82 (0.94)	10.16 (0.77)
Mean change from baseline	-0.34	-0.54	-0.38

APPENDICES

Table A1. Tanner Stage pubic hair transitions – Novolog treatment group

Baseline Tanner Stage	Week 24 Tanner Stage [1]					
	1	2	3	4	5	Total
1	55	**15**	**4**	0	0	74
2	1	7	**8**	**1**	0	17
3	0	0	6	**9**	0	15
4	0	0	0	22	**14**	36
5	0	0	0	0	37	37
Total	56	22	18	32	51	179

1 Each cell represents the number of patients with the given baseline and Week 24 Tanner stages. Numbers in bold indicate advances in Tanner stage after baseline

Table A2. Tanner Stage pubic hair transitions – Lispro treatment group

Baseline Tanner Stage	Week 24 Tanner Stage [1]					
	1	2	3	4	5	Total
1	29	**6**	0	0	0	35
2	0	7	**1**	**1**	0	9
3	0	0	8	**3**	**1**	12
4	0	0	0	11	**3**	14
5	0	0	0	0	19	19
Total	29	13	9	15	23	89

1 Each cell represents the number of patients with the given baseline and Week 24 Tanner stages. Numbers in bold indicate advances in Tanner stage after baseline

Table A3. Tanner Stage pubic hair transitions – Novolin R treatment group

Baseline Tanner Stage	Week 24 Tanner Stage [1]					
	1	2	3	4	5	Total
1	26	**2**	0	0	0	28
2	0	6	**5**	**2**	0	13
3	0	0	6	**3**	**1**	10
4	0	0	0	11	**3**	14
5	0	0	0	2	17	19
Total	26	8	11	18	21	84

1 Each cell represents the number of patients with the given baseline and Week 24 Tanner stages. Numbers in bold indicate advances in Tanner stage after baseline

---

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

/s/

Todd Sahlroot
9/2/2005 09:55:09 AM
BIOMETRICS