



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: 21-868

Drug Name: EXUBERA (insulin [rDNA origin] powder for oral inhalation)

Indication(s): Treatment of adult patients with Type 1 or Type 2 diabetes mellitus for the control of hyperglycemia

Applicant: Pfizer Global Research and Development

Date(s): Submitted December 27, 2004

Review Priority: Standard

Biometrics Division: Division of Biometrics 2 (HFD-715)

Statistical Reviewer: Joy Mele, M.S.

Concurring Reviewers: Todd Sahlroot, Ph.D.
Biometrics Team Leader
Ed Nevius, Ph.D.
Biometrics Division Director

Medical Division: Division of Metabolic and Endocrine Drug Products (HFD-510)

Clinical Team: Karen Mahoney, M.D.
Clinical Reviewer
David Orloff, M.D.
Division Director

Project Manager: Oluchi Elekwachi

Keywords: Clinical studies, NDA, active control

1. EXECUTIVE SUMMARY OF STATISTICAL FINDINGS	5
1.1 Conclusions and Recommendations	5
1.2 Brief Overview of Clinical Studies	6
1.2.1 Phase 2/3 clinical trials in patients with Type 1 diabetes	6
1.2.2 Phase 2/3 clinical trials in patients with Type 2 diabetes	7
1.3 Statistical Issues	9
2. INTRODUCTION	10
2.1 Overview	10
2.2 Data Sources	10
3. STATISTICAL EVALUATION	11
3.1 Evaluation of Efficacy	11
3.1.1 Clinical Trials in Patients with Type 1 Diabetes	11
3.1.1.1 Studies 106 and 107	11
Design	11
Patient Disposition	12
Baseline Demographics and Medical History	13
Dosing	14
Statistical Methods	16
Efficacy Results	17
3.1.2 Clinical Trials in Patients with Type 2 Diabetes	20
3.1.2.1 Overview of Phase 3 trials in Patients with Type 2 Diabetes	20
3.1.2.2 Efficacy Results for Phase 3 trials in Patients with Type 2 Diabetes	23
Statistical Methods	24
Study 108 Efficacy Results	24
Study 109 Efficacy Results	25
Study 110 Efficacy Results	25
Studies 1001 and 1002 Efficacy Results	26
Summary Graphs of Efficacy in all 5 studies in Type 2 diabetic patients	27
3.2 Evaluation of Safety	28
Hypoglycemia in Studies 107 and 108	28
Antibody levels and dosing	35
Antibody levels and hypoglycemia	35
4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS	37
4.1 Gender, Race and Age	37
4.2 Other Special/Subgroup Populations	37
4.2.1 Children	37

4.2.2 Baseline HbA1c	38
5. SUMMARY AND CONCLUSIONS	40
5.1 Statistical Issues and Collective Evidence	40
5.2 Conclusions and Recommendations	43
6.0 APPENDICES	44
6.1 Ongoing safety studies	44
6.2 Average daily dose of long-acting insulin by study, treatment and week for Type 1 studies 106 and 107	45
6.3 Average daily ratio of long-acting to short-acting insulin by study, treatment and week for Type 1 studies 106 and 107	46
6.4 Baseline FPG and HbA1c for Type 1 Studies 106 and 107	47
6.5 Oral anti-diabetic therapies used in Phase 3 studies of Type 2 diabetes	48
6.6 Type 2 FPG overtime all studies	49
6.7 Hypoglycemic events for all Phase 3 Studies	50
6.8 SAS code for modeling hypoglycemic events	50
6.9 Definition of hypoglycemia	51
6.10 Boxplot of antibodies at Week 24 by diabetes type and treatment	52

1. EXECUTIVE SUMMARY OF STATISTICAL FINDINGS

1.1 Conclusions and Recommendations

This reviewer has the following comments and conclusions based on the statistical review of the seven Phase 3 trials submitted to demonstrate the efficacy and safety of INH (Exubera). Note that Figures 3.1.3.1 and 3.1.3.2 on page 39 summarize the efficacy data.

- Approximately 85 to 90% of the Type 1 and 2 patients studied were Caucasian; so the races were not adequately represented in the database
- INH was shown to be non-inferior to SC insulin in patients with Type 1 diabetes in two Phase 3 clinical trials, 106 and 107 (see Table 3.1.1.5 and Figures 3.1.3.1 and 3.1.3.2). The HbA1c results are more favorable to INH in Study 107, a study where SC insulin was given TID and the same long-acting insulin was given in both groups. There was no evidence that long-acting insulin was titrated differently in the two groups allaying concerns that the long-acting dosing may have compensated for inadequate treatment on short-acting insulin.
- For Type 1 patients, statistically significant differences in favor of INH over SC were seen for the change from baseline in FPG in both studies. An examination of 24-hour diary records suggest a drop in glucose levels overnight in the INH group may explain the significantly lower FPG. This issue is discussed in the clinical review by the FDA medical reviewer, Dr. Mahoney.
- Across the Type 2 studies, variability in several baseline parameters was seen suggesting some heterogeneity among the patient populations which could improve generalizability of the results.
- For Type 2 patients inadequately treated with metformin or a sulfonylurea, add-on INH yielded HbA1c lowering comparable to adding either glibenclamide or metformin, respectively (Studies 1001 and 1002, see Table 3.1.2.2.8 and Figure 3.1.2.2.1). For patients inadequately treated with two oral agents, adding INH resulted in highly significant drops in HbA1c (Study 109, see Table 3.1.2.2.6 and Figure 3.1.2.2.1).
- Type 2 patients on SC insulin therapy were able to maintain their HbA1c levels when switched to INH insulin (Study 108, see Table 3.1.2.2.5 and Figure 3.1.2.2.1)
- Naïve Type 2 patients randomized to either rosiglitazone (4 mg BID, the most effective marketed dose) or INH showed statistically significantly more lowering of HbA1c and FPG on INH in Study 110 (Table 3.1.2.2.7) ; however the length of the trial at 12 weeks provided inadequate time for rosiglitazone to show a full effect (Figure 3.1.2.2.1). This reviewer concludes that Study 110 was inadequate by design and the results should not be included in labeling.
- Antibody counts were significantly higher in INH patients than in patients treated with SC insulin or oral agents. This reviewer found no relationship between change in antibody count and change in dose. A crude analysis of antibody count by severity of hypoglycemia suggests that higher levels of antibodies may be associated with more severe levels of hypoglycemia, although antibody level does not appear to be a strong predictor of moderate to severe hypoglycemia in patients treated with INH (see pages 32 to 34 for more details.)
- The survival analysis model for recurrent events used by the applicant to estimate risk ratios for hypoglycemia is not an appropriate model because it assumes that the events are independent which is clearly not the case for clusters of recurrent events within patients. No conclusions can be drawn based on the risk ratios provided by the applicant. For a discussion of this issue for Studies 107 and 108, see pages 27-34.

Since Exubera will be discussed at an upcoming advisory committee meeting, no

recommendations regarding approval based on the statistical evidence of efficacy are being made at this time.

1.2 Brief Overview of Clinical Studies

1.2.1 Phase 2/3 clinical trials in patients with Type 1 diabetes

A total of five Phase 2/3 trials plus one long term extension study were conducted in patients with Type 1 diabetes (Table 1.2.1.1). Phase 3 studies 106 and 107 are adequately designed to provide data to support the efficacy of INH compared to SC and are reviewed in detail here. Study 102 is a small Phase 2 study of short duration and was not reviewed.

Table 1.2.1.1 Phase 2/3 Open-Label, Randomized, Parallel-group Completed Clinical Trials

Study (# of centers)	Inhaled Insulin group	SC Insulin group	Duration of treatment
Phase 2 Studies			
102 (10 US) centers also in 106 and 107	Pre-meal TID + bedtime SC Ultralente 35 rand. 100% completed	Usual SC regimen (BID or TID) 37 rand. 95% completed	4 wk run-in: usual SC regimen 12 wks randomized treatment
Phase 3 Studies			
106 (33 US, 8 Can.) scr and base HbA1c 6%-11%	Pre-meal TID + bedtime SC Ultralente 170 rand. 89% completed	SC regular insulin (BID) +NPH ins 165 rand. 92% completed	4 wk run-in: BID SC reg +NPH ins 24 wks randomized treatment
107 (32 US, 8 Can.) scr and base HbA1c 6%-11%	Pre-meal TID + NPH ins AM and bed 163 rand. 94% completed	SC regular insulin (TID) + NPH ins AM and bed 165 rand. 92% completed	4 wk run-in: pre-meal SC reg +BID NPH ins 24 wks randomized treatment
111 LT Safety Extension of 106, 107, 108, 109, 110, 1009	All pts switched to INH 664 Type 1 patients Included pul fn tests	NA	Most pts had more than 18 months of extended trt -- up to 3 yrs
1026 dose-finding, pharmacology No data in EDR	TID INH +BID NPH 24 rand.	TID reg ins +BID NPH 23 rand.	4 wk run-in SC 24 wks post-prandial glucose
1027 (US, Can, Braz) Safety study Pulm. Fn. tests No data in EDR	Pre-meal TID + QD or BID ultralente or NPH, or QD ins glargine 110 rand. 84% complet. 24 wks	BID or TID reg ins + QD or BID ultralente or NPH, or QD ins glargine 116 rand. 84% complet. 24 wks	3 wk run-in SC short-acting ins 12 wk randomized treatment 12 wk follow-up SC short-acting ins Efficacy data was collected

Results from bolded studies are included in the labeling. Sample sizes are total number of patients including both adult and pediatric patients.

Study 1026, a small PD/PK trial, included standardized meal challenges as well as euglycemic

clamp studies. The applicant concluded that there was no difference in post-prandial control between INH and SC. No study report was provided for Study 1026 but information was provided in clinical summaries. This study is not reviewed here because there is insufficient HbA1c data to draw definitive conclusions on efficacy.

Study 111, a long-term uncontrolled safety study of both Type 1 and Type 2 patients, and Study 1027, a 12 week pulmonary function study, were reviewed by the FDA safety reviewers, Dr. Buenconsejo, Dr. Mahoney and Dr. Seymour.

An additional study (Study 1009) was conducted in children aged 6-11 years with Type 1 diabetes (total of 121 patients; 61 INH and 60 SC). An indication for pediatric patients is not being sought at this time so this study is not reviewed here although the applicant's results from this study are briefly stated in Section 4.2.1 (page 36).

[This space purposely left blank.]

1.2.2 Phase 2/3 clinical trials in patients with Type 2 diabetes

The applicant has conducted two Phase 2 studies and five Phase 3 studies in patients with Type 2 diabetes (Table 1.2.2.1 on following page). Only the Phase 3 studies (108, 109, 110, 1001 and 1002) are reviewed here because they have adequate size and duration to provide data to establish efficacy. The results of Study 111, the long term safety study, are included in safety analyses performed by this reviewer and by the FDA safety reviewers (Drs. Mahoney, Seymour and Buenconsejo).

Table 1.2.2.1 Phase 2/3 Open-Label, Randomized, Parallel-group Clinical Trials

Study (# of centers)	Inhaled Insulin group	Insulin or oral agent group	Duration of treatment
Phase 2 Studies			
103 (US 10) No data in EDR	pre-meal INH +bedtime ultralente 28 rand. 89% completed	BID or TID SC (dose, not freq., could be adjusted) 28 rand. 93% completed	4 wk run-in SC ins 2-day inpt instructions 12 wk rand. treatment
104 (US 9) add-on trial No data in EDR	INH+oral agent 33 rand. 100% completed	oral agent 36 rand. 100% completed	4 wk run-in usual oral agent (sulf and/or met) 2-day inpt instructions 12 wk rand. treatment
Phase 3 Studies			
108 (US and Canada 49)	pre-meal INH +bedtime ultralente 149 rand. 89% completed	SC BID reg + NPH 150 rand. 93% completed	4 wk run-in SC ins 12 wk rand. treatment
109 (US and Canada 52)	Arm 1= pre-meal INH 105 rand. 92% completed Arm 2= pre-meal INH+2 OA's 102 rand. 97% completed	Arm 3= 2 OA's SU or repaglinide + glitazone or MET 102 rand. 91% completed	4 wk run-in 2 OA's 12 wk rand. treatment
110 (US 40)	pre-meal INH 76 rand. 93% completed	ROSI 4 mg BID 67 rand. 91% completed	4 wk run-in diet+exer 24 wk rand. treatment
1001 (Eur, Scan, UK, Israel, Brazil, S. Africa 73)	pre-meal INH 225 rand. 92% completed	Metformin Add-on to SU 202 rand. 88% completed	4 wk run-in metformin 24 wk rand. treatment
1002 (Eur, Scan, UK, Israel, Brazil, S. Africa 77)	pre-meal INH 243 rand. 90% completed	Glibenclamide Add-on to MET 233 rand. 88% completed	4 wk run-in glib. 24 wk rand. treatment
111 LT safety <i>Extension of 106, 107, 108, 109, 110, 1009</i>	All pts switched to INH 626 Type 2 patients Included pul fn tests	NA	Most pts had more than 18 months of extended trt – up to 3 yrs

Results from bolded studies are included in the labeling.

Each of the five Phase 3 trials has a unique design in terms of comparator arm and the population studied. Based on the results of these trials, the applicant is seeking a broad indication for use in patients with Type 2 diabetes stating that “Exubera can be used as monotherapy or in combination with oral agents or longer-acting insulins.”

At the time of the submission of the NDA, several safety studies were ongoing. These trials are listed in a table in Appendix 6.1. Since the focus of this review is primarily efficacy, the ongoing trials are not reviewed here.

1.3 Statistical Issues

The following statistical issues arose during the process of this review:

- Randomization procedures described in the protocols and in the study reports were not the procedures actually carried out. Patients were not assigned using a minimization algorithm but instead were randomized centrally blocking on center (see page 15 for more details).
- All trials in the Exubera submission were open-label trials. The lack of blinding introduces the possibility of bias in several aspects of the trial. The first aspect considered by this reviewer was the enrollment of patients. The lack of blinding can lead to selection bias in that the randomization code could be broken based on the patients already entered in the trial. An investigator may be able to guess the treatment assignment for the next patient. Guessing is more difficult if the randomization is carried out from a central office, if stratification is not done by center and if block sizes are unknown by the investigator (i.e. not mentioned in the protocol). The first and third conditions were in place in the Exubera trials. Stratification was done by center so it is possible that the pattern of treatment assignment could be discerned and that assignment would be predictable for some patients. To test if selection bias was an issue this reviewer performed a test described by Berger and Exner on the data from Study 107 (the most important Type 1 study). Basically this test determines whether the probability of having a good response is related to the probability of being assigned to the test drug; this is essentially testing whether patients with a good prognosis are more likely to have been assigned to Exubera. This reviewer found no evidence of selection bias based on the results of this test.
- The lack of blinding can also bias the measurement of both efficacy and safety measures. HbA1c is an objective measure not likely to be affected by knowledge of treatment, particularly since dosing of patients was well-controlled by specific parameters spelled out in the protocols and inspection of the data suggests that dosing was adjusted as would be expected. There is some evidence that safety measures may have been affected by the lack of blinding. More specifically, the medical reviewer, Dr. Karen Mahoney, carefully describes the misclassification of discontinuation reasons as “withdrawn consent” or “subject request” where the data suggests that the reason was “adverse event”. She found a larger number of misclassifications in the INH group than in the comparator group suggesting the possibility that knowledge of treatment could have played a role in the naming of the discontinuation reason.
- The 12-week duration of Study 110 was insufficient to provide a fair comparison of rosiglitazone to INH (see page 26).
- Patients who discontinue from therapy due to hypoglycemia may provide HbA1c LOCF data that suggests a beneficial result when in reality the therapy was a failure for that patient. This reviewer examined the data to determine if the LOCF estimates were biased by the use of such dropout data and found no evidence that this was the case.
- Noninferiority trials were powered to rule out a treatment difference for HbA1c change from baseline of 0.5% while 0.4% is the margin usually used by FDA. This, however, was not an issue since the boundary of 0.4% was met in all relevant trials.
- The Type 2 development program consisted of five Phase 3 trials; 3 conducted in North America and 2 conducted in foreign countries. The majority of the investigator sites participated in more than one study. For Studies 1001 and 1002 (foreign studies) and for Studies 108, 109 and 110, about half of the total sites participated in more than 1 study (see page 19 for additional details). Patients were not allowed to participate in more than one study and, in addition, entry criteria differed making patients ineligible for more than one study so this reviewer was not concerned that patients may have been retested. Enrollment dates were overlapping for the studies so investigators would be seeing patients during the same timeframe from the different studies suggesting that experience from participating in one trial would not carryover and affect the conduct of a subsequent trial. Also the sample size of

each site is small and no one site would greatly influence the outcome. Though overlapping sites could affect the independence of the trial results, this reviewer did not feel, for these studies, that independence was comprised for the reasons given.

- The applicant has computed risk ratios for hypoglycemia using a recurrent events proportional hazards model that assumes events are independent. This reviewer does not think that the applicant's model is appropriate primarily because it ignores the dependency among events within patients. Also additional analyses suggest that the model is oversensitive to patients with many events; for example, dropping just one patient from an analysis of Study 107 data changed the estimate of risk from 2.25 to 1.65. This reviewer did not have sufficient time to determine the best model to describe the hypoglycemia data; however, an odds ratios for at-least one event and a comparison of number of events (e.g. Wilcoxon rank sum test) seems to describe the data well if it is agreed that time is not an important factor.

2. Introduction

2.1 Overview

Exubera is an inhaled insulin product combining a dry powder formulation of a recombinant human insulin with a customized inhalation system. At the time of this application, insulin was only available via subcutaneous (SC) injection. Exubera, then, offers an alternative non-invasive route of administration. Exubera, according to the applicant, is "rapid-acting with a faster onset of action than SC regular insulin" and so it can replace pre-meal insulin. The Phase 1 and 2 studies in the development program used a 20% insulin formulation while the Phase 3 studies and stability studies used a 60% insulin formulation. The inhaler also was modified between the Phase 2 and Phase 3 studies.

The focus of this review is primarily on efficacy in the Phase 3 trials. This reviewer did not perform a full safety review; safety is addressed by three FDA reviewers in their individual reviews (Drs. Mahoney, Seymour and Buenconsejo). Safety issues addressed in Section 3.2 arose from consultations with the clinical reviewer, Dr. Mahoney.

The applicant is seeking an indication in adults only, based on a recommendation from FDA. However, children were included in some of the Phase 3 trials. For those trials, the review is of the adult patients; results for children are briefly summarized in Section 4.2.1 on page 36 of the review.

2.2 Data Sources

The applicant provided a Common Technical Document and datasets electronically. The address for the SAS datasets in the FDA Electronic Document Room is \\Cdsesub1\n21868\N_000\2004-12-27\crt.

Study reports were well-organized and easy to navigate with essentially the same structure used for each report making it easy for the reviewer to quickly find information from several studies. Datasets were also well-organized although some data (such as detailed explanations of why patients discontinued) were not available.

The applicant responded adequately to requests, with additional information submitted to the FDA Electronic Document Room.

3. Statistical Evaluation

3.1 Evaluation of Efficacy

3.1.1 Clinical Trials in Patients with Type 1 Diabetes

3.1.1.1 Studies 106 and 107

Both adult and pediatric patients were enrolled in the studies of Type 1 diabetics. The focus for the review is on adult patients as previously mentioned.

Design

Two phase 3 clinical trials (Studies 106 and 107) were conducted in patients with Type 1 diabetes to assess the efficacy and safety of Exubera (for descriptions of other trials conducted in Type 1 patients see Table 1.2.1.1). Both trials were 6-month, open-label, out-patient trials comparing Exubera to subcutaneous human insulin therapy.

In Study 106, all patients were administered twice daily (BID) regular insulin SC plus BID SC NPH insulin given pre-breakfast and bedtime, during a lead-in period of 4 weeks. Eligible patients were randomized to one of the following treatment groups:

1. pre-meal (TID) inhaled insulin (INH) plus single Ultralente SC injection at bedtime
2. twice daily (BID) regular insulin SC plus BID SC NPH insulin; patients with a different pre-study regimen could receive dosing comparable to their prestudy dosing (e.g. pre-meal insulin)

In Study 107, all patients were administered pre-meal (TID) regular insulin SC plus BID SC NPH insulin given pre-breakfast and bedtime, during a lead-in period of 4 weeks. Eligible patients were randomized to one of the following treatment groups:

1. pre-meal (TID) inhaled insulin (INH) plus BID SC NPH insulin
2. pre-meal (TID) regular insulin SC plus BID SC NPH insulin

Inclusion criteria included the following:

- Diagnosis of Type 1 diabetes for more than one year
- Aged 12 to 65
- 2 months on a stable insulin regimen involving at least 2 daily injections
- Screening (Week -4) and pre-randomization (Week -1) HbA1c of 6% to 11%, inclusive
- Fasting plasma C-peptide ≤ 0.2 pmol/ml
- BMI ≤ 30

Exclusion criteria included the following:

- Use of insulin pump therapy within 2 months of screening
- History of severe hypoglycemia
- Significant respiratory disease
- Smoker within 6 months prior to screening
- Abnormal pulmonary function tests at Week -3
- Significant major organ system disease

HbA1c was measured on therapy at Weeks 6, 12 and 24. The primary efficacy endpoint was

HbA1c change from baseline at Week 24. Baseline was computed as the average of Weeks -1 and 0.

Secondary variables included FPG, % of patients achieving HbA1c below 7% or 8%, body weight, pre and post prandial glucose values, serum lipids and a measure of quality of life (Phase V Quality of Life (QOL) Treatment Satisfaction Questionnaire).

Patient Disposition

In each study, the plan was to enroll 160 patients in each treatment group, including pediatric patients; so the trial was powered with the inclusion of the pediatric patients. About 1/5 of the Study 106 patients and about 1/3 of the Study 107 patients were under aged 18; these patients are excluded from the tables in this review. Neither study then was adequately powered for only adult patients based on the parameters¹ used to estimate sample size.

A total of 273 adult patients were enrolled in Study 106 and a total of 208 in Study 107 (Table 3.1.1.1). Patients in Studies 106 and 107 were recruited at 41 and 40 centers, respectively in the United States and Canada with the majority of the patients (about 78%) in US centers. Sample sizes in each center ranged from 2 to 20 with most centers enrolling 10 or fewer patients.

Table 3.1.1.1 Patient Disposition

	Study 106		Study 107	
	INH	SC	INH	SC
Randomized	137	136	103	105
Never treated	0	1	1	0
Wk 6	134 (98%)	132 (97%)	102 (100%)	103 (98%)
Wk 12	127 (93%)	128 (94%)	102 (100%)	100 (95%)
Completers	120 (88%)	124 (91%)	97 (94%)	96 (91%)

In both studies, more than 90% of the patients completed 24 weeks on study; there is no notable difference between the treatment groups regarding completion rates.

Three INH patients discontinued from Study 106 due to ADE (Table 3.1.1.2); 2 due to severe hypoglycemia (Weeks 4-5) with one patient experiencing a diabetic coma plus one due to mild coughing (Week 9). One SC patient discontinued from Study 106 due to accidental injury (ADE at Week 8). In Study 107, ADE's included one INH patient who discontinued during Week 6 due to pulmonary obstruction and one SC patient who discontinued due to carcinoma.

Table 3.1.1.2 Reasons for discontinuation

	Study 106		Study 107	
	INH (n=137)	SC (n=136)	INH (n=103)	SC (n=105)
ADE	3 (2.2%)	1 (0.7%)	1 (0.9%)	1 (0.9%)
Lack of Eff.	2 (1.5%)	2 (1.5%)	0	0
Lost-to-FU	2 (1.5%)	2 (1.5%)	2 (1.9%)	1 (0.9%)
Pt request	7 (5.1%)	6 (4.4%)	1 (0.9%)	6 (5.7%)
Other	3 (2.2%)	1 (0.7%)	1 (0.9%)	1 (0.9%)
Abn. Lab	0	0	1 (0.9%)	0

The impact of the dropouts on assessment of efficacy when using LOCF was examined. There

¹ Sample size was estimated assuming, power of about 90%, no treatment difference, standard deviation of 1.2 to 1.5% for change in HbA1c, 2-sided 95% confidence interval and noninferiority margin of 0.5%.

was only one discontinued patient in the INH group who had a large drop in HbA1c carried forward; the large response was due to hypoglycemia. The impact of this one patient on the results is negligible and this reviewer did not consider alternative analyses to address the problem of carrying-forward what appears to be a beneficial result coupled with an adverse event leading to dropout.

Baseline Demographics and Medical History

The treatment groups were adequately balanced on demographic characteristics (Table 3.1.1.3). The average age of patients was 38 years (patients under 18 are excluded from the table). No patients in either trial were 65 years or older. About 5% of the patients had an BMI>30 although the cutoff for entry was 30. The majority of patients were white and had never smoked.

Table 3.1.1.3 Baseline Demographics

	Study 106		Study 107	
	INH (n=137)	SC (n=136)	INH (n=103)	SC (n=105)
Age				
Mean (SD)	38 (10)	38 (10.5)	38 (11)	38.5 (11)
Range	20-63	20-64	19-65	19-65
% >50	14%	14%	16%	17%
Gender				
% female	49%	47%	48%	44%
Race				
% White	88%	95%	87%	94%
% Hispanic	7%	3%	5%	4%
% Black	3.7%	1.5%	3%	1%
% Asian	1.5%	0	4%	0
% Amer. Ind.	0	0	1%	0
% Other	0	0.7%	0	1%
BMI				
Mean (SD)	26 (3)	26 (3)	25 (3)	26 (3)
Smoker				
Never	76%	75%	65%	70%
Former	23%	25%	35%	30%
Yes	1%	0	0	0
Duration of Diabetes (yrs)				
Mean (SD)	19 (9)	18.5 (11)	17 (11)	19 (11)
Range	1-41	1-49	2-50	1.5-49
Diabetic conditions (N)				
Retinopathy	45	39	39	35
Neuropathy	21	32	24	28
Hypertension (N)	16	20	14	16

This reviewer checked the medication history and found the groups to be well balanced; the most commonly used drugs were drugs to treat rheumatic disease or gout (about 50-60% of the patients), antibacterials and analgesics.

Dosing

During the run-in phase of both trials, all patients were given SC insulin plus NPH insulin. Patients randomized into the control arm continued on their run-in dosing regimen. In Study 106, for the patients randomized to INH, the long acting NPH daily dose was replaced with bedtime Ultralente at approximately 65-75% of the dose of the run-in NPH dose.

Regular home monitoring of glucose was expected and dosing was individualized (goal of 2-hour post-prandial increment <60 mg/dL). For INH patients, one of the following 5 dose levels was to be used prior to meals with the initial dose based on weight as follows:

Body Weight (kg)	INH dose		Equivalent SC dose
30-44	1 inhale	1 mg	3 U
45-59	2 inhales	1 mg	6 U
60-79	1 inhale	3 mg	9 U
80-99	1 inhale	1 mg	12 U
	1 inhale	3 mg	
>100	2 inhales	3 mg	18U

Extra doses could be given prior to snacks or at bedtime.

Dose titration was performed based on the recommendations shown in the table below.

Time out of target range	Target range		INH insulin dose adjustment		SC insulin dose adjustment	
	Study 106	Study 107	Study 106	Study 107	Study 106	Study 107
Pre-breakfast	80-140	80-120	Bedtime Ultralente	Bedtime NPH	Evening NPH	Bedtime NPH
Pre-lunch	80-140	80-120	Pre-brkfast INH	Pre-brkfast INH	Pre-brkfast reg SC	Pre-brkfast reg SC
Pre-supper	80-140	80-120	Pre-lunch INH	Pre-brkfast +/-pre-lunch INH	Pre-breakfast NPH	Pre-brkfast NPH +/-pre-lunch reg SC
2-hr post-prandial	NA	<180	NA	preceding meal INH	NA	preceding meal reg SC
Bedtime	100-160	100-140	Pre-supper INH	Pre-supper INH	Pre-supper reg SC	Pre-supper reg SC

On the following page is a summary of the mean dose levels for short and long acting insulin. One concern of FDA was that increases in long-acting use may compensate for inadequate treatment on short-acting insulin and then may mask notable treatment differences.

In Study 106, the treatment groups differed in both the type of long-acting insulin and the type of short-acting insulin administered so changes in either cannot be independently compared across the treatment groups. Within the INH group, baseline long-acting was NPH BID and, on-study, bedtime Ultralente was given. Within the SC group, NPH BID was given throughout the trial. So the decreases seen in long acting in the INH group are due to the differing types of long-acting (Table 3.1.14). For both treatment groups, doses of short-acting increase while on randomized treatment from Week 1 to Week 24 with a larger increase seen within the INH group.

Table 3.1.1.4 Insulin Dosing; For INH group, short-acting is measured in mg, all other doses are in units (Medians and distributions are provided for long-acting dose and ratio of doses in Appendices 6.2 and 6.3)

	Study 106		Study 107	
	INH Mean (SD)	SC Mean (SD)	INH Mean (SD)	SC Mean (SD)
Insulin Use				
Screening	n=132	n=130	n=97	n=101
Short-acting	23.8 (15.0)	26.2 (17.3)	25.9 (19.1)	25.4 (20.7)
Long-acting	34.4 (18.5)	34.6 (20.6)	31.7 (16.8)	35.1 (20.1)
Ratio	2.2 (2.4)	1.8 (1.4)	2.1 (3.0)	2.1 (1.8)
Insulin Use				
Run-in Wk -3	n=132	n=134	n=101	n=103
Short-acting	18.8 (11.9)	17.6 (8.4)	22.2 (13.9)	22.6 (12.4)
Long-acting	32.4 (15.0)	32.2 (15.4)	29.8 (14)	29.4 (12.4)
Ratio	2.7 (3.8)	2.1 (1.2)	1.8 (1.4)	1.7 (1.1)
Insulin Use				
Run-in Wk 0 ¹	n=137	n=135	n=103	n=105
Short-acting	18.4 (9.2)	18.2 (9.0)	23.7 (12.4)	24 (11)
Long-acting	34.4 (15.0)	35.7 (18.0)	30.8 (14.8)	30.3 (12.6)
Ratio	2.4 (1.8)	2.3 (1.5)	1.6 (1.4)	1.5 (0.8)
Insulin Use by Week on Study				
Week 1	n=136	n=135	n=103	n=105
Short-acting	10.8 (4.0)	18.2 (9.4)	8.0 (4.0)	24.1 (10.6)
Long-acting	23.8 (11.8)	36 (17.8)	30.5 (14)	30.5 (12.7)
Ratio	2.4 (1.4)	2.4 (1.6)	4.7 (3.4)	1.4 (0.7)
Week 6	n=134	n=130	n=103	n=102
Short-acting	11.8 (5.3)	18.2 (10.5)	9.4 (4.5)	24.1 (10.7)
Long-acting	26 (12)	37.6 (18.8)	31.0 (14.8)	31.1 (13.6)
Ratio	2.5 (1.6)	2.6 (1.7)	3.9 (2.5)	1.5 (0.7)
Week 12	n=127	n=125	n=102	n=99
Short-acting	12.4 (5.7)	17.9 (8.9)	10.2 (4.7)	24.2 (11.1)
Long-acting	25.9 (12.5)	37.5 (17.9)	31.1 (15.2)	31.1 (14)
Ratio	2.4 (1.4)	2.7 (2.8)	3.6 (2.4)	1.4 (0.7)
Week 24	n=123	n=124	n=98	n=97
Short-acting	13.4 (5.9)	18.3 (9.6)	10.8 (5)	26.0 (12.7)
Long-acting	26.1 (13)	36.8 (18.3)	31.5 (15.9)	32.5 (14.7)
Ratio	2.3 (1.4)	2.8 (3)	3.4 (2.2)	1.4 (0.7)
Change from Week 0 LOCF Long-acting	-8.8 (11)	+1.4 (10)	+0.7 (7.9)	+2.0 (7.6)
Trt Difference	-10.5 (-13, -8)		-1.3 (-3.4, 0.8)	

In Study 107, the treatment groups are comparable regarding long-acting insulin (NPH BID) and short-acting SC at baseline. All patients continued on their long-acting insulin regimen through the study; an overall small increase in dose is seen in both groups with essentially no difference

¹ Week 0 is the last week of the run-in period; the last day of Week 0 is the baseline visit.

between the groups (Table 3.1.1.4). The treatment differences in short-acting doses and the differing ratios are in keeping with what would be expected for INH compared to SC. There is no evidence that long-acting doses are increased over time in either treatment group. The dose of INH increases on average from Week 1 to about Week 12 and then is stable; an inspection of individual patients by this reviewer showed that the dose of INH is generally stable by about Week 10.

Statistical Methods

The protocol and study report stated that patients were assigned to treatment using an algorithm to balance on inclusion/exclusion criteria. Discussions with the applicant revealed that an algorithm was not used and patients were randomized to treatment via a central telephone system. Randomization was blocked within centers in sizes of 2 and 4. The protocol did not mention the block sizes or that blocking was within center so investigators should have been unaware of the blocking scheme. However, even without that knowledge, one might be able to guess the block sizes quite easily particularly since the block sizes are small and the number of patients in each site is small. Given the trials are open-label and blocking was within centers, there is a potential for selection bias. To test for selection bias, this reviewer performed an analysis described by Berger and Exner (Controlled Clinical Trials 20:319-327, 1999). This analysis was only done for Study 107.

The protocol specified that the evaluable population (compliant with a specified minimum amount of treatment) would be the primary analysis set. In addition, an analysis of the intent-to-treat (ITT) population was planned. Missing data were handled by using the last observation carried forward (LOCF) approach; other imputation schemes were also considered. For this reviewer's analysis, the primary analysis population is the ITT population. In addition, analyses of completers was done. Due to the small number of dropouts, there is no expectation that the choice of analysis population will impact on the results.

The primary outcome variable, change from baseline HbA1c, was analyzed using an analysis of covariance model which included baseline HbA1c as a covariate and terms for center (with smaller centers pooled) and treatment. As secondary analyses, the applicant considered other factors as covariates: these factors included age, BMI, baseline C-peptide, gender, race and baseline insulin antibody. This reviewer also considered dosing as a covariate. First order interactions were tested. All randomized patients were analyzed with stratification on age (<18 versus 18 and older). Only the 18 and older patients are presented here. Results for under 18 are included in the subgroup section of this review.

Both trials were powered to rule out a treatment difference of 0.5% in favor of SC over INH based on a 95% confidence on the difference in HbA1c change from baseline at Week 24.

Efficacy Results

Given that the trials are open-label and that randomization was blocked on center, this reviewer was concerned that investigators could have selected patients based on knowledge of the next assignment. For Study 107, the results for a test for selection bias showed that baseline characteristics and outcome were not related to the probability of being assigned to the INH group; therefore this reviewer believes that there is no evidence that there was bias in the selection of patients for Study 107. The analysis was not done for Study 106 since Study 107 was considered the most important of the two studies.

Baseline was computed as the average of Weeks -1 and 0. For each study, the treatment groups were comparable at baseline for HbA1c and FPG (Table 3.1.1.5). Plots of the baseline values can be found in Appendix 6.4.

The confidence intervals for the HbA1c treatment differences (Table 3.1.1.5) indicate no statistically significant difference between the treatment groups. The results in Study 106 are more favorable to SC while the opposite is true for 107 with more favorable results for INH; although the differences are not clinically relevant with confidence limits within the noninferiority margin set by the protocol (0.5) and within a margin of 0.4 which is generally accepted by the FDA .

Table 3.1.1.5 Results for HbA1c and FPG

	Study 106			Study 107		
	INH Mean (SD)	SC Mean (SD)	LSM Diff 95% CI	INH Mean (SD)	SC Mean (SD)	LSM Diff 95% CI
HbA1c Baseline	n=136 7.9 (0.9)	n=132 8.0 (1.0)		n=103 7.8 (0.9)	n=103 7.8 (1.0)	
Change from baseline Wk 24 Completer	-0.25 (0.8) (n=120)	-0.4 (0.7) (n=123)	+0.1 (-0.1, +0.3)	-0.3 (0.8) (n=97)	-0.2 (0.8) (n=96)	-0.1 (-0.3, +0.1)
Wk 24 LOCF	-0.2 (0.8)	-0.4 (0.8)	+0.1 (-0.04, +0.3)	-0.3 (0.8)	-0.2 (0.8)	-0.1 (-0.3, +0.1)
FPG Baseline	n=131 191 (61)	n=129 198 (61)		n=101 178 (69)	n=97 191 (68)	(p=0.18)
Change from baseline By week						
12	-31 (99)	-13 (96)	-22 (-42, -0.8)	-30 (89)	-11 (85)	-26 (-50, -3)
24	-31 (91)	-1 (95)	-33 (-54, -12)	-25 (97)	+5 (83)	-38 (-63, -14)
Wk 24 LOCF	-29 (95)	-7 (100)	-27 (-48, -7)	-21 (98)	+5 (89)	-35 (-59, -12)

For the least squares mean difference (LSM), negative values favor INH. Statistical model included treatment, baseline, age as a stratifier (<18 and ≥18) and an interaction term for age and treatment.

The FPG results show significantly larger decreases in FPG for INH than for SC. The decreases in FPG are not correlated with decreases in HbA1c with r^2 values of less than 0.01 for INH and less than 0.1 for SC for each treatment group at Week 24.

Percentages of patients with final HbA1c level below 7% and 8% were named as secondary variables; percentage of patients below 8% at end of study is included in the applicant's proposed labeling. This reviewer has provided these results for all patients and by subgroups defined by baseline values of 8% and 7%. Results by subgroup indicate whether patients were able to maintain baseline levels or improve from baseline. There is no statistically significant difference between the groups in either study for both measures of overall percentages. For the subgroups, the most notable differences are a treatment difference of 38% (p=0.02, Fisher's exact test), in favor of SC in Study 106 for patients with baseline values less than 7% and a treatment difference of 20% (p=0.07, Fisher's exact test) in favor of INH in Study 107 for patients with baseline values less than 7%. The inconsistency of the results makes it difficult to draw any conclusions from these analyses. Overall it seems that patients with low values (under 7 or 8) at baseline tend to stay low even when changing treatment regimens. Less than half the patients with values above 8 at baseline are able to achieve HbA1c levels below 8% after 24 weeks of titrated treatment.

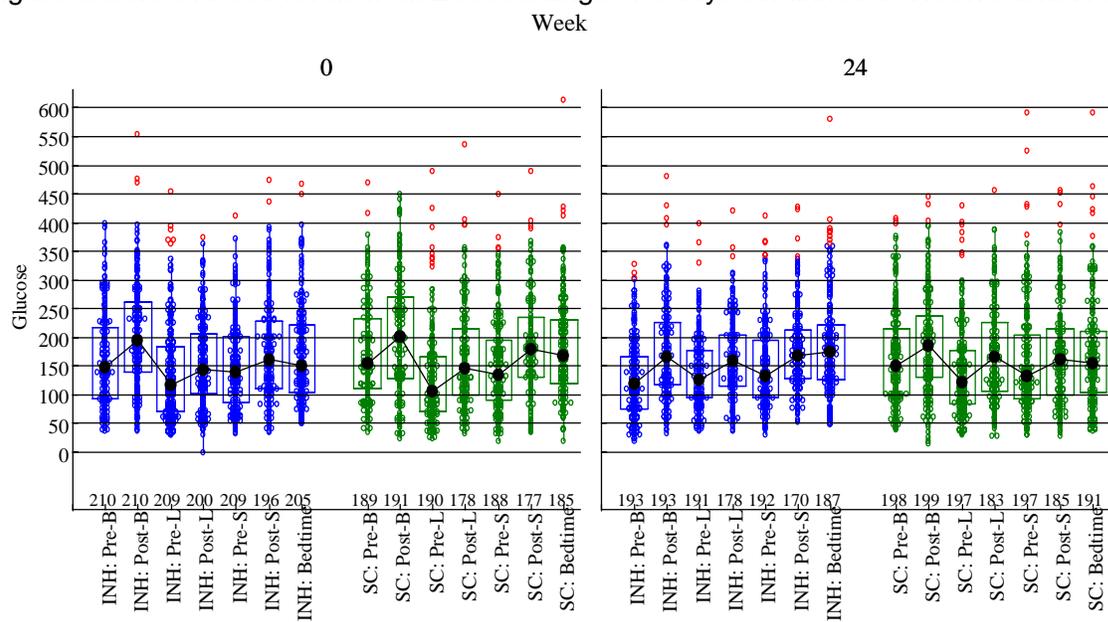
Table 3.1.1.6 Percentage of patients with HbA1c below 8 and below 7 at Week 24 LOCF by study and treatment and by baseline category

	Study 106		Study 107	
	INH (n=136)	SC (n=132)	INH (n=103)	SC (n=103)
% pts w/ last HbA1c <8%	64%	68%	75%	66%
By Subgroup				
Baseline<8%	84% (62/74)	91% (64/70)	92% (55/60)	85% (57/67)
Baseline≥8%	40% (25/62)	42% (26/62)	51% (22/43)	31% (11/36)
% pts w/ last HbA1c <7%	17%	20%	28%	30%
By Subgroup				
Baseline<7%	50% (12/24)	88% (14/16)	63% (12/19)	68% (13/19)
Baseline≥7%	10% (11/112)	10% (12/116)	20% (17/84)	21% (18/84)

The significant treatment effects seen for fasting plasma glucose levels but not for HbA1c changes led this reviewer to look, with the medical reviewer, at changes in glucose levels as recorded in 24-hour diaries over a 3-day period. The results are illustrated in the boxplots on the following page (Figure 3.1.1.1).

At baseline, the groups are comparable with medians varying within a day from about 125 to 200. At Week 24, in the INH group, the pre-breakfast value (median of about 130) is notably lower than the bedtime value (median of about 180) while in the SC group, the pre-breakfast values are similar to bedtime values and baseline values. This difference between the treatment groups in pre-breakfast values is about the same difference seen for FPG as reported in Table 3.1.1.5 on the previous page.

Figure 3.1.1.1 Glucose levels from 24-recordings of 3 days combined at Week 0 and Week 24



3.1.2 Clinical Trials in Patients with Type 2 Diabetes

For this review, the designs and the descriptions of the patient populations for all five of the Phase 3 trials of Type 2 diabetes are presented in the following section (3.1.2.1) in order to allow the reader to more clearly see the similarities and differences among the five trials. The results are presented in Section 3.1.2.2 with separate sections for each of the studies.

3.1.2.1 Overview of Phase 3 trials in Patients with Type 2 Diabetes

The applicant has conducted five Phase 3 clinical trials in Type 2 diabetics: Studies 108, 109, 110, 1001 and 1002 (see Summary Table 1.2.2.1 on page 7). Study 108 was an active-controlled trial in patients previously treated with SC insulin with the goal being to show that the effect of Exubera was non-inferior to the effect of SC insulin; the design of this trial is similar to the studies in Type 1 diabetics. Study 110, in seemingly naïve patients who were without medication for at least one month prior to screening and during the run-in, compared Exubera to rosiglitazone (4 mg BID, the most effective marketed dose of rosiglitazone). Study 109 was a three arm study of Exubera plus two oral agents (combination therapy), Exubera monotherapy and the combination of two oral agents. Studies 1001 and 1002 were add-on studies where patients inadequately treated with sulfonylureas or metformin, respectively, were randomized to either add-on Exubera or oral agent (metformin or glibenclamide). Studies 109 and 110 were superiority trials and the other three studies were non-inferiority trials.

Three trials (108, 109 and 110) were conducted in North America and two trials (1001 and 1002) were conducted in Europe, South America, the Middle East and Africa (Table 3.1.2.1.1). Three were of 24-week duration and two of 12-week duration; generally FDA recommends 24 to 26-week duration for a trial to assess Type 2 antidiabetic drugs.

Table 3.1.2.1.1 Studies in patients with Type 2 diabetes

Study	Dates	Locations (# centers)	Duration of randomized treatment	Total # of patients randomized
108	9/99 to 12/00	US (38) Canada (11)	24 weeks	299
109	6/99 to 9/00	US (40) Canada (12)	12 weeks	309
110	10/99 to 3/01	US (40)	12 weeks	145
1001	2/00 to 12/01	Europe (43) Scandinavia (13) UK (6) Israel (4) Brazil (4) South Africa (3)	24 weeks	427
1002	3/00 to 5/02	Europe (44) Scandinavia (13) UK (7) Israel (5) Brazil (5) South Africa (3)	24 weeks	476

Studies 108, 109 and 110 were conducted at many of the same sites; Studies 1001 and 1002 were also conducted at many of the same sites. Patients could not enter more than one trial according to all the study protocols so with regard to enrolled patients, the studies would be considered independent. The trials are, however, not independent regarding investigators. Since the trials are overlapping and the patient entry criteria varied based on diabetic treatment experience, it seems unlikely that the collection of data by an investigator would be biased by their prior participation in the trial. Also since a large number of sites participated in each trial, the

number of patients at each site is small. No one site would be expected to have a strong influence on the outcome of the trial; nevertheless, this reviewer did examine data by center, though data by center is not presented in the review.

Patients in all trials were asked to follow the ADA diet, to do moderate exercise 3 times per week (no specific questions regarding diet and exercise were asked) and to monitor home glucose 4 times per day.

Patients enrolled in these five trials had to meet the following entry criteria:

- No significant respiratory disease
- No smoking within 6 months of screening nor on study
- Fasting C-peptide of 0.2 pmol/ml or more
- No more than 1 episode of severe hypoglycemia and no ER or hospital visits for hypoglycemia in the 6 months previous to screening (not a criterion in 1001 and 1002)

Additional entry criteria as well as study design characteristics are summarized below.

Table 3.1.2.1.2 Entry criteria and treatment arms in Phase 3 studies of Type 2 diabetics

Study	HbA1c	Diag Diab	BMI \leq	Trt Prior to Screening	4 Week Run-in Trt	Randomized Trt	
						INH	Control
108	6-11% at Wks-4+-1	≥ 1 year	40	Stable SC for 2 mos+	SC BID reg+NPH	Pre-meal + Ultralente	SC BID reg+NPH
109	8-11% at Wks-4+-1	≥ 1 year	35	Stable doses of 2 OA's for 2 mos+	2 OA's; SU or repaglinide + glitazone or MET	Mono INH Comb INH+2 OA's	2 OA's SU or repaglinide + glitazone or MET
110	8-11% at Wks-4+-1 ¹	≥ 2 mos	35	No antidiab. drug for 1 month+	diet+exer	Pre-meal	ROSI 4 mg BID
1001	8-12% at Wk -1	≥ 6 mos	Not in protocol	Poorly controlled on SU for 2 mos+	SU	Pre-meal Add-on to SU	Metformin Add-on to SU
1002	8-12% at Wk -1	≥ 6 mos	Not in protocol	MET for 2 mos+	MET	Pre-meal Add-on to MET	Glibenclamide Add-on to MET

SU=sulfonylurea MET=metformin

In all studies, the initial dose of INH was based on body weight; titration of the INH dose was based on self-measured premeal glucose levels and a target range of 80-140 mg/dL. A table of the oral antidiabetic medications and doses is provided in Appendix 6.5.

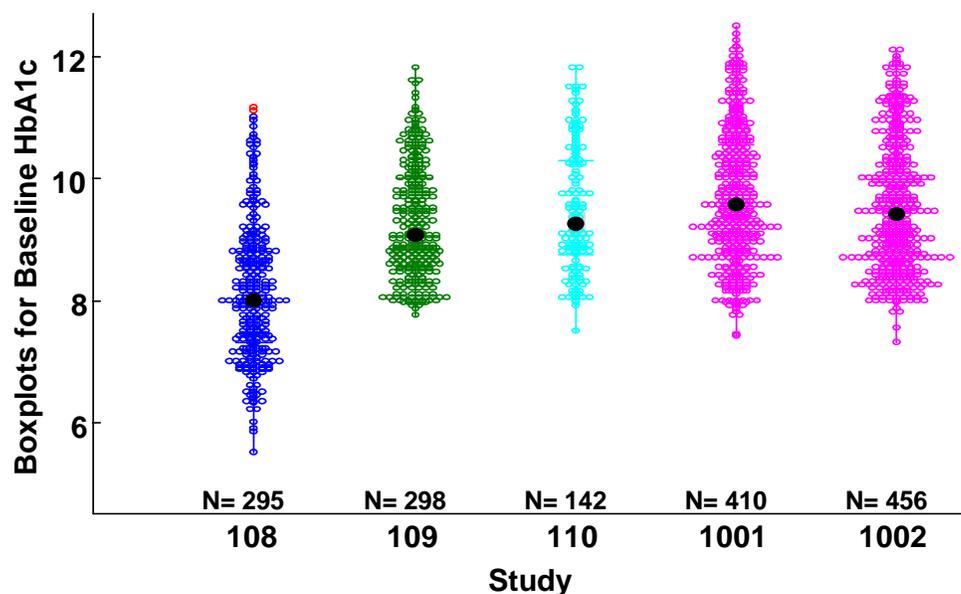
The entry criteria for HbA1c are similar for the studies with the exception of Study 108 where patients could be enrolled with values below 8. For Study 108, patients were not necessarily considered inadequately treated at baseline as they were in the other 4 studies. The distributions of the observed HbA1c baseline data (Figure 3.1.2.1.1 on following page) illustrate the differences among the studies (note that the treatment groups had comparable distributions within each study so data with the groups combined is shown).

Figure 3.1.2.1.1 HbA1c baseline values for five studies of Type 2 diabetics; each symbol

¹ The study report stated that the acceptable range for HbA1c was 6 to 11 while the protocol stated 8 to 11; from the data, it is clear that the study report is incorrect.

represents a patient

VISIT
0



Within studies, treatment groups are fairly well balanced with regard to baseline demographics with the only notable exception being gender within Study 110. Across studies, variability in all baseline parameters is seen suggesting some heterogeneity among the patient populations; this being a positive observation which could improve generalizability.

Table 3.1.2.1.3 Baseline demographics in Type 2 studies

	108		109			110		1001		1002	
	INH	SC	INH	INH+ OA's	OA's	INH	ROSI	INH	MET	INH	SU
Age											
Mean	59	56	57	58	56	53	54	61	60	55	55
Range	36-80	23-78	35-77	38-77	33-80	28-76	29-80	37-79	35-79	35-74	36-77
% ≥65	29%	25%	23%	22%	21%	17%	19%	34%	36%	16%	16%
Gender											
% female	32%	34%	29%	36%	36%	36%	55%	45%	49%	44%	42%
Race											
White	77%	74%	79%	78%	82%	77%	70%	96%	95%	93%	95%
Black	12%	10%	8%	9%	5%	9%	15%	2%	2%	3%	2%
Asian	3%	3%	2%	0%	2%	1%	0%	1%	1%	2%	2%
Other	1%	5%	3%	5%	3%	0%	1%	2%	2%	3%	1%
Hispanic	8%	8%	8%	8%	7%	12%	13%	0%	0%	0%	0%
BMI											
Mean	31	30	30	30	30	32	33	29	29	32	31
Range	21-51	21-38	22-39	18-38	18-38	20-44	22-48	20-48	20-57	19-47	22-47
Dur. of Diabetes (yrs)											
Mean	13.8	13.2	9.3	9.8	9.6	4.3	3.1	9.6	8.8	8.4	7.8
Range	.4-59	.9-43	2-25	1-37	1-33	.1-22	.1-18	.7-37	.5-33	.6-36	.3-30

The comparator in Study 110, rosiglitazone, has found to be more effective in women and overweight men in studies previously reviewed by FDA; so the larger number of women in the

rosiglitazone group of Study 110 may favor rosiglitazone over INH. However, gender subgroup analyses of the data from this study showed no gender effect and suggests no bias due to this small imbalance.

This reviewer checked the coding for “other” race and found that a total of 3 patients enrolled in the five studies were Native Americans; clearly this ethnic group is under-represented in the Type 2 database for Exubera.

3.1.2.2 Efficacy Results for Phase 3 trials in Patients with Type 2 Diabetes

In this section, patient disposition and discontinuation reasons are summarized for all five trials in Tables 3.1.2.1.3 and 3.1.2.1.4, respectively. Following these tables, the efficacy results with a brief description of the statistical methods for the five Phase 3 trials are presented separately. This section closes with some graphics summarizing the efficacy results.

Studies 108, 109, 1001 and 1002 all recruited more patients then specified in the protocol with the largest increases in sample sizes seen for Studies 1001 and 1002 where the trials were powered for 180 patients in each group. An inspection by this reviewer of the enrollment dates shows no irregularities; for example, there are no breaks in the dates suggesting that enrollment was stopped and started for a period of time. This reviewer does not think that over-enrollment biases the results.

For Study 110, investigators had difficulties enrolling adequate numbers of patients so the protocol was modified from requiring 150 patients per group to requiring 124 patients per group. This new goal of a total of 248 patients also could not be met and enrollment into the trial was halted at a total of 145 patients. Based on the protocol assumption of a 20% difference in response rates, the trial, as executed, was underpowered; however, a larger difference was observed and the treatment difference on the primary outcome (percent of patients with HbA1c below 8%) was statistically significant.

In all studies, the dropout rates are at most 12% and are generally less than 10% (Table 3.1.2.2.3). These low dropout rates do not impact the interpretation of the outcome data therefore this reviewer did not perform any sensitivity analyses.

Table 3.1.2.2.3 Patient disposition in Type 2 studies

	108		109			110		1001		1002	
	INH	SC	INH	INH+ OA's	OA's	INH	ROSI	INH	MET	INH	SU
Rand.	149	150	105	102	102	76	69	225	202	243	233
Never trt.	0	0	1	3	0	1	1	3	1	4	2
Week 6	99%	99%	99%	91%	97%	99%	94%	96%	98%	95%	95%
Week 12	98%	98%	NA	NA	NA	NA	NA	94%	95%	93%	91%
Completer	89%	93%	92%	91%	97%	93%	91%	92%	88%	90%	88%

For studies 1001 and 1002, week 12 in the table above actually refers to Week 14.

The primary reasons for dropout across the studies were adverse event (ADE) and subject request (bolded numbers in Table 3.1.2.2.4). Subject request referred to both lost-to-follow-up and voluntary withdrawal. The most common ADE among Type 2 patients across all controlled studies was increased coughing; in Studies 1001 and 1002, 2 patients out of the 12 discontinuing for an ADE had increased coughing as a reason (one moderate and one severe).

Table 3.1.2.2.4 Reasons for discontinuation in Type 2 studies; Number of patients

	108		109			110		1001		1002	
	INH	SC	INH	INH+ OA's	OA's	INH	ROSI	INH	MET	INH	SU
Total SS	149	150	105	102	102	76	69	225	202	243	233
Death	2	0	0	0	0	0	0	0	0	0	3
ADE	2	2	1	1	0	1	3	4	7	8	3
LOE	1	1	3	0	2	0	1	0	2	0	4
Lab abn.	0	0	0	0	0	0	0	1	0	0	0
Subj req.	10	6	2	2	2	2	0	7	5	4	6
Oth/pr. viol	2	0	1	1	2	1	1	3	9	8	10
Total	17	9	7	4	6	4	5	15	23	20	26

The “other” reason referred to protocol violations (including not satisfying entrance criteria) or to other reasons. In Studies 1001 and 1002, about half the patients in the “other” category were listed as protocol violators.

Statistical Methods

The sponsor performed an analysis of covariance (ANCOVA) of HbA1c change from baseline with baseline HbA1c as a covariate and with terms for treatment and center (small centers pooled). This reviewer did not include center in the model and obtained essentially the same results; all results presented in the efficacy tables and graphs were computed by this reviewer. ITT data with the last observation substituted for missing data (LOCF) was analyzed. For all endpoints, a 95% confidence interval was computed and a difference of 0.4% for HbA1c was considered to be clinically relevant.

Study 108 Efficacy Results

The goal of Study 108 was to show that meal-time INH was comparable to conventional SC insulin given twice daily in patients stable on SC therapy at baseline. This goal was met for HbA1c change from baseline with the confidence interval showing no clinically relevant difference (Table 3.1.2.2.5). The treatment difference on FPG is statistically significant.

Table 3.1.2.2.5 Study 108 Efficacy results at Week 24; means and SD's and LSM difference

	INH (TID) (n=146)	SC (BID) (n=149)	INH-SC diff (95% CI)	p-value
HbA1c				
Baseline	8.1 (1.1)	8.2 (1.1)		
Change	-0.7 (1.2)	-0.6 (1.1)	-0.1 (-0.3,+0.2)	0.45
%<8% at EP	76%	69%	OR 1.4 (0.9, 2.4)	0.19
BSL<8%	92% (65/71)	88% (61/69)		
BSL≥8%	61% (46/75)	53% (42/80)		
FPG				
Baseline	152 (37)	159 (45)		
Change	-20 (55)	-9 (52)	-16 (-26, -5)	0.004

Negative differences favor INH, model with trt and baseline as terms. OR=odds ratio

Study 109 Efficacy Results

For Study 109, patients on oral therapy were randomized either to INH alone (i.e. switching from OA's), to continue on their OA regimen or to add-on INH. The protocol described a step-down procedure to adjust for making two comparisons (the combination versus OA and OA monotherapy versus INH). Usually in combination studies, the combination is expected to beat each component for approval of the combination product. For this study, the combination does beat each component at a <0.0001 nominal level (Table 3.1.2.2.6) for both HbA1c and FPG. In addition to showing that the combination product was effective, the applicant wished to show that switching from oral therapy (in this case two oral therapies) to INH resulted in more effective lowering of HbA1c which they were able to demonstrate (CI's not shown here).

Table 3.1.2.2.6 Study 109 Efficacy results at Week 12; means and SD's and LSM difference

	INH (n=102)	OA (n=96)	INH+OA (n=100)
HbA1c			
Baseline	9.3 (0.9)	9.3 (1.0)	9.2 (1.0)
Change	-1.5 (1.0)	-0.3 (0.9)	-1.9 (0.9)
Trt diff vs INH+OA (95% CI)	-0.5 (-0.7, -0.3)	-1.7 (-1.9, -1.5)	
%<8% at EP	56%	19%	86%
FPG			
Baseline	203 (43)	203 (44)	195 (49)
Change	-26 (58)	-1.7 (41)	-49 (50)
Trt diff vs INH+OA (95% CI)	-28 (-41, -16)	-53 (-65, -40)	

Negative differences favor the combination INH+OA

Study 110 Efficacy Results

Study 110 differed from the other studies in that the percentage of patients with an HbA1c of less than 8% at Week 12 was the primary efficacy variable. A logistic regression analysis was planned to analyze the primary efficacy variable.

The efficacy results showed significant effects on HbA1c for INH compared to rosiglitazone (4 mg BID, highest effective dose) for both the primary measure and change from baseline; the decrease in HbA1c for rosiglitazone is larger than what has been seen in earlier trials in naïve patients where mean decreases are usually less than 1%.

Table 3.1.2.2.7 Study 110 Efficacy results; means and SD's and LSM difference

	INH (n=75)	ROSI (n=67)	INH-ROSI diff (95% CI)	p-value
HbA1c				
Baseline	9.5 (1.1)	9.4 (0.9)		
Change	-2.3 (1.1)	-1.4 (1.2)	-0.9 (-1.2,-0.5)	<0.0001
%<8% at EP	83%	58%	OR 3.4 (1.6, 7.4)	0.0014
FPG				
Baseline	208 (56)	199 (50)		
Change	-64 (57)	-56 (42)	-2.6 (-16, +11)	0.71

Negative differences favor INH. OR=odds ratio

No significant treatment differences are seen for FPG. Due to the differing results for HbA1c and FPG, this reviewer looked at the correlation between the measures both at baseline and change at endpoint. For each treatment group, significant correlation is seen both at baseline and endpoint, as would be expected for these measures of glucose control. It is not clear to this reviewer why no statistical difference is seen for FPG.

Studies 1001 and 1002 Efficacy Results

Patients inadequately treated on a sulfonylurea in Study 1001 or on metformin in Study 1002 were randomized to add-on therapy of INH or an OA (metformin in Study 1001 and glibenclamide in Study 1002). For each trial, two objectives were named; to show superiority in patients with baseline HbA1c above 9.5% and to show non-inferiority using all patients. This reviewer is only presenting the results for all patients in the table below. In both studies, no clinically relevant difference was seen between the treatment groups on both HbA1c and FPG .

Table 3.1.2.2.8 Efficacy results for Studies 1001 and 1002; means and SD's and LSM difference

Study 1001				
	INH (n=214)	MET (n=196)	INH-MET diff (95% CI)	p-value
HbA1c				
Baseline	9.7 (1.1)	9.7 (1.2)		
Change	-2.1 (1.1)	-1.9 (1.2)	-0.2 (-0.4, -0.04)	0.02
%<8% at EP	64%	58%	OR 1.3 (0.9, 1.9)	0.22
FPG				
Baseline	220 (55)	219 (55)		
Change	-48 (55)	-50 (55)	+2.4 (-6, +11)	0.56
Study 1002				
	INH (n=234)	GLIB (n=222)	INH-GLIB diff (95% CI)	p-value
HbA1c				
Baseline	9.5 (1.1)	9.6 (1.1)		
Change	-2.1 (1.2)	-2.1 (1.1)	-0.2 (-0.4, +0.01)	0.13
%<8% at EP	77%	73%	OR 0.8 (0.5, 1.2)	0.28
FPG				
Baseline	203 (56)	216 (54)		
Change	-42 (54)	-53 (54)	+2 (-6, +10)	0.59

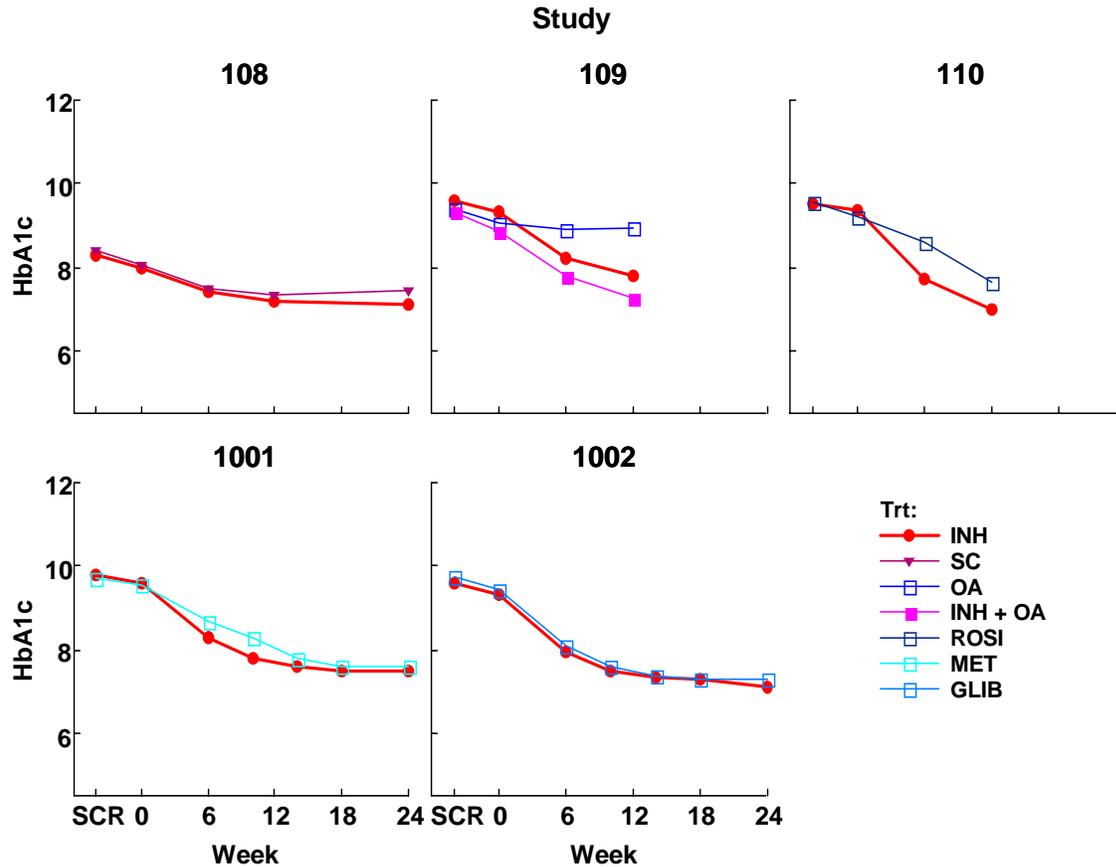
Negative differences favor INH, model with trt and baseline as terms. OR=odds ratio

In patients with baseline HbA1c values above 9.5% (about 100 patients in each group), essentially the same results were seen in each of the studies. The treatment difference was statistically significant with mean differences of -0.38 (CI $-0.63, -0.14$) in Study 1001 and -0.37 (CI $-0.62, -0.12$) in Study 1002. Since a difference of about 0.4 is often named as the smallest clinically relevant difference in diabetes trials, it is clear that these results can only be considered marginally favorable to INH. [The treatment difference in the lower stratum is very small being between -0.1 and $+0.1$.]

Summary Graphs of Efficacy in all 5 studies in Type 2 diabetic patients

Figure 3.1.2.2.1 below shows mean HbA1c overtime in each of the 5 trials in Type 2 diabetics. A slight decrease in mean HbA1c is seen during the run-in period even though patients are maintained on their screening regimen during this period. The solid red line represents treatment with INH. Graphs of FPG are provided in Appendix 6.6.

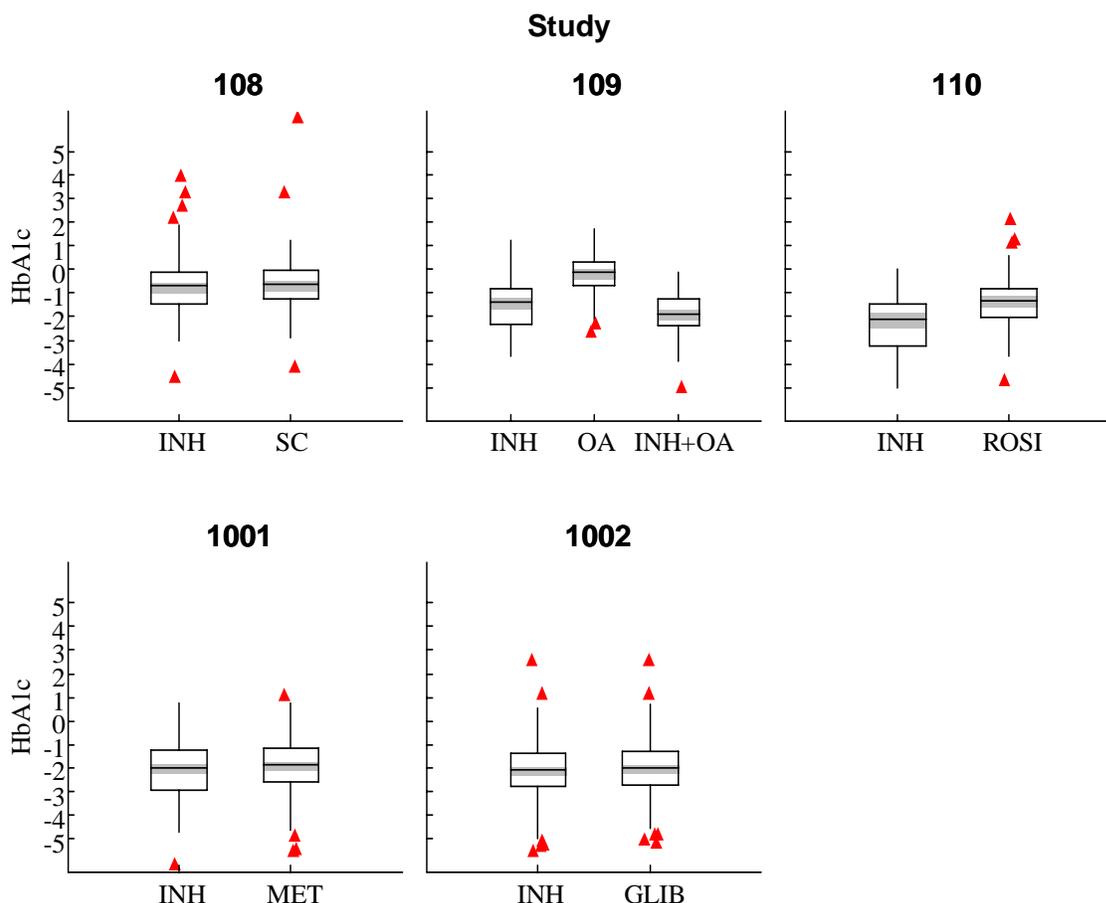
Figure 3.1.2.2.1 Mean HbA1c overtime in Phase 3 studies of Type 2 diabetics



This reviewer had some concern about the length of the trials given that FDA generally recommends at least 24-week trials. From the graphs it is clear that 24 weeks appears to be adequate to obtain a stable response. After 12 weeks of therapy in the 24-week trials, it appears that the maximum mean response is achieved for INH so perhaps in Studies 109 and 110, the INH response at Week 12 is representative of the response we might see with further treatment on INH. This reviewer thinks we can not make this assumption for the rosiglitazone response in Study 110. Previous studies reviewed by this statistician showed more lowering up to about Week 18 in naïve patients on rosiglitazone 4 mg BID. In Study 110, the comparison at Week 12 is unfairly biased against rosiglitazone.

Boxplots of HbA1c change from baseline are shown below in order to illustrate the range and distribution of effects in each of the Phase 3 studies of Type 2 diabetics. The comparability of the distributions is clearly evident in Studies 108, 1001 and 1002 where non-inferiority was demonstrated.

Figure 3.1.2.2.2 Boxplots of HbA1c change from baseline at endpoint for all five Phase 3 studies in Type 2 diabetics



3.2 Evaluation of Safety

Three issues regarding safety were examined by this reviewer: 1) rates of hypoglycemia in Studies 107 and 108, 2) the relationship between insulin antibodies and dose and 3) the relationship between insulin antibodies and the incidence of severe hypoglycemia. Note that full safety reviews were performed by three FDA reviewers; Dr. Buenconsejo (statistical reviewer) and Drs. Mahoney and Seymour (medical reviewers).

Hypoglycemia in Studies 107 and 108

The applicant has reported risk ratios for hypoglycemia computed using “a counting process approach for recurrent time-to-event data”. Based on these estimates, the applicant makes comparative statements about the risk of hypoglycemia. For several studies, the risk ratios appear to be inconsistent with the incidence of events. To examine the use of recurrent time-to-

event modeling in the analysis of hypoglycemia data, this reviewer examines the hypoglycemic event data for Studies 107 and 108; two studies of particular interest to the FDA clinical reviewers (see Appendix 6.7 for a summary of events for all Phase 3 studies). Note that this reviewer considers this examination of the analysis of the hypoglycemic events as preliminary because there was insufficient time to adequately test the appropriateness and the fit of the models given that this reviewer discovered this issue only about a week before this review was due to the FDA advisory committee staff. The final statistical review will more thoroughly examine the models and the appropriateness of these models for hypoglycemic data, in general.

The applicant has provided this reviewer with the SAS coding used to model recurrent hypoglycemic events. (See Appendix 6.8 for the applicant's coding and for modifications by this reviewer). Their model is an Anderson-Gill type proportional hazards model for recurrent events where events are assumed to be independent, i.e. there are no modifications made to the model to account for recurrent events within patient or dependency among events. This reviewer has run one additional model which calls for a robust sandwich estimate for the covariance matrix(henceforth referred to as the robust variance) to correct for possible correlations. Other models, such as one including, for example, a time-dependent variable (e.g. prior number of events), will be considered in a later version of this review.

Hypoglycemia with a severity scale was defined in the protocol. Late in the IND stage of development, an additional definition of hypoglycemia was provided by an FDA medical reviewer (see Appendix 6.9 for both definitions). The latter definition was applied retrospectively and so data was not collected specifically for this endpoint. For the study reports, the applicant has presented results for the protocol-defined overall and severe hypoglycemia while in the NDA summaries, the applicant has focussed only on the FDA-defined hypoglycemic events. Note that, by definition, FDA events are a subset of the overall protocol-defined hypoglycemic events and that all protocol-defined severe events, also are counted as FDA events.

For Study 107, the severe hypoglycemic events are examined while for Study 108, all hypoglycemic events are examined. For both studies, the FDA-defined events are also discussed. Study 107 is a study of Type 1 diabetics so essentially all patients have frequent hypoglycemic events throughout the trial in both groups (average of about 9-10 per month in both groups); to assess safety, the emphasis is primarily on the severe events. In Study 108, a trial of Type 2 diabetics, there are very few severe hypoglycemic events (a total of 5 in this trial) and so the emphasis in the Type 2 population is on the overall hypoglycemic event rate. The results for Study 107 are described first below, followed by the results of Study 108.

Results for severe hypoglycemic events in the applicant's study report for 107 were shown for adults and children combined. The reviewer's table on the following page shows the results for children (<18) and adults separately along with the combined results.

Table 3.2.1 Study 107 Tabulation of severe hypoglycemic events

	% (n/N) of pts. with at least 1 event	Number of patients with "n" events								
		0	1	2	3	4	5	6	12	
All pts										
INH	16% (26/162)	136	15	4	3	1	1	1	1	1
SC	14% (22/162)	140	16	5	1	0	0	0	0	0
Adults										
INH	17% (18/103)	85	10	3	2	1	1	0	1	1
SC	13% (13/103)	90	8	4	1	0	0	0	0	0
Children										
INH	7% (8/59)	51	5	1	1	0	0	1	0	0
SC	8% (9/59)	50	8	1	0	0	0	0	0	0

Chi square tests comparing the percentage of patients having at least one severe event show no statistically significant differences between the groups shown in the second column in Table 3.2.1 above with p-values all greater than 0.20. Chi square tests on the counts shown for recurrent events in the table above also yielded non-significant results with p-values greater than 0.60.

These results strongly contrast with the applicant's reporting of a statistically significant doubling of risk of severe hypoglycemia (risk ratio of 2.02 with 95% CI of 1.3 to 3.15; also reported in Diabetes Care 28:1630-1635, 2005) for all INH-treated patients compared to SC treated patients. Comparable risk of overall hypoglycemia was observed (applicant's risk ratio of 0.94, CI of 0.91 to 0.97). This reviewer characterizes the groups as comparable, even though the CI does not include 1, because of the size of the estimate and because the repeated measures procedure is powerful enough to detect small clinically insignificant differences.

The number of "FDA-defined" events is summarized below for adults only (Table 3.2.3). A chi square test on at least 1 FDA-defined event yields a p-value of 0.49 (odds ratio of 0.73 with 95% CI of 0.29 to 1.8). A Wilcoxon rank-sum test on the number of events observed for each patient yields a p-value of 0.09 (median of 6 events for INH and 8 events for SC). Note that there are four SC patients with a large number of events.

Table 3.2.2 Study 107 Adults Tabulation of FDA-defined hypoglycemic events

	% (n/N) of pts. with at least 1 event	Number of patients with "n" events																
		0	1	2	3	4	5	6-10	11-15	16-20	21-30	31-40	41-50	51	53	56	78	
Adults																		
INH	88% (91/103)	12	10	11	8	2	4	22	13	8	8	3	2	0	0	0	0	0
SC	91% (94/103)	9	5	11	11	2	3	16	15	8	11	7	1	1	1	1	1	1

Using a survival analysis of recurrent FDA-defined events, the applicant reports a risk ratio of 0.72 with 95% CI of 0.67 to 0.79; a statistically significant reduced risk of hypoglycemia for INH compared to SC. So a conclusion opposite to the one reached for severe hypoglycemia is reached using the FDA definition which includes mild, moderate and severe events.

The applicant's survival analyses yield statistically significant results whereas the reviewer's chi square and Wilcoxon tests show no treatment difference. The drawback to the reviewer's tests is that they do not take time into account. This reviewer wonders though whether time is an important factor in these studies. To qualify for the studies, patients should not have experienced more than 1 severe hypoglycemic event during run-in and more than 2 in the previous 6 months

significant treatment differences seen for the whole population or for adults and children separately (Table 3.2.3). The time to recurrent events models all suggest an increased risk of severe hypoglycemia for INH over SC with risk estimates ranging from 1.6 to 2.25; however, when the dependency among recurrent records is accounted for in the model, the risk is not statistically significant for adults and children separately.

Table 3.2.3 Results from different PH models for severe hypoglycemia in Study 107

Model	Population	p-value	Risk Ratio (95% CI)
Applicant AG model	All	0.002	2.02 (1.3, 3.5)
Reviewer robust variance-AG model	All	0.044	2.02 (1.0, 4.0)
Time to first event model	All	0.55	1.19 (0.67, 2.1)
Applicant AG model	Adults	0.003	2.25 (1.3, 3.9)
Reviewer robust variance-AG model	Adults	0.07	2.25 (0.95, 5.3)
Time to first event model	Adults	0.36	1.4 (0.68, 2.85)
Applicant AG model	Children	0.23	1.6 (0.7, 3.6)
Reviewer robust variance-AG model	Children	0.37	1.6 (0.6, 4.7)
Time to first event model	Children	0.87	0.92 (0.35, 2.39)

A risk ratio below 1 favors INH.

In addition this reviewer ran analyses to determine the impact of the one adult patient with 12 events on the results by leaving this patient out of the analysis of adults. Results from both models show non-significant results; for the applicant's model, the p-value goes from 0.003 with the patient with 12 records to 0.09, leaving that one patient out (Table 3.2.4).

Table 3.2.4 Results from PH models

Model	Population	p-value	Risk Ratio (95% CI)
Applicant AG model w/o pt w/12	Adults	0.09	1.65 (0.9, 2.9)
Reviewer RV-AG model w/o pt w/12	Adults	0.21	1.65 (0.8, 3.6)

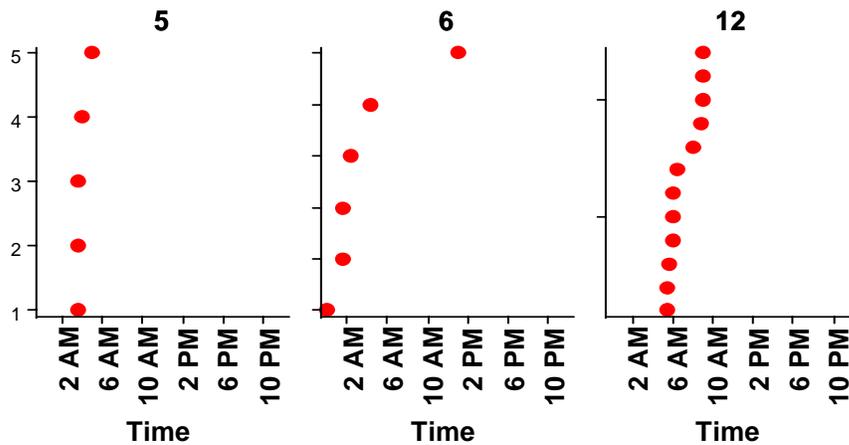
A risk ratio below 1 favors INH.

For the FDA-defined events, the applicant reports a highly significant ($p < 0.0001$) risk ratio of 0.72 using the AG model; with the RV-AG model, the results are less significant with $p = 0.033$ and a CI of 0.54 to 0.97. Dropping out of the analyses four patients with more than 50 events each gives a risk ratio of 0.85 ($p = 0.25$, RV-AG model).

More work is required to determine which model best fits the data collected, including considering other models, so this reviewer can only note at this point that clearly including an adjustment for correlation among events increases the variance and the width of the confidence while the estimates of risk are unchanged. Inspection of the data suggests that recurrent events should not be considered independent and that the reviewer's RV-AG model would be preferred over the applicant's model. What remains uncertain is whether the RV-AG model is a good fit or whether alternative models should be considered.

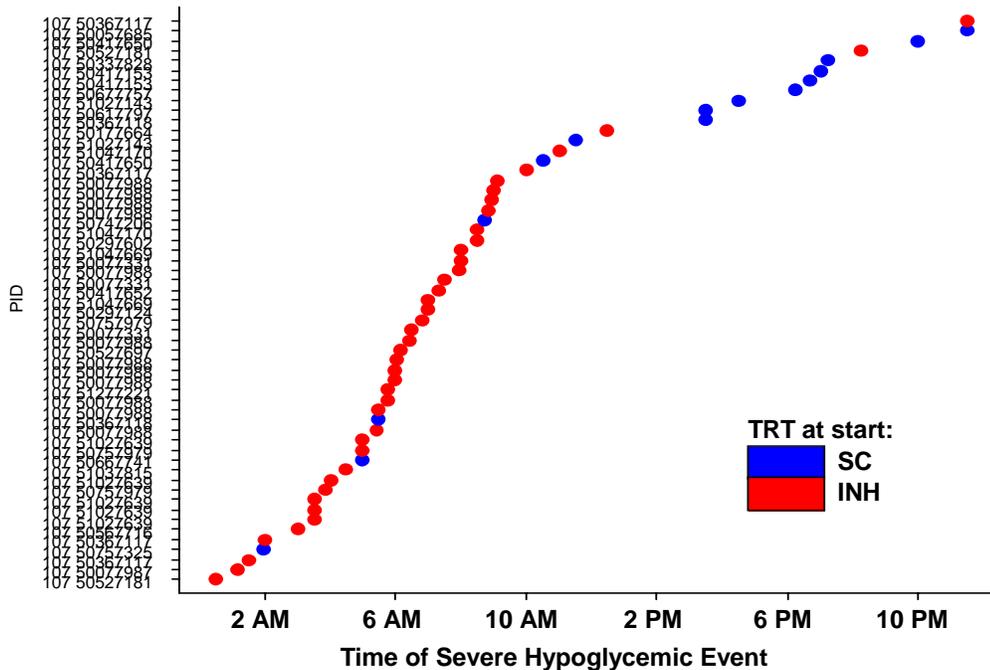
In addition to examining the time on study of the hypoglycemic events, this reviewer looked at the time of day of the events. For the 3 patients with the most events (patients with 5 and 12 events are adults while the patient with 6 was a 15-year old), all except 1 occurred in the early morning. This is consistent with the results from the 24-hour monitoring (see Figure 3.1.1.1) where a drop in glucose was seen in the morning.

Figure 3.2.2 Time of day of severe hypoglycemic events for 3 INH patients with the most events
3 INH Patients with Multiple Severe Hypoglycemic Events



Looking at the data for all adults who experienced a severe hypoglycemic event, there is a preponderance of events in the early morning hours.

Figure 3.2.3 Study 107 Time of day of severe hypoglycemic events for all 58 events in adults
Adults



So overall the risk of severe hypoglycemic events for adults in Study 107 appears to be comparable for the treatment groups although more work is needed to identify the best statistical model for the data. A few INH patients had several events; the records for these patients should be more carefully examined to determine if special circumstances led to the large number of events. For FDA-defined events, there appears to be a reduced risk in INH patients according to

the recurrent events model though other analyses suggest no treatment differences. In the models for recurrent events, it is clear that patients with many events carry considerable weight in the estimate of the risk ratio. Dropping out a single patient with many events has a large effect on the estimate suggesting that the model is too sensitive and may overestimate the risk (in either direction) when there are patients that may be considered outliers.

The 107 data clearly shows a higher risk in the morning for severe hypoglycemic events in INH patients than SC patients.

In Study 108, only 5 episodes of severe hypoglycemia were observed; 2 INH patients each experienced one event and 1 INH patient experienced 2 events, only 1 SC patient experienced a severe hypoglycemic event (the difference between the groups was about 1%, $p > 0.30$, chi square test). Only 2 (1 INH and 1 SC) of the 5 events occurred in the morning, both at 8AM.

A majority of the patients in each treatment group experienced at least one protocol-defined hypoglycemic event (Table 3.2.5, 76% INH vs. 71% SC, $p = 0.34$, chi square test). Almost ¼ of the patients experience an FDA-defined hypoglycemic event (23% INH vs. 19% SC, $p = 0.34$, chi square test, OR 1.3 with CI 0.75 to 2.3) For both types of events, the distribution of recurrent events is similar for the two groups ($p > 0.40$, Wilcoxon rank sum test).

Table 3.2.5 Study 108 Hypoglycemic events

	% (n/N) of pts. with at least 1 event ¹	Number of patients with “n” events							
		0	1	2-5	6-10	11-15	16-20	21-25	>25
Protocol-defined all									
All pts									
INH	76% (111/146)	35	20	38	16	12	10	3	12
SC	71% (106/149)	43	13	40	15	14	5	4	15
FDA-defined									
All pts									
INH	23% (34/146)	112	20	6	1	3	2	2	0
SC	19% (28/149)	121	12	4	5	1	2	0	4

For Study 108 (Type 2 patients), the applicant reported a significantly reduced risk of overall hypoglycemia (Table 3.2.6) for patients treated with INH compared to patients treated with SC. A PH model using a robust variance showed no statistically significant difference between the groups ($p = 0.50$). Similar results are seen for the FDA-defined events.

Table 3.2.6 Risk ratios for protocol-defined overall hypoglycemic events and for “FDA-defined” events

	Risk ratio (CI)	p-value
Protocol-defined		
Applicant’s AG model	0.89 (0.82, 0.96)	0.003
Reviewer’s RV-AG model	0.89 (0.62, 1.3)	0.50
“FDA-defined”		
Applicant’s AG model	0.8 (0.6, 1.07)	0.14
Reviewer’s RV-AG model	0.8 (0.4, 1.6)	0.53

Again, as for Study 107, more work is needed to see which model is a good fit for the data. Clearly though ignoring the dependency of events within patients is not appropriate.

This reviewer observed no patterns by time of day or days after randomization (even for patients

¹ This reviewer was not able to replicate the exact numbers reported by the applicant. The applicant reported 109 (76%) events for INH patients and 104 (72%) for SC patients.

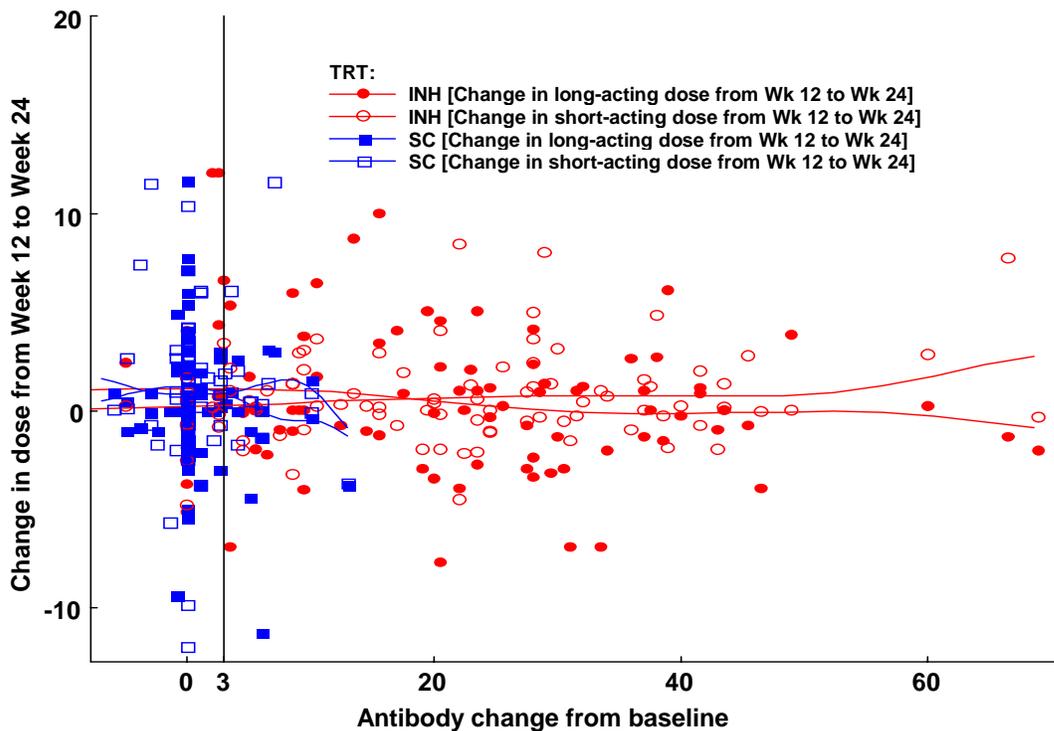
with many events).

Antibody levels and dosing

Higher insulin antibody levels were seen for patients treated with INH, particularly Type 1 patients, compared to patients treated with SC or OA's (see Appendix 6.10 for a boxplot of antibody data from combined studies). The applicant examined the impact of antibody production on dose requirements, glycemic control and adverse clinical outcomes and concluded that there was no association between level of antibodies and the aforementioned outcomes.

Under the guidance of Dr. Mahoney, this reviewer examined further the relationship of antibody level to dosing and to hypoglycemia. For dosing, we restricted our review to Study 107, a study in Type 1 patients, because the largest effect on antibodies is seen in Type 1 patients and Study 107 had two treatment arms that only differed in type of short-acting insulin used (INH versus SC). Since increases in antibodies usually occur after a few months of therapy, we looked at the association of change in dose for both long and short-acting insulin between Weeks 12 and 24 with the endpoint antibody count. We saw no indication that either long or short-acting doses were increased to compensate for increased antibody production; the smoothed lines in Figure 3.2.4 illustrate the lack of relationship between dose change and antibody change. The applicant combined all Phase 2 and 3 studies and also reported no association between dose and antibody level.

Figure 3.2.4 Change in short and long-acting insulin doses from Week 12 to Week 24 versus change in antibodies at endpoint in Study 107 (Type 1 patients)



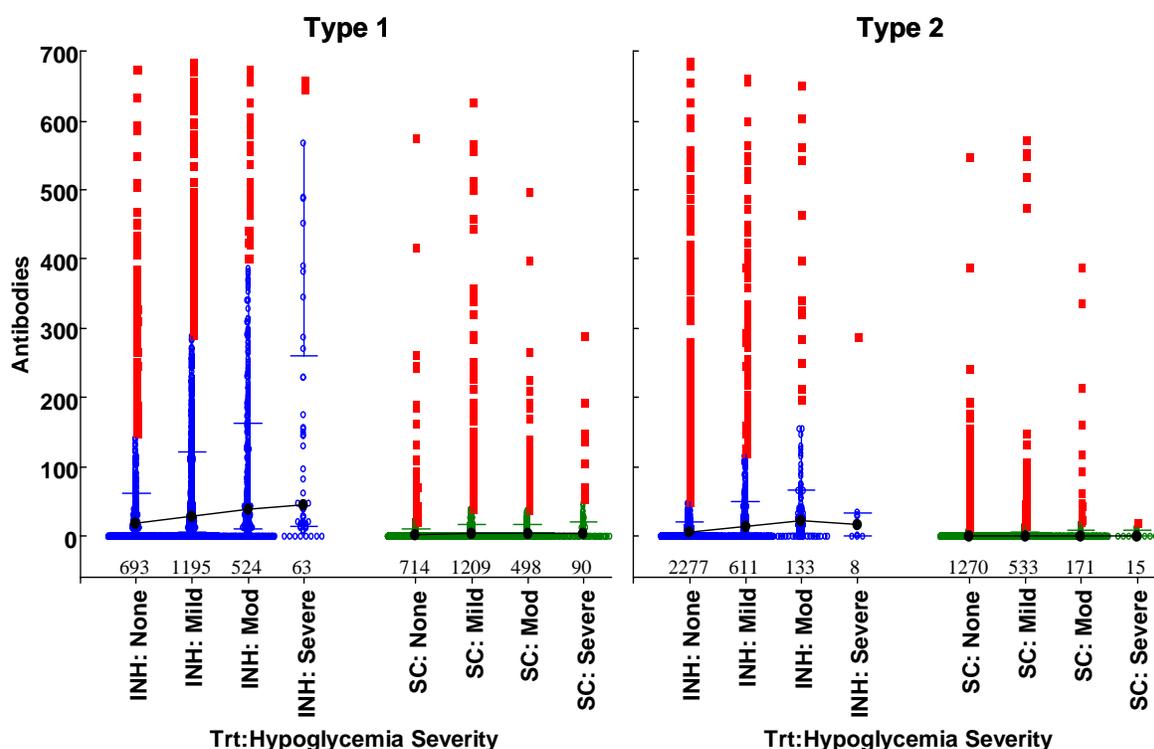
Antibody levels and hypoglycemia

This reviewer also looked at the relationship between hypoglycemia and antibody levels. For the applicant's examination of this issue, a more stringent definition of hypoglycemia (see Appendix 6.9 for the protocol definition) was used with glucose level at 36 or less OR patient required

assistance. The applicant looked at monthly incidence of hypoglycemia versus end-of-study antibodies and reported no association.

For our FDA analysis, we looked at the antibody count versus the level of hypoglycemia (0=no hypoglycemia to 3=severe hypoglycemia) in the combined Phase 2/3 studies containing antibody data; antibody data was matched to hypoglycemia event by date (only data on treatment was considered). The distributions of antibodies for Type 1 and Type 2 patients are shown in the boxplots below; essentially no antibodies are seen for patients treated with oral agents. This is a rather crude depiction of the data in that there is no regard for multiple observations per patient; however further examination of the data by time and by patient for patients with multiple severe events indicated that the overall impression from this graph is reasonable.

Figure 3.2.5 Boxplots of antibody levels by observed hypoglycemia severity for INH and SC patients (the graphs are truncated such that outliers ranging from 700 to over 5,000 are not visible, outlier ranges were unrelated to hypoglycemic severity) RED=OUTLIER



It is clear that higher levels of severity of hypoglycemia are associated with higher levels of antibodies. The relationship between an increase in antibody level and the chance of developing moderate or severe hypoglycemia is statistically significant in patients treated with INH¹ but the risk is not appreciable. For example, increases in antibody levels as high as 500 do not even result in a doubling of the risk (odds ratio of about 1.4 for all patients, 1.2 for Type 1 patients only). So though there appears to be a relationship between the two measures, antibody level does not appear to be a strong predictor of moderate to severe hypoglycemia in patients treated with INH.

¹ Results based on a logistic regression model. Note that change in antibody level is not predictive of moderate or severe hypoglycemia in patients treated with SC (p>0.15).

4. Findings in Special/Subgroup Populations

The Phase 3 trials in the Exubera development program each utilized a unique trial design, so studies should not be pooled to examine subgroups. This reviewer checked subgroups in each study and will report here only notable interactions of subgroup and treatment for HbA1c change from baseline.

4.1 Gender, Race and Age

For Type 1 patients, there was no difference in treatment effects due to age with elderly defined as 50 or more (there are no patients over 65, see Table 3.1.1.3) or due to gender. The vast majority of patients were Caucasian so effects by race could not be assessed.

For Type 2 patients, there was no difference in treatment effects due to age with elderly defined as 65 or more or due to gender. As for the Type 1 patients, the vast majority of patients were Caucasian so effects by race could not be assessed.

4.2 Other Special/Subgroup Populations

4.2.1 Children

The applicant is not seeking an indication in children, however data from children was provided in Type 1 Studies 106, 107 and 1009. The results for Studies 106 and 107 show essentially no difference between INH and SC in patients under 18 years.

Table 4.2.1.1 HbA1c results for children in Studies 106 and 107

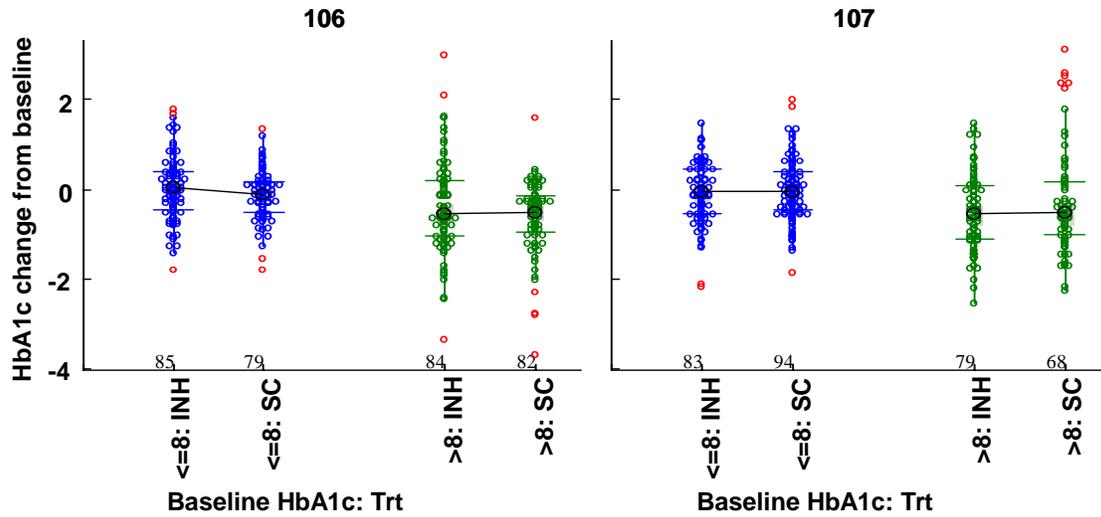
	Study 106			Study 107		
	INH Mean (SD)	SC Mean (SD)	LSM Diff 95% CI	INH Mean (SD)	SC Mean (SD)	LSM Diff 95% CI
HbA1c Baseline	n=33 8.6 (1.0)	n=29 8.5 (0.8)		n=59 8.3 (0.9)	n=59 8.3 (0.9)	
Change from baseline Wk 24 LOCF	0 (1.2)	-0.3 (0.7)	+0.3 (-0.09, 0.7)	-0.2 (0.8)	0 (1.1)	-0.2 (-0.5, 0.1)

Results from a 12-week study of children aged 6-11 (Study 1009) showed results similar to Study 107 with a treatment difference of -0.2 and 95% CI of -0.5 to 0.03.

4.2.2 Baseline HbA1c

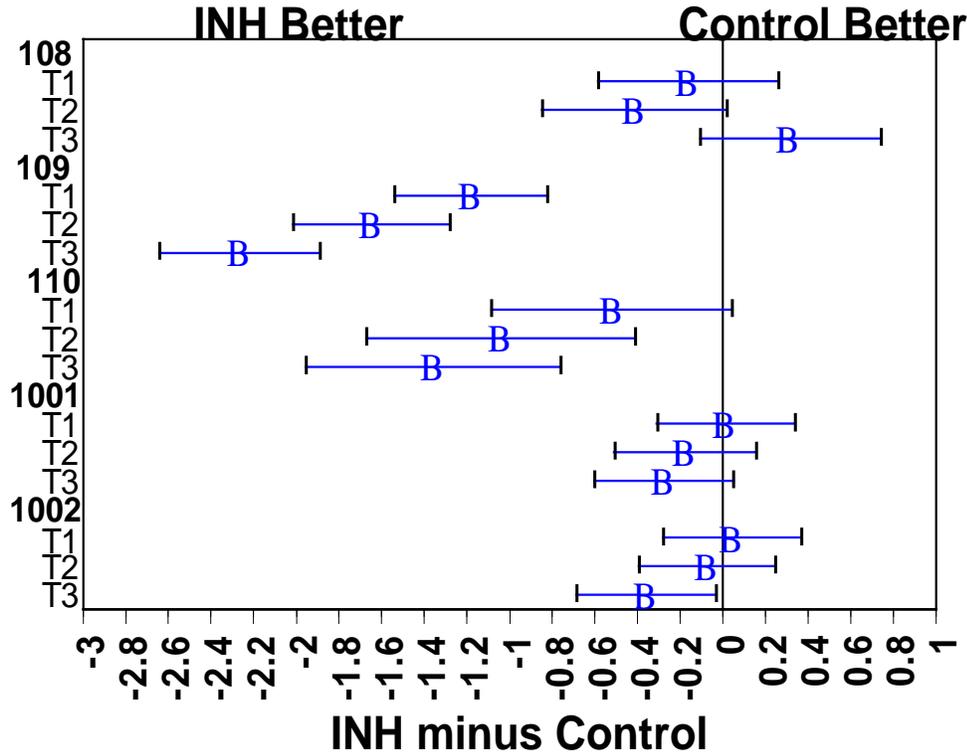
In Type 1 patients, more lowering of HbA1c is seen for larger baseline values regardless of treatment as can be seen in the boxplots of Figure 4.2.2.1. Tests for interaction yielded p-values greater than 0.30.

Figure 4.2.2.1 Boxplots of Week 24 HbA1c change from baseline by median baseline HbA1c for Studies 106 and 107



In all five Phase 3 trials of Type 2 patients, changes in HbA1c are baseline-related with p-values for the interaction of HbA1c change and baseline ranging from 0.001 to 0.05 in the individual studies. This reviewer looked at the treatment effects by tertiles (tertiles are defined within study) and found no statistically significant interaction in Studies 1001 and 1002 while, for the other 3 studies, the interaction remained significant (Figure 4.2.2.2 on the following page).

Figure 4.2.2.2 Treatment difference and 95% CI by baseline study and baseline tertiles.



With increasing baseline, the treatment effects appear to favor INH over comparator, with the exception of Study 108 where the comparator is SC insulin. It is worth recalling that in Studies 109 and 110, the comparators are fixed doses of oral agents so it is not surprising that INH (a titrated drug) would appear more efficacious as baseline increased. This analysis suggests that against titrated medications, INH appears to be “equally” efficacious and offers no significant advantage over SC insulin or titrated oral agents (metformin and glibenclamide).

5. Summary and Conclusions

5.1 Statistical Issues and Collective Evidence

The following statistical issues¹ arose during the process of this review:

- Randomization procedures described in the protocols and in the study reports were not the procedures actually carried out. Patients were not assigned using a minimization algorithm but instead were randomized centrally blocking on center (see page 15 for more details).
- All trials in the Exubera submission were open-label trials. The lack of blinding introduces the possibility of bias in several aspects of the trial. The first aspect considered by this reviewer was the enrollment of patients. The lack of blinding can lead to selection bias in that the randomization code could be broken based on the patients already entered in the trial. An investigator may be able to guess the treatment assignment for the next patient. Guessing is more difficult if the randomization is carried out from a central office, if stratification is not done by center and if block sizes are unknown by the investigator (i.e. not mentioned in the protocol). The first and third conditions were in place in the Exubera trials. Stratification was done by center so it is possible that the pattern of treatment assignment could be discerned and that assignment would be predictable for some patients. To test if selection bias was an issue this reviewer performed a test described by Berger and Exner on the data from Study 107 (the most important Type 1 study). Basically this test determines whether the probability of having a good response is related to the probability of being assigned to the test drug; this is essentially testing whether patients with a good prognosis are more likely to have been assigned to Exubera. This reviewer found no evidence of selection bias based on the results of this test.
- The lack of blinding can also bias the measurement of both efficacy and safety measures. HbA1c is an objective measure not likely to be affected by knowledge of treatment, particularly since dosing of patients was well-controlled by specific parameters spelled out in the protocols and inspection of the data suggests that dosing was adjusted as would be expected. There is some evidence that safety measures may have been affected by the lack of blinding. More specifically, the medical reviewer, Dr. Karen Mahoney, carefully describes the misclassification of discontinuation reasons as “withdrawn consent” or “subject request” where the data suggests that the reason was “adverse event”. She found a larger number of misclassifications in the INH group than in the comparator group suggesting the possibility that knowledge of treatment could have played a role in the naming of the discontinuation reason.
- The 12-week duration of Study 110 was insufficient to provide a fair comparison of rosiglitazone to INH (see page 26).
- Patients who discontinue from therapy due to hypoglycemia may provide HbA1c LOCF data that suggests a beneficial result when in reality the therapy was a failure for that patient. This reviewer examined the data to determine if the LOCF estimates were biased by the use of such dropout data and found no evidence that this was the case.
- Noninferiority trials were powered to rule out a treatment difference for HbA1c change from baseline of 0.5% while 0.4% is the margin usually used by FDA. This, however, was not an issue since the boundary of 0.4% was met in all relevant trials.
- The Type 2 development program consisted of five Phase 3 trials; 3 conducted in North America and 2 conducted in foreign countries. The majority of the investigator sites participated in more than one study. For Studies 1001 and 1002 (foreign studies) and for Studies 108, 109 and 110, about half of the total sites participated in more than 1 study (see

¹ The statistical issues listed here are identical in wording to the list provided in Section 1.3 of this review.

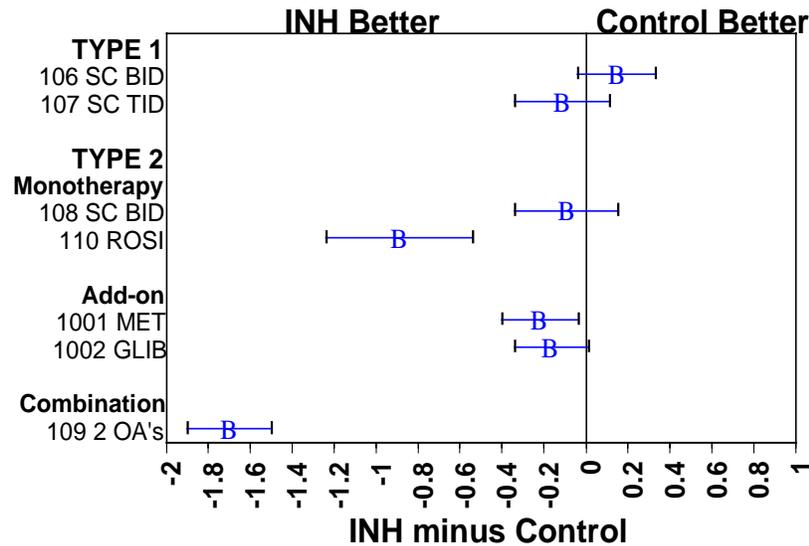
page 19 for additional details). Patients were not allowed to participate in more than one study and, in addition, entry criteria differed making patients ineligible for more than one study so this reviewer was not concerned that patients may have been retested. Enrollment dates were overlapping for the studies so investigators would be seeing patients during the same timeframe from the different studies suggesting that experience from participating in one trial would not carryover and affect the conduct of a subsequent trial. Also the sample size of each site is small and no one site would greatly influence the outcome. Though overlapping sites could affect the independence of the trial results, this reviewer did not feel, for these studies, that independence was compromised for the reasons given.

- The applicant has computed risk ratios for hypoglycemia using a recurrent events proportional hazards model that assumes events are independent. This reviewer does not think that the applicant's model is appropriate primarily because it ignores the dependency among events within patients. Also additional analyses suggest that the model is oversensitive to patients with many events; for example, dropping just one patient from an analysis of Study 107 data changed the estimate of risk from 2.25 to 1.65. This reviewer did not have sufficient time to determine the best model to describe the hypoglycemia data; however an odds ratios for at-least one event and a comparison of number of events (e.g. Wilcoxon rank sum test) seems to describe the data well if it is agreed that time is not an important factor.

[This space purposely left blank.]

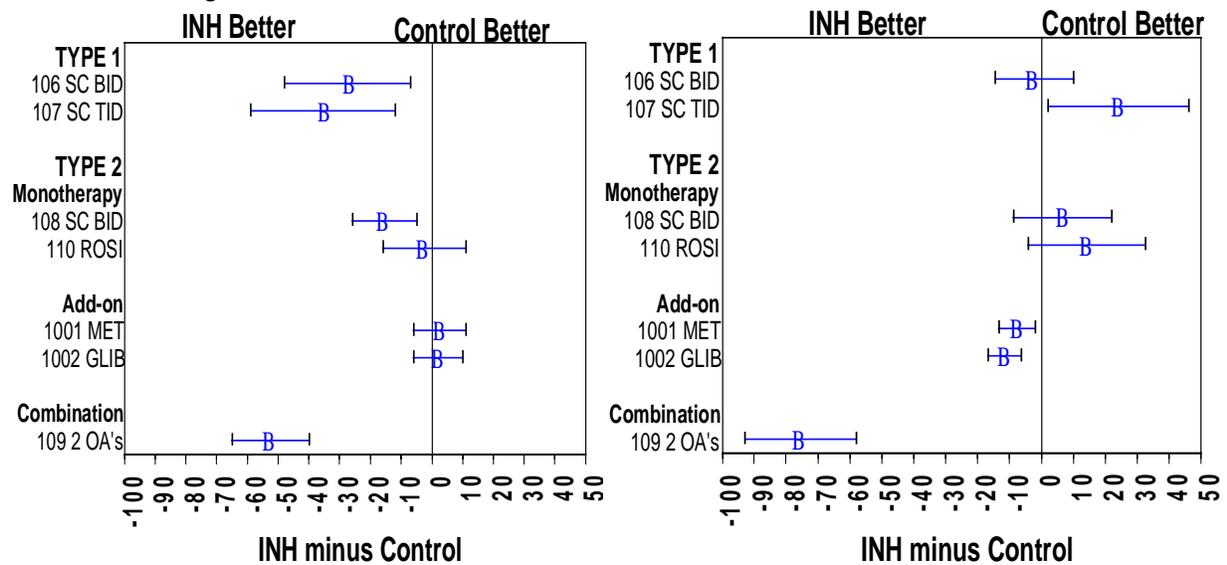
The collective evidence of efficacy from the seven Phase 3 trials in patients with Type 1 and Type 2 diabetes is summarized by the mean treatment difference for HbA1c change from baseline shown in Figure 3.1.3.1 below. The left edge of the graph shows the study number with the control group. For more details regarding these results see Section 3.1.2.2 of this review. A discussion of conclusions drawn from these results follows in Section 5.2.

Figure 3.1.3.1 HbA1c Week 24 change from baseline; LS mean treatment difference with 95% confidence interval



The results for FPG and post-prandial increment, two important secondary endpoints, are summarized in the figures below.

Figure 3.1.3.2 FPG (left graph) and Post-prandial increment (right graph) Week 24 change from baseline; LS mean treatment difference with 95% confidence interval



5.2 Conclusions and Recommendations

This reviewer has the following comments and conclusions¹ based on the statistical review of the seven Phase 3 trials submitted to demonstrate the efficacy and safety of INH (Exubera). Note that Figures 3.1.3.1 and 3.1.3.2 on page 41 summarize the efficacy data.

- Approximately 85 to 90% of the Type 1 and 2 patients studied were Caucasian; so the races were not adequately represented in the database
- INH was shown to be non-inferior to SC insulin in patients with Type 1 diabetes in two Phase 3 clinical trials, 106 and 107 (see Table 3.1.1.5 and Figures 3.1.3.1 and 3.1.3.2). The HbA1c results are more favorable to INH in Study 107, a study where SC insulin was given TID and the same long-acting insulin was given in both groups. There was no evidence that long-acting insulin was titrated differently in the two groups allaying concerns that the long-acting dosing may have compensated for inadequate treatment on short-acting insulin.
- For Type 1 patients, statistically significant differences in favor of INH over SC were seen for the change from baseline in FPG in both studies. An examination of 24-hour diary records suggest a drop in glucose levels overnight in the INH group may explain the significantly lower FPG. This issue is discussed in the clinical review by the FDA medical reviewer, Dr. Mahoney.
- Across the Type 2 studies, variability in several baseline parameters was seen suggesting some heterogeneity among the patient populations which could improve generalizability of the results.
- For Type 2 patients inadequately treated with metformin or a sulfonylurea, add-on INH yielded HbA1c lowering comparable to adding either glibenclamide or metformin, respectively (Studies 1001 and 1002, see Table 3.1.2.2.8 and Figure 3.1.2.2.1). For patients inadequately treated with two oral agents, adding INH resulted in highly significant drops in HbA1c (Study 109, see Table 3.1.2.2.6 and Figure 3.1.2.2.1).
- Type 2 patients on SC insulin therapy were able to maintain their HbA1c levels when switched to INH insulin (Study 108, see Table 3.1.2.2.5 and Figure 3.1.2.2.1)
- Naïve Type 2 patients randomized to either rosiglitazone (4 mg BID, the most effective marketed dose) or INH showed statistically significantly more lowering of HbA1c and FPG on INH in Study 110 (Table 3.1.2.2.7) ; however the length of the trial at 12 weeks provided inadequate time for rosiglitazone to show a full effect (Figure 3.1.2.2.1). This reviewer concludes that Study 110 was inadequate by design and the results should not be included in labeling.
- Antibody counts were significantly higher in INH patients than in patients treated with SC insulin or oral agents. This reviewer found no relationship between change in antibody count and change in dose. A crude analysis of antibody count by severity of hypoglycemia suggests that higher levels of antibodies may be associated with more severe levels of hypoglycemia, although antibody level does not appear to be a strong predictor of moderate to severe hypoglycemia in patients treated with INH (see pages 35 to 36 for more details.)
- The survival analysis model for recurrent events used by the applicant to estimate risk ratios for hypoglycemia is not an appropriate model because it assumes that the events are independent which is clearly not the case for clusters of recurrent events within patients. No conclusions can be drawn based on the risk ratios provided by the applicant. For a discussion of this issue for Studies 107 and 108, see pages 27-34.

Since Exubera will be discussed at an upcoming advisory committee meeting, no

¹ This section is identical to Section 1.1 of this review.

recommendations regarding approval based on the statistical evidence are being made at this time.

6.0 Appendices

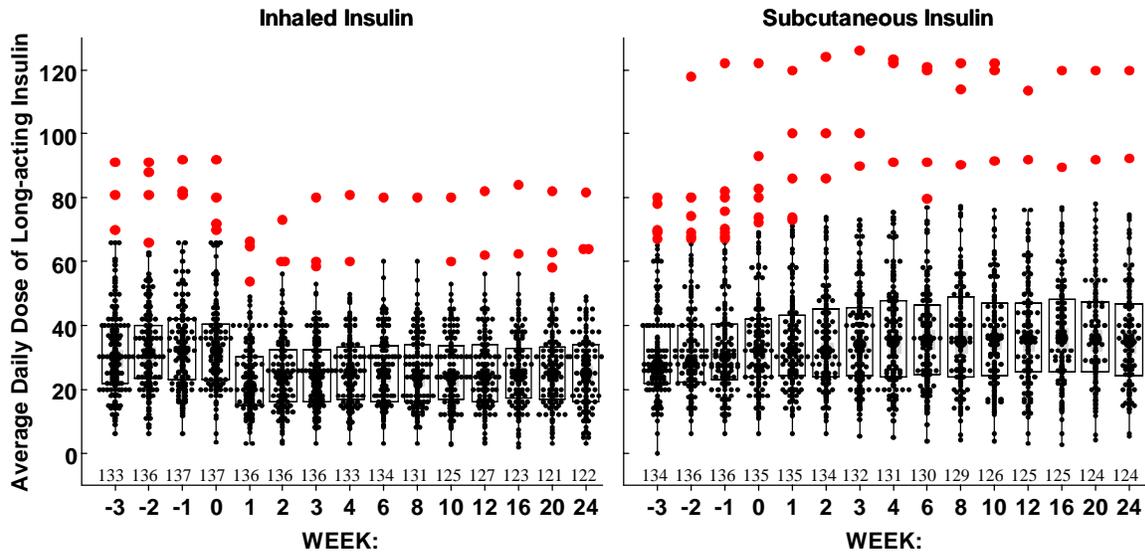
6.1 Ongoing safety studies

Clinical safety studies ongoing at the time of the NDA submission

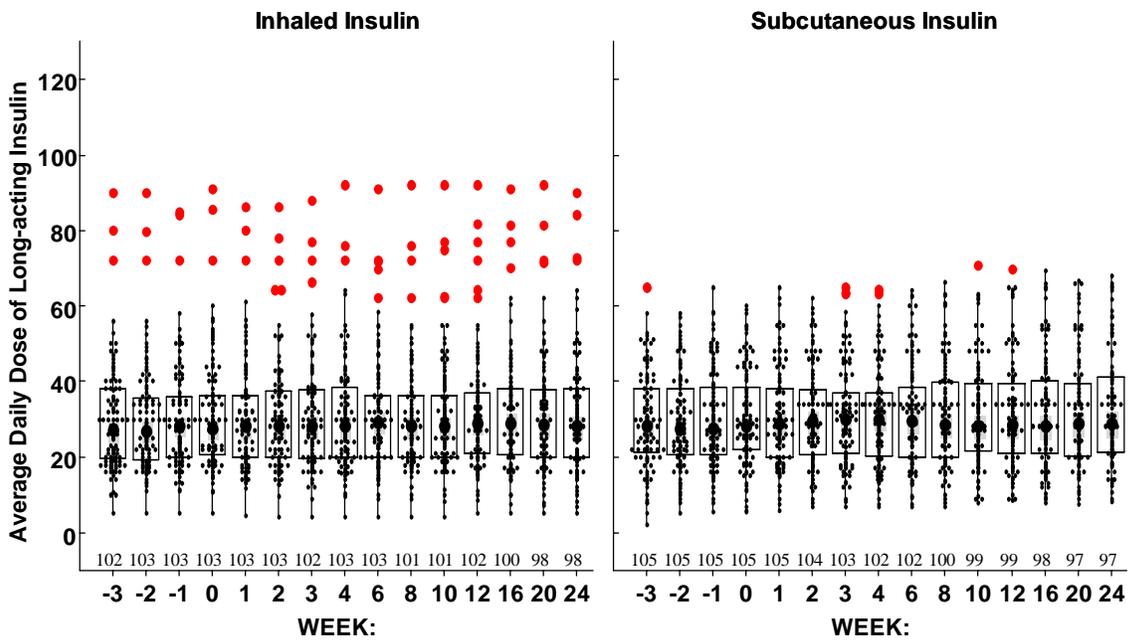
Study (# of centers)	Inhaled Insulin group	Comparator	Duration of treatment
1017 Type 2 (US) primarily an efficacy study	add-on INH to MET+SU INH + MET (switch from SU to INH)	add-on ROSI to MET+SU Total of 74 pts. rand. in all 3 arms	1 year started 4/03 completion expected 11/06
1022 Type 1 (US, Can, Arg., Brazil, Mexico) Safety study Pulm. Fn. tests	INH regimen not described 290 rand. Enrollment is complete Synopsis, no report	SC regimen not described 290 rand.	4 wk run-in on SC 2 year Median Dur 456 days ~75% > 1 yr exposure
1028 Type 1+2 (US, Can, Ger., Braz., Mex., Costa Rico) Chronic asthma Safety Study	Pre-meal TID + QD or BID ultralente or NPH, or QD ins glargine 46 rand. Enrollment ongoing	BID or TID reg ins + QD or BID ultralente or NPH, or QD ins glargine 49 rand.	3 wk run-in SC short- acting ins 1 year randomized treatment 6 wk runout on SC
1029 Type 2 (US, Can, PR, Brazil)	INH 291 rand.	SC 291 rand.	3 year extended treatment period completion expected 5/09
1030 Type 1+2 (US, Can, Ger., Braz., Mex., Costa Rico) Chronic obs. pul. dis. Safety Study	Pre-meal TID + QD or BID ultralente or NPH, OR QD ins glargine OR Oral antidiabetic agent 30 rand. Enrollment is complete	BID or TID reg ins + QD or BID ultralente or NPH, OR QD ins glargine OR Oral antidiabetic agent 27 rand.	3 wk run-in SC short- acting ins 1 year randomized treatment 6 wk runout on SC
1036 Type 1+2 Extension of 102, 103, 104 Pulm. Fn tests	All pts on INH regimen 173 rand. in orig prot 62 ongoing	NA	4 year extension study

6.2 Average daily dose of long-acting insulin by study, treatment and week for Type 1 studies 106 and 107

Study 106

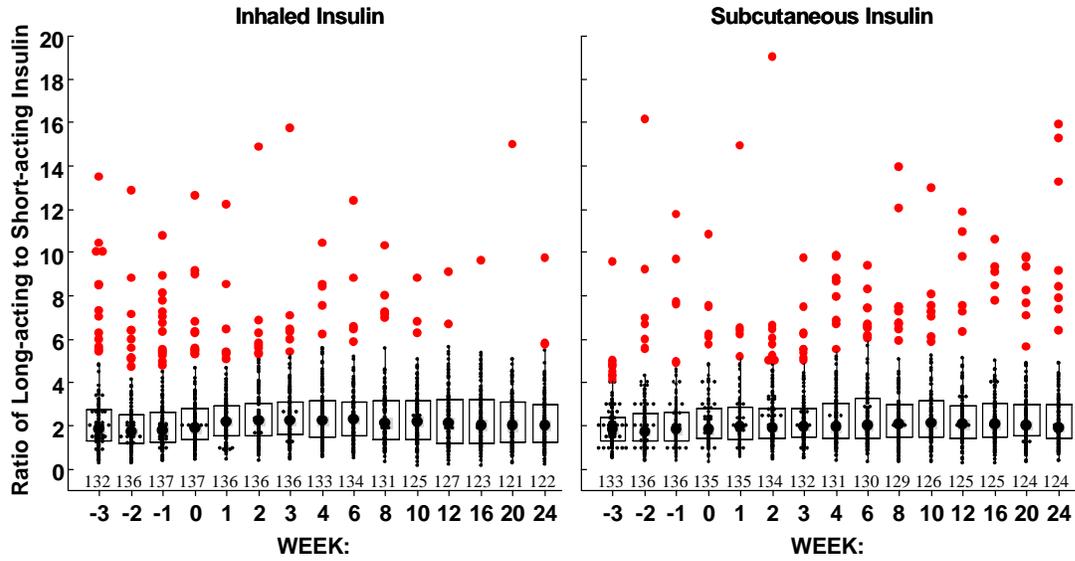


Study 107

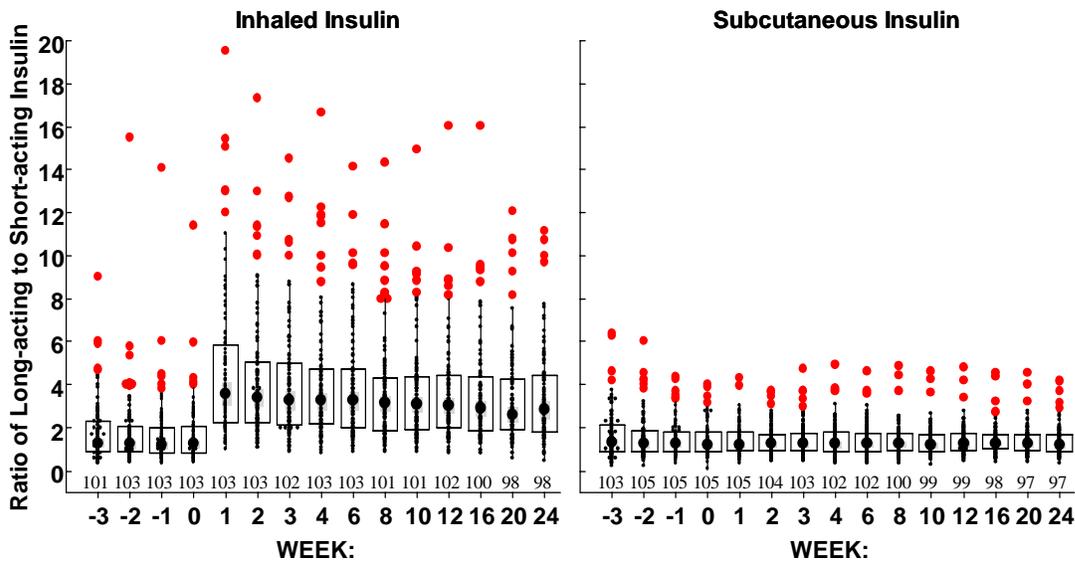


6.3 Average daily ratio of long-acting to short-acting insulin by study, treatment and week for Type 1 studies 106 and 107

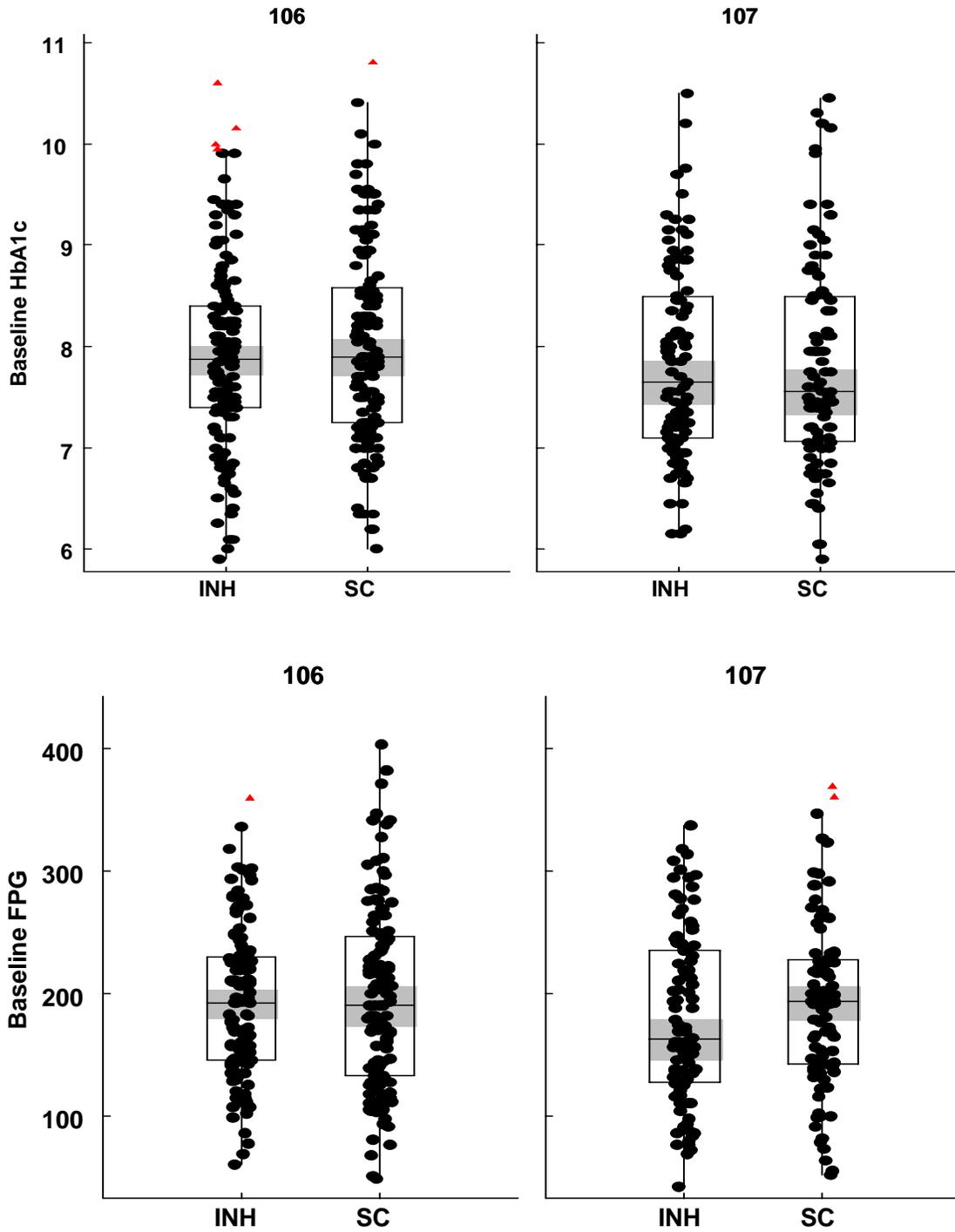
Study 106



Study 107



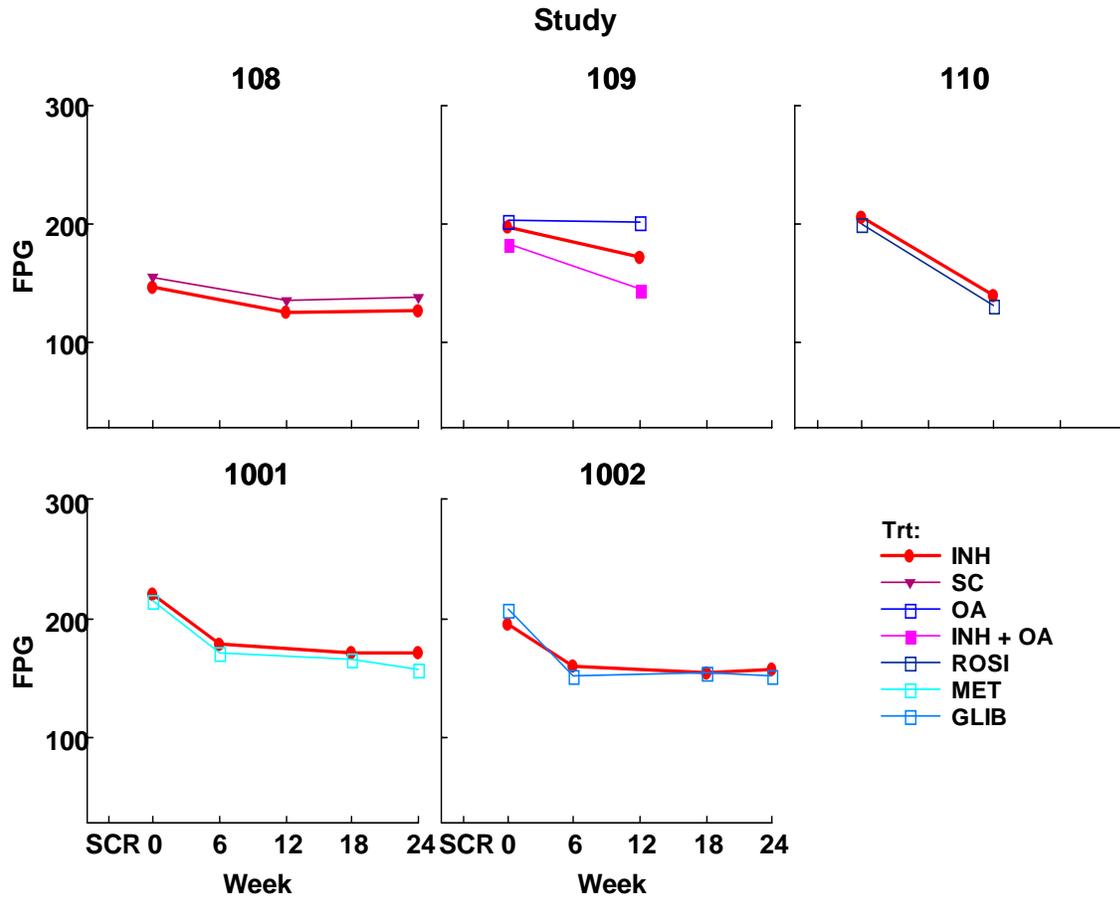
6.4 Baseline FPG and HbA1c for Type 1 Studies 106 and 107



6.5 Oral anti-diabetic therapies used in Phase 3 studies of Type 2 diabetes

Study	Screening	Run-in	On-study
109	<p>Stable regimen of 2 OA's</p> <p>1. Sulf. or repaglinide glipizide\geq10mg daily or glimepiride\geq4mg daily or glyburide\geq10mg daily or repaglinide\geq2mg TID</p> <p>2. glitazone or metformin metformin\geq1.7mg daily or troglitazone\geq400mg daily rosiglitazone 4mg daily pioglitazone\geq30mg daily</p>	<p>Maintain screening dose</p> <p>Pts on troglitazone were switched to another glitazone when it was removed from the market</p>	<p>Maintain screening dose for patients remaining on 2 OA's</p>
110	<p>NA (diet and exercise only)</p>	<p>NA (diet and exercise only)</p>	<p>For pts rand. to rosiglitazone: fixed dose of 4 mg BID (maximum marketed dose)</p>
1001	<p>Sulfonylurea alone</p> <p>glibenclamide (stand.)\geq10mg daily or glibenclamide (micro.)\geq7mg daily or glicazide\geq160mg daily or glimepiride\geq3mg daily</p>	<p>Maintain screening dose</p>	<p>Maintain screening dose of sulf. for all pts</p> <p>For pts rand. to metformin: starting dose of 500 mg and titrated, based on FPG, to max of 2.5 g daily over a minimum of 6 weeks or maximum of 18 weeks</p>
1002	<p>Metformin\geq1.5g daily</p>	<p>Maintain screening dose</p>	<p>Maintain screening dose of met. for all pts</p> <p>For pts rand. to glibenclamide: starting dose of 2.5 mg and titrated, based on FPG, to max of 5 mg twice daily over a minimum of 6 weeks or maximum of 24 weeks</p>

6.6 Type 2 FPG overtime all studies



6.7 Hypoglycemic events for all Phase 3 Studies

The table below shows the number of events observed for all 3 definitions of hypoglycemia for the Phase 3 studies reviewed in this document. In each box, the top number is the sample size, the middle number is the number of patients with at least one event and the bottom number is the total number of events.

	Total		"FDA"		Severe	
TYPE 1						
	INH	Control	INH	Control	INH	Control
Study 106	136	132	136	132	136	132
At least 1 event	134 (99%)	130 (98%)	118 (87%)	116 (88%)	24 (18%)	19 (14%)
All events	6303	6454	1357	1315	43	35
Study 107	103	103	103	103	103	103
At least 1 event	102 (99%)	101 (98%)	91 (88%)	94 (91%)	18 (17%)	13 (13%)
All events	5306	5915	971	1327	43	19
TYPE 2						
Study 108	146	149	146	149	146	149
At least 1 event	111 (76%)	106 (71%)	34 (23%)	28 (19%)	3	1
All events	1109	1301	80	104	4	1
Study 109	102	INH+OA 100	102	INH+OA 100	102	INH+OA 100
At least 1 event	68 (67%)	78 (78%)	17 (17%)	21 (21%)	1	0
All events	365	477	23	48	1	0
Study 110	75	67	75	67	75	67
At least 1 event	36 (48%)	5 (7.5%)	9 (12%)	0	0	0
All events	153	9	15	0	0	0
Study 1001	214	196	214	196	214	196
At least 1 event	112 (52%)	53 (27%)	15 (7%)	7 (4%)	1	0
All events	374	180	20	9	1	0
Study 1002	234	222	234	222	234	222
At least 1 event	75 (32%)	69 (31%)	13 (6%)	12 (5%)	0	0
All events	214	203	27	23	0	0

The highlighted numbers were statistically significantly different based on a recurrent events proportional hazards model with a robust sandwich estimate for the covariance matrix to account for dependency among the events within patients. These analyses were performed by the reviewer.

6.8 SAS code for modeling hypoglycemic events

Applicant's code

```
proc phreg;
  model (startday,stopday)*status(0) = trt / alpha=0.05 rl;
```

Reviewer's code

```
proc phreg covsandwich(aggregate);
  model (startday,stopday)*status(0) = trt / alpha=0.05 rl;
  id patid;
```

6.9 Definition of hypoglycemia

The following is excerpted directly from the protocol of Study 107 (the same definition was used in other studies as well); the underlining and bolding was added by the reviewer.

Subjects are instructed to check blood glucose in the event they experience symptoms of hypoglycemia. Subjects will record all hypoglycemic episodes on their weekly glucose monitoring worksheets.

Hypoglycemia is defined as one of the following:

- Characteristic symptoms of hypoglycemia with no blood glucose check. Clinical picture must include prompt resolution with food intake, subcutaneous glucagon, or intravenous glucose.
- Characteristic symptoms of hypoglycemia with blood glucose check showing glucose 59 mg/dl or less. Symptoms associated with a blood glucose of 60 mg/dl or greater cannot be reported as hypoglycemia.
- Any glucose measurement 49 mg/dl or less, with or without symptoms.

Every hypoglycemic event must be characterized with respect to its severity.

In order to characterize the event as severe, all 3 of the following criteria must be met:

1. The subject was unable to treat himself or herself.
2. The subject exhibited at least one of the following neurological symptoms:

- a) memory loss,
- b) confusion,
- c) uncontrollable behavior,
- d) irrational behavior,
- e) unusual difficulty in awakening,
- f) suspected seizure,
- g) seizure,
- h) loss of consciousness.

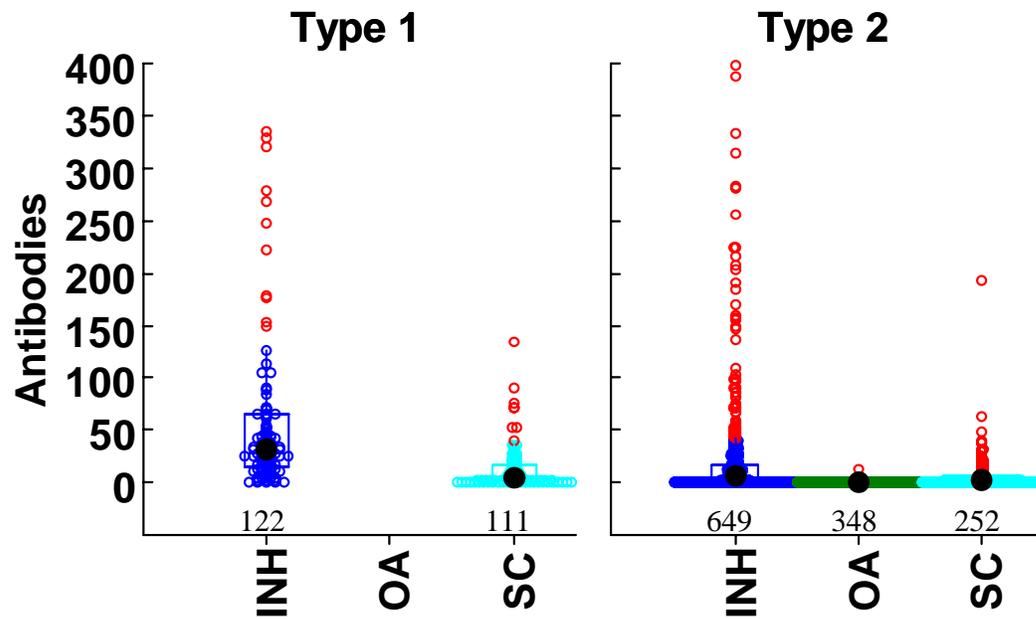
3. Either:

- a) If blood glucose was measured and was 49 mg/dl or less or,
- b) If the blood glucose was not measured, the clinical manifestations were reversed by oral carbohydrates, subcutaneous glucagon, or intravenous glucose.

Events that do not meet all 3 criteria for severe hypoglycemia are characterized as mild-moderate.

The “FDA” definition of hypoglycemia is blood glucose level of less than or equal to 36 mg/dL AND/OR requiring assistance, which was applied retrospectively. According to the program code provided by the applicant, an hypoglycemic event was counted as an FDA event if the patient had an hypoglycemic event as defined in the protocol (see above) and the glucose value recorded on the CRF-Hypoglycemic Episode Report was 39 or less and/or the answer to “Was the subject unable to self treat?” on the CRF was yes.

6.10 Boxplot of antibodies at Week 24 by diabetes type and treatment



A few outliers have been excluded from these boxplots in order to more clearly show the boxplots.