CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

213051Orig1s000

STATISTICAL REVIEW(S)

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

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA:	NDA 213051
Sequence #:	0001
Drug Name:	Semaglutide
Indication(s):	Type 2 Diabetes Mellitus (T2DM)
Applicant:	Novo Nordisk
Date(s):	Received: March 20, 2019 Primary Review: August 15, 2019 PDUFA Due Date: September 20, 2019
Review Priority:	Expedited
Biometrics Division:	Division of Biometrics II
Statistical Reviewer:	Robert Abugov, Ph.D.
Concurring Reviewers:	Yun Wang, Ph.D. (Team Leader) Mark Rothmann, Ph.D. (Acting Division Director)
Medical Division:	Division of Metabolism and Endocrinology Products
Clinical Team:	Andrea Lungu, M.D. (Medical Officer) Mitra Rauschecker, M.D. (Acting Team Leader) Lisa Yanoff, M.D., Ph.D. (Acting Division Director)
Project Manager:	Peter Franks

Keywords: NDA Review, Clinical Studies

Contents

List of Tables

List of Figures

1 EXECUTIVE SUMMARY

Six randomized trials demonstrate effectiveness of semaglutide 7 or 14 mg administered as oral tablets (po) once daily (qd) for the improvement of glycemic control in patients with type 2 diabetes mellitus (T2DM). Four of the trials, 4224, 4233, 4234, and 4280, were conducted with a placebo control. One of the trials, 4224, included both a placebo and an active (liraglutide injection 1.8 mg qd) control, and two trials, 4222 and 4223, were conducted using respective active controls empagliflozin 25 mg po qd and sitagliptin 100 mg po qd.

Compared to the control, five of the six trials showed statistically significant effects of 7 and/or 14 mg semaglutide for the primary endpoint, week 26 change from baseline HbA1c. In trial 4224, however, no significant difference was seen between semaglutide 14 mg and liraglutide 1.8 mg, another drug in the same class. Of the five remaining trials testing statistical efficacy for the key secondary endpoint, week 26 change from baseline body weight, five showed statistically significant effects of both 7 and 14 mg semaglutide against control, and one trial 4223 showed superiority of the 14 mg, but not the 7 mg dose, to placebo. Supportive endpoints, proportion of patients with HbA1c less than 7%, and week 26 change from baseline fasting plasma glucose, all trended in the direction expected if semaglutide is efficacious. Please refer to Section 3.2.4 for detailed efficacy results.

No major statistical issues have been identified in this submission. The collective evidence from primary, secondary, and supportive endpoints in these studies consistently supports effectiveness of oral semaglutide for the treatment of type 2 diabetes. Safety risks associated with use of oral semaglutide are acceptable. Therefore, I recommend approval of oral semaglutide for glycemic control in patients with type 2 diabetes mellitus. See Section 5.3 for further details.

2 INTRODUCTION

2.1 Overview

2.1.1 Drug Class and Indication

Semaglutide is a glucagon-like peptide-1 receptor agonist (GLP-1 RA) proposed as an adjunct to diet and exercise to improve glycemic control in adults with T2DM. It is a GLP-1 analogue modified for resistance to metabolic degradation through albumin binding and additionally modified for resistance to enzymatic degradation by dipeptidyl peptidase-4 (DPP-4). Tablets for oral administration are co-formulated with the absorption enhancer salcaprozate sodium (aka sodium N-8-[(2-hydroxybenzoyl) amino] caprylate, abbreviated as SNAC).

2.1.2 History of Drug Development

Semaglutide for subcutaneous injection was originally approved on December 5, 2017 as an adjunct to diet and exercise to improve glycemic control in adults with T2DM. The present submission proposes this same indication for an orally administered version of this product.

IND 114464 for semaglutide po was opened on September 9, 2013. In minutes from the pre-filing meeting, submitted to DARRTS on December 28, 2018, the applicant agreed to use of the treatment policy estimand for all primary analyses of efficacy. The applicant further agreed that analyses for the Summary of Clinical Efficacy and Integrated Summary of Effectiveness would be conducted without pooling of data across trials. In addition, the Agency agreed that the Integrated Study of Safety need only evaluate studies on oral semaglutide, with a phase 3a pool, placebo pool, and a placebo-dose pool. To facilitate separate analyses, FDA requested that studies conducted only in Japan be flagged in the safety datasets.

FDA further requested that the NDA include the analysis programs, each indexed with a table including the input datasets, macros used, and the output file names, and that core variable names be retained in the analysis datasets across the different studies in the submission.

2.1.3 Data Sources

Data sources for the current review are located at

\\cdsesub1\evsprod\NDA213051\0001\m5\datasets .

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

Data and analysis programs provided by the applicant were consistently well organized. Adequacy of trial design and analyses are further discussed in this review.

3.2 Evaluation of Efficacy

3.2.1 Study Design and Endpoints

This review will focus on six randomized controlled confirmatory trials conducted in adult T2DM patients with a variety of background medications and concentrations of baseline hemoglobin A1c (Table 1).

Table 1. Randomized Confirmatory Trials Reviewed for Effectiveness

Trial	Background	Inclusion HbA1c	Blinding	Treatments	N	Duration (Weeks)
P1 (4233)		7.0-9.5%	DB	S3 S7 S14 Pbo	175 175 175 178	26
P2 (4223)	Met	7.0-10.5%	OL	S14 SGLT2i (E)	411 410	52
P3 (4222)	Met SU±Met	7.0-10.5%	DB	S3 S7 S14 DPP-4i (Si)	466 465 465 467	78
P4 (4224)	Met±SGLT2i	7.0-9.5%	DB DD	S14 GLP1 RA (L1.8) Pbo	284 283 142	52
P5 (4234)	Met±SU SU±Met BasIns BasIns±Met	7.0-9.5% renal imp	DB	S14 Pbo	163 161	26
P8 (4280)	Ins±Met Japan: Met w/BasIns only	7.0-9.5%	DB	S3 S7 S14 Pbo	184 182 181 184	52

source: reviewer

Met metformin, SU sulfonylurea, SGLT2i sodium-glucose co-transporter 2 inhibitor, BasIns basal insulin, Ins basal with or without bolus, or premixed insulin, OAD oral anti-diabetic drug using any of the above therapies, plus thiazolidinedione with or without metformin, renal imp moderate renal impairment, DB double-blind, OL open-label, DD double-dummy, S3 S7 S14 semaglutide maintenance doses 3 mg, 7 mg, or 14 mg qd po, Pbo placebo, DPP-4i dipeptidyl peptidase-4 inhibitor, Si sitagliptin 100 mg qd po, E empagliflozin 25 mg qd po, L1.8 liraglutide injection qd escalated to 1.8 mg, DPP-4i (Si)

Trials omitted from efficacy evaluation in the present review (Table 2) include trial 4221 for cardiovascular outcomes, trial 4257 in which ramped dosing occurred at eight-week intervals, and two randomized trials, 4281 and 4282, conducted solely in Japan.

Table 2. Randomized Confirmatory Trials Omitted from Efficacy Review

Trial	Background	Inclusion HbA1c	Blinding	Treatments	N	Duration (Weeks)
P6 (4221)	SOC - Incretins	not applicable CVOT	DB: Pbo, S not dose	S14 Pbo	1591 1592	58[1]
P7 (4257)	All	7.5-9.5%	OL	S_flex DPP-4i (Si)	253 251	52
P9 (4281) Japan	Monotherapy	on OAD 6.5-9.5% no OAD 7.0-10%	DB: Pbo, S OL: L .9	S3 S7 S14 GLP1 RA (L.9) Pbo	49 49 49 48 49	52
P10 (4282) Japan	Met+other	7.0-10.5%	OL	S3 S7 S14 GLP1 RA (D)	131 132 130 165	52

source: reviewer
S_flex semaglutide flexible dose ramped at eight week intervals, L.9 liraglutide escalated to .9 mg, D dulaglutide 0.75 mg q7d, , CVOT cardiovascular outcomes trial
1 Mean time on trial product in event driven cardiovascular outcomes trial

For subjects randomized to semaglutide, the semaglutide dose was escalated every four weeks, from 3 to 7 to 14 mg, until the randomized maintenance dose was reached.

The primary efficacy endpoint in all of the trials was week 26 change from baseline percent concentration of HbA1c (ΔHbA1c). The secondary efficacy endpoint was week 26 change from baseline body weight (ΔBW). Endpoints proposed for inclusion on the product label, but not included in the analysis hierarchy, included descriptive week 26 proportion of individuals with HbA1c < 7% and week 26 change from baseline fasting plasma glucose (ΔFPG).

During the trials, patients with persistent and unacceptable hyperglycemia were offered intensifications of treatment as rescue (Table 3) according to local or American Diabetes Association/ European Association for the Study of Diabetes guidelines. GLP-1 receptor agonists, DPP-4 inhibitors and amylin analogues were not allowed as rescue medications.

Table 3. Criteria for Provision of Rescue Medications

Trial	Rescue Criteria	Time
P1 (4233)	Investigator discretion	W8 to end-of-trial
P2 (4223)	Investigator discretion	W8 to end-of-trial
P3 (4222)	FPG>14.4 FPG>13.3 FPG>11.1 HbA1c>8.5%	W8 to W13 W14 to W25 W26 to end-of-trial W26 to end-of-trial
P4 (4224)	FPG>13.3 FPG>11.1 HbA1c>8.5%	W8 to W13 W14 to end-of-trial W26 to end-of-trial
P5 (4234)	Investigator discretion	W12 to end-of-trial
P8 (4280)	HbA1c>8.5%	W26 to end-of-trial

source: reviewer
FPG fasting plasma glucose mmol/L, W week

3.2.2 Statistical Methodologies

The primary and secondary endpoints were analyzed using ANCOVA with dependent variable change from baseline and with independent variables treatment, region, baseline value, and stratification factors used in the randomization (Table 4). Adjusted means were calculated at mean baseline values, and used coefficients proportional to the frequencies in the data of the various categorical variables.

Table 4. ANCOVA Statistical Models for Primary and Secondary Endpoints

Trial	Dependent Variable	Independent Variables
P1 (4233)	Week 26 change from baseline	treatment region[1] baseline
P2 (4223)	Week 26 change from baseline	treatment region[1] baseline
P3 (4222)	Week 26 change from baseline	treatment region[1] baseline background medication descent[1]
P4 (4224)	Week 26 change from baseline	treatment region[1] baseline background medication descent[2]
P5 (4234)	Week 26 change from baseline	treatment region[1] baseline background medication renal function[3] interaction[4]
P8 (4280)	Week 26 change from baseline	treatment region[1] baseline background medication[5] descent[2] interaction[4]

source: reviewer
1 Categories include North America, South America, Africa, Asia, Europe,
2 Japanese vs non-Japanese
3 eGFR 30-44 or 45-59 mL/min/1.73 m2
4 interaction between stratification factors
5 (metformin, no metformin) and (basal insulin, basal + bolus insulin, premixed insulin)

The treatment policy estimand was used to evaluate efficacy, with missing data at week 26 imputed using pattern-mixture models. For each study, imputation groups were defined according to randomized treatment and treatment status (remain on treatment without rescue vs discontinued treatment or rescued). Then, for each of these groups, missing data were imputed using a regression based on non-missing data collected from other individuals in the same group, with dependent variable change from baseline and independent variables as specified in Table 4 above, with the exception of treatment, which not included since it was used to define each group, and region, which was defined as North American vs not North American. The regression results, with associated covariances, were used to simulate 1000 complete datasets, and each of these datasets was analyzed using the primary ANCOVA analysis, with the 1000 analysis results combined according to Rubin's rule.

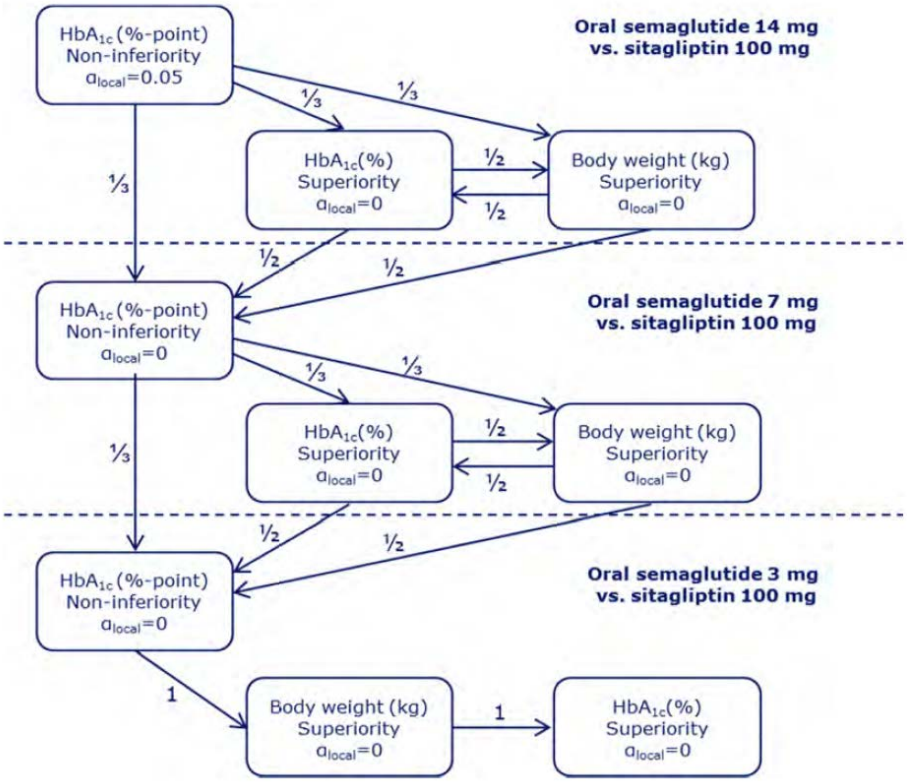
To evaluate the sensitivity of statistically significant results to the missing-at-random pattern-mixture imputations described immediately above, tipping point analysis were conducted for week 26 change from baseline percent HbA1c and change from baseline body weight. The tipping point analyses imposed penalties which reduced the treatment effect in imputed data, and the penalties were increased until the primary analyses combining results from the 1000 simulated datasets 'tipped,' i.e. failed to reject the null hypothesis. The applicant then discussed the plausibility of imposed penalties at these tipping points.

As implied by the treatment policy estimand, efficacy was analyzed using the full analysis set (FAS), comprised of all randomized patients regardless of initiation of correct treatment, rescue, dropout, or protocol violations.

To control the probability of overall type 1 error at the two-sided .05 level of significance, weighted Bonferroni-based closed testing procedures were used. In all trials evaluating non-inferiority hypotheses, upper 95% confidence limits were compared to a 0.4 margin. For trials with multiple doses of semaglutide, results from the 14 mg dose were tested first, followed by results from the 7 mg and 3 mg doses (Figure 1, Figure 4). For single dose study 4223, after testing HbA1c for non-inferiority, alpha was allocated equally to tests on HbA1c and body weight (Figure 2). In study 4224, the primary and secondary endpoints were tested sequentially against placebo and then tested against liraglutide (Figure 3).

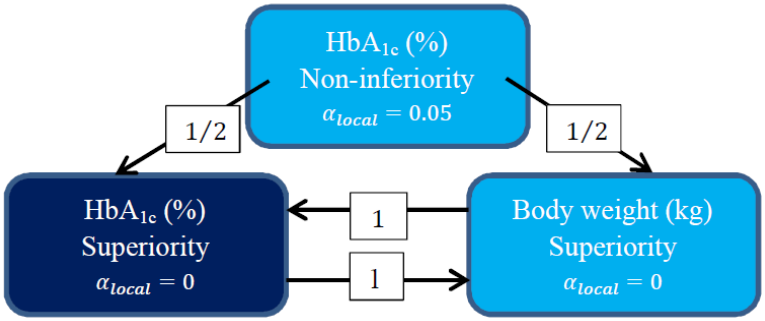
For study 4234, the superiority test comparing S14 to placebo for change from baseline HbA1c was followed by the superiority test for change from baseline body weight, each tested at the .05 level of significance.

Figure 1. Control of Type 1 Error, Trial 4222



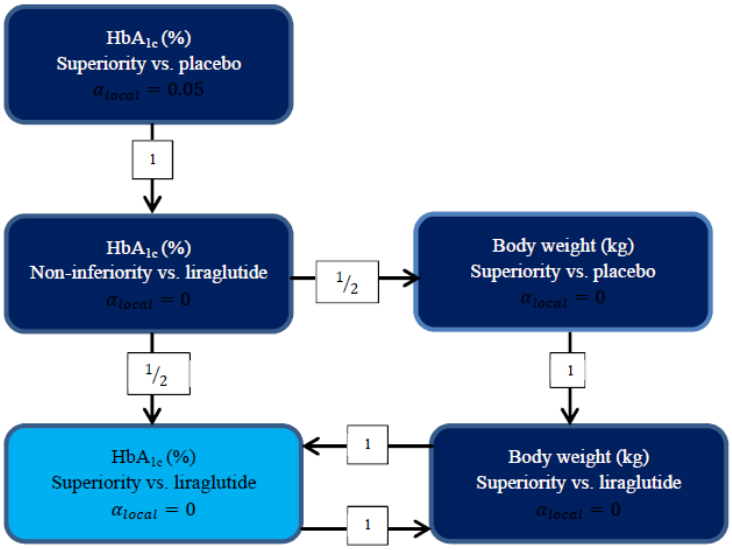
source: CSR studies P3

Figure 2. Control of Type 1 Error, Trial 4223



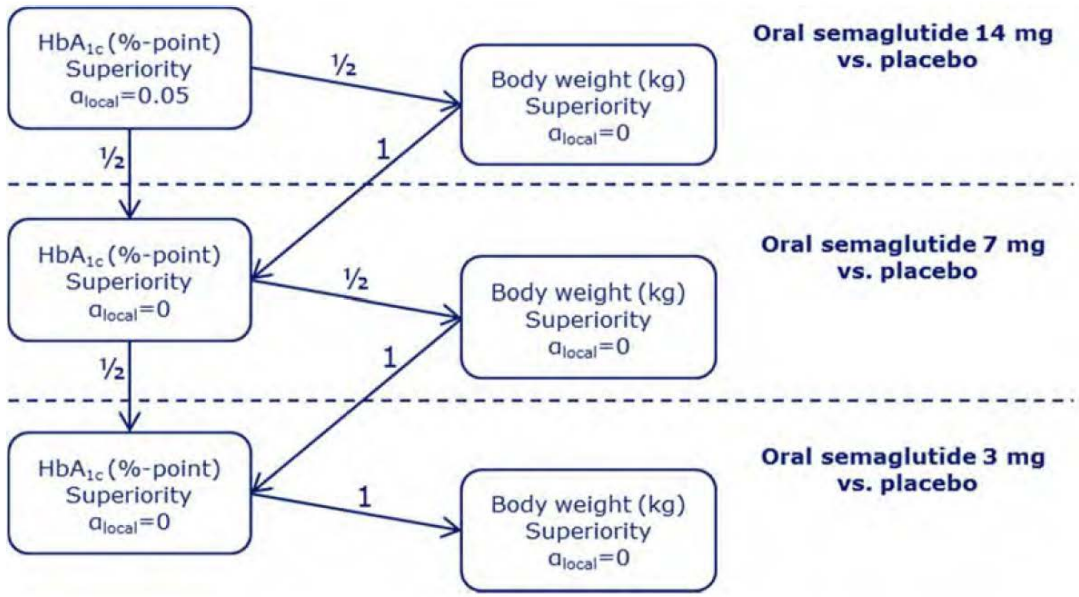
source: Figure 2-1 of SAP

Figure 3. Control of Type 1 Error, Study 4224



source: Figure 17-1 of protocol

Figure 4. Control of Type 1 Error, Trials 4233 and 4280



source: CSR studies P1 and P8

3.2.3 Patient Disposition, Demographic and Baseline Characteristics

The number and percent of patients with missing HbA1c at week 26 was consistently less than 10% (Table 5, Table 6, and Table 7)[1]. In all studies, semaglutide exhibited a numerically monotonic dose-response trend for treatment discontinuation due to adverse events.

Table 5. Patient Disposition at Week 26, Placebo Controlled Studies 4233 and 4224

Trial	**4233**				**4224**		
	S14	**S7**	**S3**	**Pbo**	**S14**	**Lira**	**Pbo**
Randomized (FAS)	175 (100)	175 (100)	175 (100)	178 (100)	285 (100)	284 (100)	142 (100)
Discontinued Treatment	24 (13.7)	18 (10.3)	12 (6.9)	19 (10.7)	38 (13.3)	28 (9.9)	14 (9.9)
Adverse Event	13 (7.4)	7 (4)	4 (2.3)	4 (2.2)	30 (10.5)	21 (7.4)	6 (4.2)
Protocol Violation	0 (0)	1 (0.6)	2 (1.1)	0 (0)	1 (0.4)	0 (0)	0 (0)
Study Withdrawal	5 (2.9)	5 (2.9)	0 (0)	3 (1.7)	2 (0.7)	3 (1.1)	2 (1.4)
Other	6 (3.4)	5 (2.9)	6 (3.4)	12 (6.7)	5 (1.8)	4 (1.4)	6 (4.2)
Rescue Medication	2 (1.1)	4 (2.3)	13 (7.4)	27 (15.2)	10 (3.5)	9 (3.2)	11 (7.7)
Death	1 (0.6)	0 (0)	0 (0)	0 (0)	2 (0.7)	2 (0.7)	1 (0.7)
HbA1c Measured	160 (91.4)	160 (91.4)	167 (95.4)	168 (94.4)	278 (97.5)	272 (95.8)	134 (94.4)
HbA1c Not Measured	15 (8.6)	15 (8.6)	8 (4.6)	10 (5.6)	7 (2.5)	12 (4.2)	8 (5.6)
Retrieved Data	11 (6.3)	8 (4.6)	16 (9.1)	34 (19.1)	41 (14.4)	26 (9.2)	17 (12)

source: disposit.sas

[1] Disposition given as number of patients and, in parentheses, percent of full analysis set

Table 6. Patient Disposition at Week 26, Active Controlled Studies 4222 and 4223

Trial	4222				4223	
	S14	S7	S3	Sita	S14	Empa
Randomized (FAS)	465 (100)	465 (100)	466 (100)	467 (100)	411 (100)	410 (100)
Discontinued Treatment	58 (12.5)	44 (9.5)	48 (10.3)	32 (6.9)	53 (12.9)	28 (6.8)
Adverse Event	39 (8.4)	20 (4.3)	15 (3.2)	13 (2.8)	36 (8.8)	10 (2.4)
Protocol Violation	3 (0.6)	5 (1.1)	4 (0.9)	4 (0.9)	0 (0)	3 (0.7)
Pregnancy	0 (0)	0 (0)	1 (0.2)	0 (0)	0 (0)	0 (0)
Study Withdrawal	4 (0.9)	3 (0.6)	11 (2.4)	1 (0.2)	5 (1.2)	6 (1.5)
Other	12 (2.6)	16 (3.4)	17 (3.6)	14 (3)	12 (2.9)	9 (2.2)
Rescue Medication	5 (1.1)	11 (2.4)	25 (5.4)	13 (2.8)	8 (1.9)	5 (1.2)
Death	0 (0)	2 (0.4)	1 (0.2)	0 (0)	0 (0)	0 (0)
HbA1c Measured	436 (93.8)	438 (94.2)	435 (93.3)	446 (95.5)	392 (95.4)	395 (96.3)
HbA1c Not Measured	29 (6.2)	27 (5.8)	31 (6.7)	21 (4.5)	19 (4.6)	15 (3.7)
Retrieved Data	38 (8.2)	31 (6.7)	43 (9.2)	25 (5.4)	44 (10.7)	18 (4.4)

source: disposit.sas

Table 7. Patient Disposition at Week 26, Placebo Controlled Studies 4234 and 4280

Trial	**4234**		**4280**			
	S14	**Pbo**	**S14**	**S7**	**S3**	**Pbo**
Randomized (FAS)	163 (100)	161 (100)	181 (100)	182 (100)	184 (100)	184 (100)
Discontinued Treatment	29 (17.8)	20 (12.4)	31 (17.1)	24 (13.2)	15 (8.2)	13 (7.1)
Adverse Event	24 (14.7)	10 (6.2)	22 (12.2)	12 (6.6)	9 (4.9)	3 (1.6)
Pregnancy	0 (0)	0 (0)	2 (1.1)	5 (2.7)	2 (1.1)	2 (1.1)
Protocol Violation	1 (0.6)	3 (1.9)	0 (0)	1 (0.5)	0 (0)	0 (0)
Study Withdrawal	0 (0)	2 (1.2)	2 (1.1)	2 (1.1)	0 (0)	2 (1.1)
Other	4 (2.5)	5 (3.1)	5 (2.8)	4 (2.2)	4 (2.2)	6 (3.3)
Rescue Medication	7 (4.3)	16 (9.9)	4 (2.2)	2 (1.1)	5 (2.7)	9 (4.9)
Death	1 (0.6)	2 (1.2)	2 (1.1)	0 (0)	0 (0)	0 (0)
HbA1c Measured	154 (94.5)	155 (96.3)	173 (95.6)	174 (95.6)	176 (95.7)	176 (95.7)
HbA1c Not Measured	9 (5.5)	6 (3.7)	8 (4.4)	8 (4.4)	8 (4.3)	8 (4.3)
Retrieved Data	28 (17.2)	28 (17.4)	27 (14.9)	18 (9.9)	14 (7.6)	13 (7.1)

source: disposit.sas

There were no obvious differences between treatments for baseline characteristics in the submitted studies (Table 8, Table 9, Table 10).

Table 8. Patient Demographics, Placebo Controlled Studies 4233 and 4224, Full Analysis Set

Trial			**4233**				**4224**		
			S14	**S7**	**S3**	**Pbo**	**S14**	**Lira**	**Pbo**
Randomized (FAS)			175 (100)	175 (100)	175 (100)	178 (100)	285 (100)	284 (100)	142 (100)
Sex Female		**N (%)**	89 (50.9)	82 (46.9)	86 (49.1)	89 (50)	138 (48.4)	135 (47.5)	68 (47.9)
Age	**mean**		54	55.5	54.9	53.8	56.2	56.4	56.5
	range		22-78	31-79	23-84	28-79	27-81	35-78	29-83
	< 65	**N (%)**	144 (82.3)	135 (77.1)	136 (77.7)	145 (81.5)	232 (81.4)	220 (77.5)	109 (76.8)
	≥ 65	**N (%)**	31 (17.7)	40 (22.9)	39 (22.3)	33 (18.5)	53 (18.6)	64 (22.5)	33 (23.2)
Race	**White**	**N (%)**	130 (74.3)	131 (74.9)	135 (77.1)	132 (74.2)	208 (73)	212 (74.6)	99 (69.7)
	Asian		29 (16.6)	30 (17.1)	31 (17.7)	31 (17.4)	39 (13.7)	36 (12.7)	19 (13.4)
	African American		10 (5.7)	11 (6.3)	6 (3.4)	10 (5.6)	12 (4.2)	9 (3.2)	8 (5.6)
	Native American		1 (0.6)	1 (0.6)	1 (0.6)	1 (0.6)	0 (0)	1 (0.4)	1 (0.7)
	Pacific Islander		1 (0.6)	0 (0)	0 (0)	0 (0)	0 (0)	1 (0.4)	0 (0)
	Other		4 (2.3)	2 (1.1)	2 (1.1)	4 (2.2)	23 (8.1)	17 (6.0)	12 (8.5)
Ethnicity	**Hispanic**	**N (%)**	46 (26.3)	31 (17.7)	52 (29.7)	51 (28.7)	3 (1.1)	8 (2.8)	3 (2.1)
	Not Hispanic		122 (69.7)	133 (76)	116 (66.3)	121 (68)	17 (6)	18 (6.3)	5 (3.5)
	N/A		7 (4)	11 (6.3)	7 (4)	6 (3.4)	268 (94)	266 (93.7)	137 (96.5)
Country	**USA**	**N (%)**	47 (26.9)	53 (30.3)	53 (30.3)	60 (33.7)	75 (26.3)	76 (26.8)	28 (19.7)

source: disposit.sas

Table 9. Patient Demographics, Active Controlled Studies 4222 and 4223, Full Analysis Set

Trial			4222				4223	
			S14	S7	S3	Sita	S14	Emp
Randomized (FAS)			465 (100)	465 (100)	466 (100)	467 (100)	411 (100)	410 (100)
Sex Female		N (%)	218 (46.9)	220 (47.3)	212 (45.5)	229 (49)	205 (49.9)	201 (49)
Age	mean		57.3	57.7	58.2	58.2	57.4	57.8
	range		21-81	26-79	27-84	18-84	30-84	27-82
	< 65	N (%)	342 (73.5)	335 (72)	339 (72.7)	346 (74.1)	306 (74.5)	300 (73.2)
	≥ 65	N (%)	123 (26.5)	130 (28)	127 (27.3)	121 (25.9)	105 (25.5)	110 (26.8)
Race	White	N (%)	317 (68.2)	330 (71)	344 (73.8)	333 (71.3)	355 (86.4)	353 (86.1)
	Asian		61 (13.1)	69 (14.8)	56 (12)	59 (12.6)	28 (6.8)	21 (5.1)
	African American		45 (9.7)	38 (8.2)	38 (8.2)	39 (8.4)	26 (6.3)	33 (8)
	Native American		5 (1.1)	3 (0.6)	4 (0.9)	6 (1.3)	0 (0)	0 (0)
	Pacific Islander		0 (0)	0 (0)	1 (0.2)	0 (0)	0 (0)	0 (0)
	N/A		17 (3.7)	14 (3)	10 (2.1)	18 (3.9)	0 (0)	0 (0)
	Other		20 (4.3)	11 (2.4)	13 (2.8)	12 (2.6)	2 (0.5)	3 (0.7)
Ethnicity	Hispanic	N (%)	75 (16.1)	77 (16.6)	76 (16.3)	93 (19.9)	91 (22.1)	108 (26.3)
	Not Hispanic		377 (81.1)	378 (81.3)	385 (82.6)	366 (78.4)	320 (77.9)	302 (73.7)
	N/A		13 (2.8)	10 (2.2)	5 (1.1)	8 (1.7)	0 (0)	0 (0)
Country	USA	N (%)	134 (28.8)	134 (28.8)	139 (29.8)	131 (28.1)	115 (28)	127 (31)

source: disposit.sas

Table 10. Patient Demographics, Placebo Controlled Studies 4234 and 4280, Full Analysis Set

Trial			4234		4280			
			S14	Pbo	S14	S7	S3	Pbo
Randomized (FAS)			163 (100)	161 (100)	181 (100)	182 (100)	184 (100)	181 (100)
Sex Female		N (%)	80 (49.1)	88 (54.7)	96 (53)	79 (43.4)	82 (44.6)	96 (53)
Age	mean		70.9	69.8	61.1	60.5	60.8	61.1
	range		49-92	45-86	35-85	26-82	27-82	35-85
	< 65	N (%)	26 (16)	39 (24.2)	108 (59.7)	121 (66.5)	110 (59.8)	108 (59.7)
	≥ 65	N (%)	137 (84)	122 (75.8)	73 (40.3)	61 (33.5)	74 (40.2)	73 (40.3)
Race	White	N (%)	158 (96.9)	152 (94.4)	94 (51.9)	95 (52.2)	89 (48.4)	94 (51.9)
	Asian		1 (0.6)	0 (0)	66 (36.5)	66 (36.3)	66 (35.9)	66 (36.5)
	African American		4 (2.5)	9 (5.6)	11 (6.1)	10 (5.5)	15 (8.2)	11 (6.1)
	Native American		0 (0)	0 (0)	1 (0.6)	0 (0)	1 (0.5)	1 (0.6)
	N/A		0 (0)	0 (0)	8 (4.4)	10 (5.5)	12 (6.5)	8 (4.4)
	Other		0 (0)	0 (0)	1 (0.6)	1 (0.5)	1 (0.5)	1 (0.6)
Ethnicity	Hispanic	N (%)	7 (4.3)	14 (8.7)	30 (16.6)	24 (13.2)	18 (9.8)	30 (16.6)
	Not Hispanic		156 (95.7)	147 (91.3)	143 (79)	148 (81.3)	154 (83.7)	143 (79)
	N/A		0 (0)	0 (0)	8 (4.4)	10 (5.5)	12 (6.5)	8 (4.4)
Country	USA	N (%)	43 (26.4)	51 (31.7)	52 (28.7)	50 (27.5)	58 (31.5)	52 (28.7)

source: disposit.sas

3.2.4 Results and Conclusions

3.2.4.1 *Primary Endpoint: Week 26 Change from Baseline Percent HbA1c*

Differences in HbA1c change between all semaglutide doses and placebo were statistically significant in all six of the studies evaluated (Table 11). In the active controlled trials, for HbA1c change, S14, S7, and S3 were significantly superior to sitagliptin, S14 was significantly superior to empagliflozin, and S14 was not significantly superior to liraglutide (Table 12). Because the upper 95% confidence limits were less than 0.4%, all doses were considered 'non-inferior' to active controls (Table 12). Tipping point analyses indicated that statistical significance was not sensitive to variation in plausible assumptions for the imputation of missing data.

Numerically monotonic trends in dose-response were consistently seen in trials with multiple doses of semaglutide.

Table 11. Change from Baseline Percent HbA1c Compared to Placebo, Week 26

Trial	Trt	N	Baseline	Δ Baseline	Difference from Placebo		
					LSMean	95% CI	P-value
4233	S3	175	7.9	-0.89	-0.59	(-0.82, -0.36)	<.0001
	S7	175	8.0	-1.16	-0.86	(-1.11, -0.62)	<.0001
	S14	175	8.0	-1.40	-1.10	(-1.34, -0.87)	<.0001
	Pbo	178	7.9	-0.30			
4224	S14	285	8.0	-1.23	-1.06	(-1.23, -0.89)	<.0001
	Pbo	142	7.9	-0.17			
4234	S14	163	8.0	-1.02	-0.84	(-1.05, -0.63)	<.0001
	Pbo	161	7.9	-0.18			
4280	S3	184	8.2	-0.57	-0.50	(-0.72, -0.29)	<.0001
	S7	182	8.2	-0.94	-0.88	(-1.10, -0.66)	<.0001
	S14	181	8.2	-1.25	-1.19	(-1.40, -0.97)	<.0001
	Pbo	184	8.2	-0.06			

source: hba1c 4222pa.sas, hba1c 4223pa.sas, hba1c 4224pa.sas, hba1c 4233pa.sas, hba1c 4234pa.sas, hba1c 4280pa.sas, compile primary.sas

Table 12. Change from Baseline Percent HbA1c Compared to Active Controls, Week 26

Trial	Trt	N	Baseline	Δ Baseline	Difference from Active Control		
					LSMean	95% CI	P-value
4224	S14	285	8.0	-1.23	-0.13	(-0.27, 0.01)	.065
	Lira 1.8	284	8.0	-1.10			
4222	S3	466	8.3	-0.60	0.17	(0.04, 0.29)	0.008
	S7	465	8.4	-1.04	-0.27	(-0.39, -0.14)	<.0001
	S14	465	8.3	-1.29	-0.52	(-0.64, -0.39)	<.0001
	Sita 100	467	8.3	-0.77			
4223	S14	411	8.1	-1.31	-0.44	(-0.57, -0.31)	<.0001
	Empa 25	410	8.1	-0.86			

source: hba1c 4222pa.sas, hba1c 4223pa.sas, hba1c 4224pa.sas, compile primary.sas

In summary, with the single exception of S14 versus liraglutide in study 4224, both S14 and S7 were superior to placebo and all active controls (Table 11 and Table 12). With regard to liraglutide, the upper 95% confidence limit for the difference from S14 was less than 0.4, and therefore S14 may be considered 'non-inferior' to liraglutide.

3.2.4.2 *Key Secondary Endpoint: Week 26 Change from Baseline Body Weight*

The placebo-controlled trials demonstrated a statistically significant impact of S14 on week 26 change from body weight (Table 13). The impacts of S3 and S7 on change from baseline body weight were statistically significant in trial 4280 (insulin + OAD background) but not in trial 4233 (semaglutide alone). Tipping point analyses indicated statistical significance was not sensitive to variation in plausible assumptions for the imputation of missing data.

In the active-controlled trials, semaglutide had statistically significant effect on body weight compared to sitagliptin and liraglutide (Table 14), but not compared to empagliflozin.

Numerically monotonic dose-response trends were consistently seen in trials with multiple doses of semaglutide (Table 13 and Table 14).

Table 13. Change from Baseline Body Weight Compared to Placebo, Week 26. (Measurement is in kilograms)

Trial	Trt	N	Baseline	Δ Baseline	Difference from Placebo		
					LSMean	95% CI	P-value
4233	S3	175	86.9	-1.46	-0.07	(-0.90, 0.76)	.9
	S7	175	89.0	-2.27	-0.88	(-1.89, 0.13)	.086
	S14	175	88.1	-3.68	-2.29	(-3.14, -1.45)	<.0001
	Pbo	178	88.6	-1.39			
4224	S14	285	92.9	-4.36	-3.84	(-4.66, -3.03)	<.0001
	Pbo	142	93.2	-0.52			
4234	S14	163	91.3	-3.43	-2.50	(-3.24, -1.76)	<.0001
	Pbo	161	90.4	-0.93			
4280	S3	184	85.9	-1.37	-0.95	(-1.84, -0.05)	.04
	S7	182	87.1	-2.42	-1.99	(-3.00, -0.97)	.0001
	S14	181	84.6	-3.68	-3.26	(-4.17, -2.35)	<.0001
	Pbo	184	86.0	-0.43			

source: bw 4222pa.sas, bw 4223pa.sas, bw 4224pa.sas, bw 4233pa.sas, bw 4234pa.sas, bw 4280pa.sas, compile primary.sas

Table 14. Change from Baseline Body Weight Compared to Active Controls, Week 26. (Measurement is in kilograms)

Trial	Trt	N	Baseline	Δ Baseline	Difference from Active Control		
					LSMean	95% CI	P-value
4224	S14	285	92.9	-4.36	-1.22	(-1.88, -0.56)	.0003
	Lira 1.8	284	95.5	-3.15			
4222	S3	466	91.6	-1.17	-0.57	(-1.05, -0.10)	.02
	S7	465	91.3	-2.16	-1.57	(-2.05, -1.10)	<.0001
	S14	465	91.2	-3.09	-2.50	(-2.98, -2.03)	<.0001
	Sita 100	467	90.9	-0.59			
4223	S14	411	92.0	-3.83	-0.09	(-0.66, 0.48)	.76
	Empa 25	410	91.3	-3.74			

source: bw 4222pa.sas, bw 4223pa.sas, bw 4224pa.sas, compile secondary.sas

3.2.4.3 *Supportive Endpoint: Week 26 Percent HbA1c Less Than 7%*

Compared to placebo, raw proportion of patients with HbA1c < 7% at week 26 was generally higher among patients treated with semaglutide. Similarly, compared to active control, S14 and S7 exhibited numerically higher raw proportion of patients with HbA1c < 7% at week 26.

Numerically monotonic dose-response trends were consistently seen in trials with multiple doses of semaglutide.

Table 15. Raw Proportion of HbA1c < 7%, Week 26

Trial	Raw Proportion (n HbA1c<7% / N)						
	S14	**S7**	**S3**	**Pbo**	**Sita 100**	**Empa 25**	**Lira 1.8**
4233	0.77 (123/160)	0.69 (110/160)	0.55 (92/167)	0.31 (52/168)			
4224	0.68 (188/278)			0.14 (19/134)			0.62 (168/272)
4234	0.58 (89/154)			0.23 (35/155)			
4280	0.58 (101/173)	0.43 (74/174)	0.28 (50/176)	0.07 (12/176)			
4222	0.56 (246/436)	0.44 (192/438)	0.27 (116/435)		0.32 (144/446)		
4223	0.67 (262/392)					0.40 (158/395)	

source: HbA1c raw.sas

The proportion of patients with HbA1c < 7% at week 26 was not included in the test hierarchy; therefore true statistical significance of differences between treatment and control could not be assessed without increasing overall type 1 error. Nominal p-values and confidence intervals,[2] however, were calculated, and they informally support evidence for semaglutide benefit provided by the primary and secondary endpoints, with nominally significant differences between placebo and all doses of semaglutide (Table 16).

[2] True type 1 error is higher than indicated by 'nominal' p-values, and the true width of confidence intervals is wider than indicated by 'nominal' confidence intervals. For nominal inference, strength of evidence in favor of treatment is therefore weaker - p-values and associated confidence intervals should be interpreted with great caution and should not be used as the main justification for regulatory decisions.

Table 16. Adjusted Proportion of HbA1c < 7% Compared to Placebo, Week 26

Trial	Trt	N	Proportion[1]	Difference from Placebo			
				Proportion	OR	OR 95% CI[2]	p-value[2]
4233	S3	175	0.540	0.265	3.09	(1.91, 4.99)	<.0001
	S7	175	0.760	0.485	8.36	(4.86, 14.41)	<.0001
	S14	175	0.687	0.412	5.79	(3.50, 9.59)	<.0001
	Pbo	178	0.275				
4224	S14	285	0.684	0.572	17.1	(9.50, 30.77)	<.0001
	Pbo	142	0.113				
4234	S14	163	0.576	0.378	5.5	(3.20, 9.44)	<.0001
	Pbo	161	0.198				
4280	S3	184	0.250	0.194	5.61	(2.77, 11.37)	<.0001
	S7	182	0.424	0.368	12.37	(6.12, 25.00)	<.0001
	S14	181	0.573	0.517	22.52	(11.14, 45.51)	<.0001
	Pbo	184	0.056				

source: hba1cr 4222.sas, hba1cr 4223.sas, hba1cr 4224.sas, hba1cr 4233.sas, hba1cr 4234.sas, hba1cr 4280.sas, compile secondary.sas
1 back-transformed from modeled odds-ratio
2 p-values and confidence intervals nominal and incorrectly biased toward significance

Differences between S14 and liraglutide for percent HbA1c < 7% were not statistically significant (Table 17). In study 4222, in which patients were on a background of metformin or sulfonylurea with or without metformin, the difference between low dose semaglutide (S3) and sitagliptin also was not statistically significant. Nominally significant differences were seen between sitagliptin and S7 or S14, and between empagliflozin and S14.

Table 17. Adjusted Proportion of HbA1c < 7% Compared to Active Controls, Week 26

Trial	Trt	N	Proportion	Difference from Active Control			
				Proportion	OR	OR 95% CI[1]	p-value[1]
4224	S14	285	0.684	0.060	1.31	(0.91, 1.89)	.15
	Lira 1.8	284	0.624				
4222	S3	466	0.220	-0.055	0.74	(0.54, 1.02)	.065
	S7	465	0.428	0.153	1.97	(1.46, 2.66)	<.0001
	S14	465	0.551	0.276	3.23	(2.39, 4.37)	<.0001
	Sita 100	467	0.275				
4223	S14	411	0.670	0.296	3.39	(2.47, 4.65)	<.0001
	Empa 25	410	0.375				

source: hba1cr 4222.sas, hba1cr 4223.sas, hba1cr 4224vL.sas, compile secondary.sas
1 p-values and confidence intervals nominal and incorrectly biased toward significance

3.2.4.4 *Supportive Endpoint: Week 26 Change from Baseline Fasting Blood Glucose*

Week 26 change from baseline fasting blood glucose was not included in the test hierarchy; therefore only nominal p-values and confidence intervals could be calculated. With the exception of the difference between placebo and S3 in study 4280, the analyses informally support a benefit of semaglutide compared to placebo, with nominally significant differences between placebo and semaglutide (Table 18).

Compared to empagliflozin and liraglutide, S14 did not provide a nominally significant benefit for change from baseline fasting plasma glucose (Table 19). However, compared to sitagliptin, S7 and S14 did provide nominally significant benefits.

Numerically monotonic trends in dose-response were consistently seen in trials with multiple doses of semaglutide (Table 18, Table 19).

3.2.4.5 *Efficacy Summary*

Compared to placebo, the submitted trials clearly demonstrate that S7 and S14 reduce percent HbA1c and body mass. Supportive evidence also indicates that these dosages reduce fasting plasma glucose while increasing the percent of patients with HbA1c < 7%. Concerns regarding failed superiority of S14 to L1.8 for HbA1c endpoints in study 4224 (Table 12, Table 17) are mitigated by the fact that (i) the week 26 change from baseline for semaglutide in trial 4224 is similar to that in the other trials (Table 11 through Table 19) and (ii) that liraglutide is in the same product class as semaglutide.

Table 18. Change from Baseline Fasting Plasma Glucose Compared to Placebo, Week 26

Trial	Trt	N	Baseline	Δ Baseline	Difference from Placebo		
					LSMean	95% CI[1]	p-value[1]
4233	S3	175	158.3	-16.16	-12.92	(-21.36, -4.49)	.003
	S7	175	161.9	-27.88	-24.64	(-35.10, -14.18)	<.0001
	S14	175	158.1	-32.87	-29.63	(-38.26, -21.00)	<.0001
	Pbo	178	160.0	-3.23			
4224	S14	285	167.1	-36.05	-29.52	(-35.95, -23.09)	<.0001
	Pbo	142	166.7	-6.53			
4234	S14	163	163.6	-27.73	-21.12	(-30.55, -11.69)	<.0001
	Pbo	161	163.5	-6.61			
4280	S3	184	158.4	-4.03	-9.34	(-19.42, 0.74)	.069
	S7	182	153.3	-19.53	-24.84	(-34.82, -14.85)	<.0001
	S14	181	150.1	-24.20	-29.51	(-39.53, -19.49)	<.0001
	Pbo	184	149.5	5.31			

source: fpg 4222pa.sas, fpg 4223pa.sas, fpg 4224pa.sas, fpg 4233pa.sas, fpg 4234pa.sas, fpg 4280pa.sas, compile fpg.sas
1 p-values and confidence intervals nominal and incorrectly biased toward significance

Table 19. Change from Baseline Fasting Plasma Glucose Compared to Active Controls, Week 26

Trial	Trt	N	Baseline	Δ Baseline	Difference from Active Control		
					LSMean	95% CI[1]	p-value[1]
4224	S14	285	167.1	-36.05	-2.42	(-7.40, 2.57)	.34
	Lira 1.8	284	167.6	-33.64			
4222	S3	466	174.2	-13.55	1.87	(-3.55, 7.28)	.5
	S7	465	170.3	-21.28	-5.86	(-11.40, -0.31)	.039
	S14	465	167.9	-30.54	-15.12	(-20.58, -9.66)	<.0001
	Sita 100	467	171.8	-15.42			
4223	S14	411	171.5	-35.92	0.35	(-4.30, 5.00)	.9
	Empa 25	410	174.0	-36.28			

source: fpg 4222pa.sas, fpg 4223pa.sas, fpg 4224pa.sas, compile fpg.sas
1 p-values and confidence intervals nominal and incorrectly biased toward significance

3.3 Evaluation of Safety

Percent incidence of severe and clinically significant hypoglycemic adverse events did not show a clear trend with respect to dose of semaglutide (Table 20). To account for potential time-in-study differences between treatments, and to increase sample size for number of events, overall rates (including incidence and recurrence) were also examined, and, like percent incidence, did not show a clear trend with respect to dose of semaglutide (Table 21).

Perhaps due to the larger number of events, it could be seen that alert rates were numerically higher in S14 than placebo, and, except for study 4233, alert rates were numerically higher in S14 than in S7 or S3.

Table 20 and Table 21 include all patients continuing to receive randomized treatment. However, on Table 2 of the proposed product label, the applicant did not count patients on randomized treatment who used rescue medications, unnecessarily breaking the randomization while failing to account for influence of treatment on use of rescue medication. Incorporation into the product label of results from Table 20 is therefore recommended.

Table 20. Hypoglycemia Percent Incidence While on Randomized Treatment

Study	Class[1]	Percent Incidence (26 Weeks, 52 Weeks Only for Study 4280)						
		S14	S7	S3	Pbo	Sita	Empa	L1.8
4222	Alert	24.1% (112/465)	19.4% (90/464)	18.2% (85/466)		19.3% (90/466)		
	Clin Signif	9% (42/465)	7.1% (33/464)	6.4% (30/466)		9.4% (44/466)		
	Severe	0.4% (2/465)	0% (0/464)	0% (0/466)		0.9% (4/466)		
4223	Alert	8.5% (35/410)					8.3% (34/409)	
	Clin Signif	2.7% (11/410)					2% (8/409)	
	Severe	0.2% (1/410)					0.2% (1/409)	

source: safety hypoglyc perc inc.sas

Table 20. Hypoglycemia Percent Incidence While on Randomized Treatment (continued)

Study	Class[1]	Percent Incidence (26 Weeks, 52 Weeks Only for Study 4280)						
		S14	S7	S3	Pbo	Sita	Empa	L1.8
4224	Alert	7.7% (22/285)			4.2% (6/142)			10.2% (29/284)
	Clin Signif	0.7% (2/285)			1.4% (2/142)			3.5% (10/284)
	Severe	0% (0/285)			0% (0/142)			0% (0/284)
4233	Alert	4% (7/175)	5.1% (9/175)	5.1% (9/175)	0.6% (1/178)			
	Clin Signif	0% (0/175)	0% (0/175)	3.4% (6/175)	1.1% (2/178)			
	Severe	0% (0/175)	0.6% (1/175)	0% (0/175)	0% (0/178)			
4234	Alert	20.9% (34/163)			11.2% (18/161)			
	Clin Signif	5.5% (9/163)			3.1% (5/161)			
	Severe	0% (0/163)			0% (0/161)			
4280	Alert	53.6% (97/181)	52.5% (95/181)	52.7% (97/184)	53.3% (98/184)			
	Clin Signif	29.8% (54/181)	26% (47/181)	32.6% (60/184)	32.1% (59/184)			
	Severe	1.1% (2/181)	0.6% (1/181)	2.7% (5/184)	0.5% (1/184)			

source: safety hypoglyc perc inc.sas

1 ADA 2018 Hypoglycemic Class: Alert episodes are associated with plasma glucose ≤ 70 mg/dL and ≥ 54 mg/dL, clinically significant episodes are associated with plasma glucose < 54 mg/dL, and severe episodes require assistance of another person for recovery.

Table 21. Hypoglycemia Event Rate While on Randomized Treatment

Study	Class	Events per Treatment-Year (# Events / # Treatment-years)						
		S14	S7	S3	Pbo	Sita	Empa	L1.8
4222	Alert	0.769 (468/609)	0.432 (271/627)	0.377 (234/621)		0.439 (283/645)		
	Clin Signif	0.128 (78/609)	0.102 (64/627)	0.113 (70/621)		0.141 (91/645)		
	Severe	0.003 (2/609)	0 (0/627)	0 (0/621)		0.006 (4/645)		
4223	Alert	0.231 (84/364)					0.193 (74/384)	
	Clin Signif	0.036 (13/364)					0.023 (9/384)	
	Severe	0.003 (1/364)					0.003 (1/384)	
4224	Alert	0.173 (44/255)			0.092 (12/131)			0.227 (59/260)
	Clin Signif	0.008 (2/255)			0.015 (2/131)			0.042 (11/260)
	Severe	0 (0/255)			0 (0/131)			0 (0/260)
4233	Alert	0.098 (8/81)	0.121 (10/83)	0.140 (12/86)	0.012 (1/85)			
	Clin Signif	0 (0/81)	0 (0/83)	0.07 (6/86)	0.023 (2/85)			
	Severe	0 (0/81)	0.012 (1/83)	0 (0/86)	0 (0/85)			
4234	Alert	1.289 (96/74)			0.553 (42/76)			
	Clin Signif	0.242 (18/74)			0.066 (5/76)			
	Severe	0 (0/74)			0 (0/76)			
4280	Alert	3.911 (602/154)	3.538 (565/160)	3.668 (624/170)	3.148 (545/173)			
	Clin Signif	1.052 (162/154)	1.196 (191/160)	1.334 (227/170)	1.161 (201/173)			
	Severe	0.013 (2/154)	0.006 (1/160)	0.029 (5/170)	0.006 (1/173)			

source: safety hypoglyc event rate.sas

The submitted confirmatory trials did not address the potential impact of absorption enhancer SNAC on the pharmacodynamics of other orally administered medications. One possible approach is to evaluate the impact of SNAC on hypothyroidism patients since thyroxine, like semaglutide, is administered orally before breakfast. In the absence of laboratory data on TSH, T4, or T3, change from baseline LDL cholesterol was used as an indicator of administered thyroxine effectiveness. An ANCOVA was conducted, with dependent variable change from baseline LDL, and with independent factors hypothyroidism (patients with hypothyroidism on oral thyroxine medication versus patients who did not have hypothyroidism and were not on oral thyroxine medication), treatment [semaglutide (S3, S7, or S14) versus control], baseline LDL, and treatment by hypothyroidism interaction. A separate ANCOVA was performed for each study.

With caveats that data collection for this purpose was not optimal in these studies, that unrecorded compensatory adjustments in thyroxine dose may have obscured treatment effects, that semaglutide, as a GLP-1 receptor agonist, may itself reduce LDL, that variables directly reflecting measures of thyroid activity, such as TSH, T4, and T3, were not measured, and that the analyses may be inadequately powered, the results failed to establish an impact of SNAC on the absorption of concurrently administered oral thyroxine replacements. In particular, for each of the studies, among patients administered semaglutide, hypothyroidism had no significant effect on week 26 change from baseline LDL, indicating that semaglutide + SNAC did not impact potential effects on LDL of ongoing treatment with thyroxine. In contrast, for study 4223, a different diabetes treatment, empagliflozin, may have been associated with a greater increase in LDL cholesterol (15 mg/dL vs 3 mg/dL) among patients with hypothyroidism compared to patients without hypothyroidism (nominal p-value = .006).

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

Whether efficacy depended on subgroups such as race, gender, age class, or region was first examined by adding to the primary analysis model terms for the particular subgroup under examination as well as the interactions of that subgroup with each dependent factor (blue lines in Figure 5 through Figure 11).

In the traditional subgroup analyses, there were random highs and random lows in sample estimates of subgroup treatment effects due to small sample size and large variability for some subgroups. Therefore, we also derived shrinkage estimates of subgroup treatment effects using a Bayesian hierarchical model based on summary sample estimates (red lines in Figure 5 through Figure 11, example computer code provided in Appendix). The total variability in the sample estimates is the sum of the within subgroup variability of the sample estimator and the across subgroups variability in underlying/true parameter values. A shrinkage estimate of the subgroup treatment effect, which borrows information from the other subgroups while estimating the treatment effect for a specific subgroup, is a “weighted” average of the sample estimate and the overall estimate. The weights are based on the ratio of the between-subgroup variability to the within-subgroup variability. The greater the ratio is, the smaller the weight on the overall estimate

(the less the shrinkage). We used the same flat prior distribution to derive shrinkage estimates for all subgroups. The Bayesian hierarchical model assumptions are:

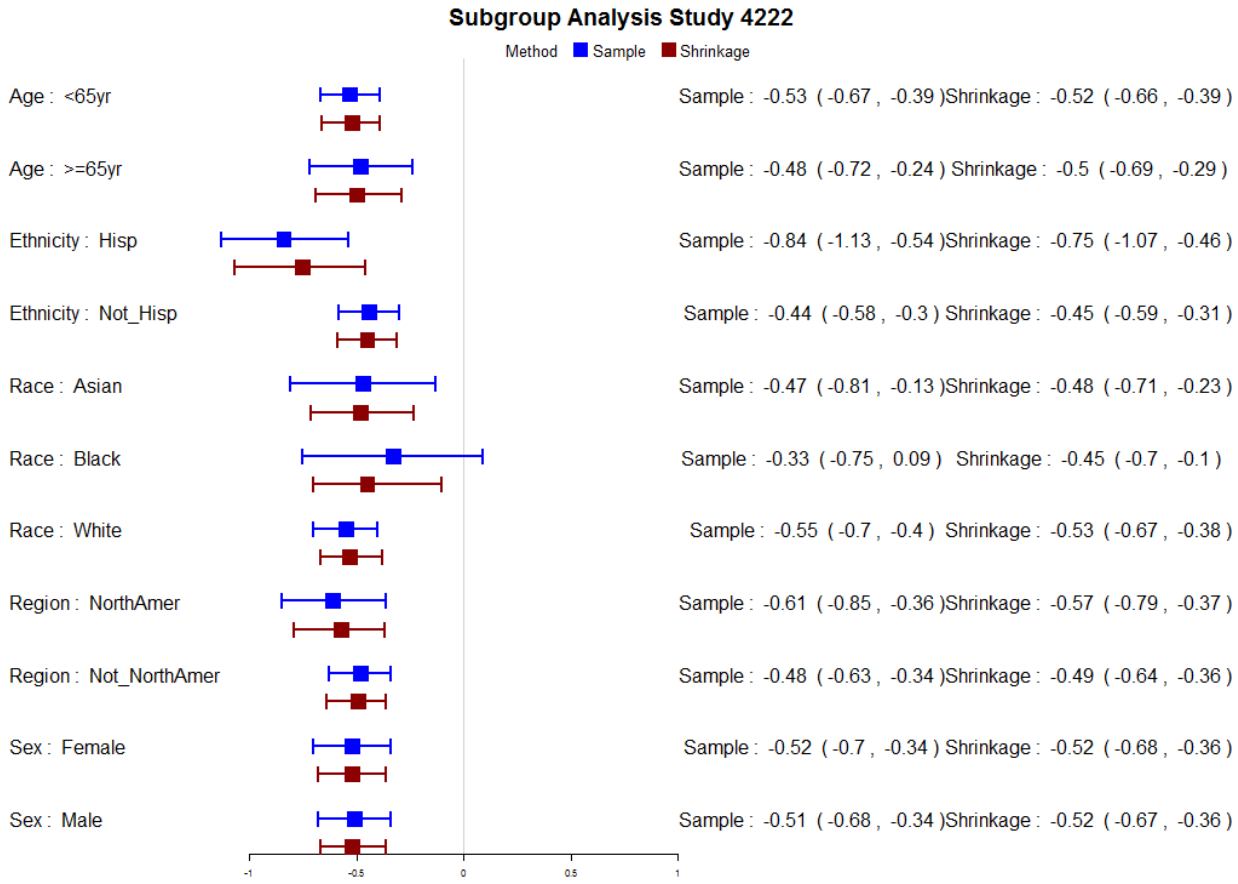
For $i = 1, 2\ldots$, where Y_i represents the observed sample estimate of treatment effect in a subgroup level i, assume $Y_i \sim N(\mu_i, \sigma_i^2)$ where

- σ_i^2 are the observed variance for sample estimates
- $\mu_i \sim N(\mu, \tau^2)$
- $\mu \sim N(0, 4)$, $1/\tau^2 \sim$ Gamma(0.001, 0.001)

In general, the treatment effect of semaglutide was consistent across the different subgroups. For change from baseline HbA1c, gender, age class (<65, ≥65), race, and region (North America vs not North America) had no qualitative impact on the effect of semaglutide 14 mg compared to placebo or active controls (Figure 5 through Figure 11, sample sizes on Table 22 and Table 23).

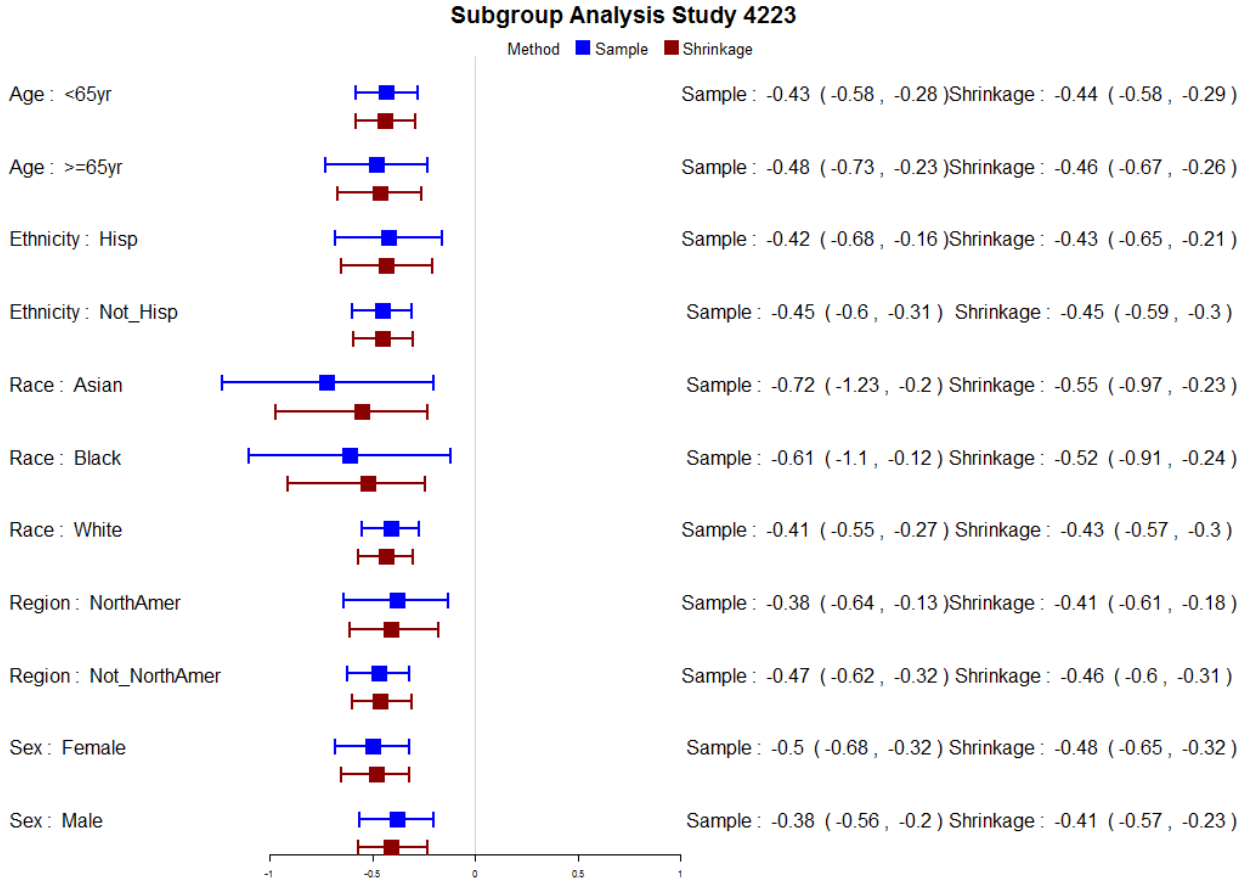
Study 4234 indicated that semaglutide 14 mg may be less effective in Hispanics than in non-Hispanics, with the point estimate indicating potential harm (Figure 9). Concerns regarding ethnicity are mitigated by the fact that (i) there were only seven Hispanics in study 4234, (ii) all of the other studies showed positive benefits regardless of ethnicity (Figure 5, Figure 6, Figure 7, Figure 8, and Figure 10), and (iii) inverse-variance weighted meta-analysis of placebo-controlled studies 4224, 4233, 4234, and 4280 showed positive effects of semaglutide regardless of ethnicity (Figure 11).

Figure 5. Subgroup Analyses, Semaglutide 14 mg vs Sitagliptin, Study 4222



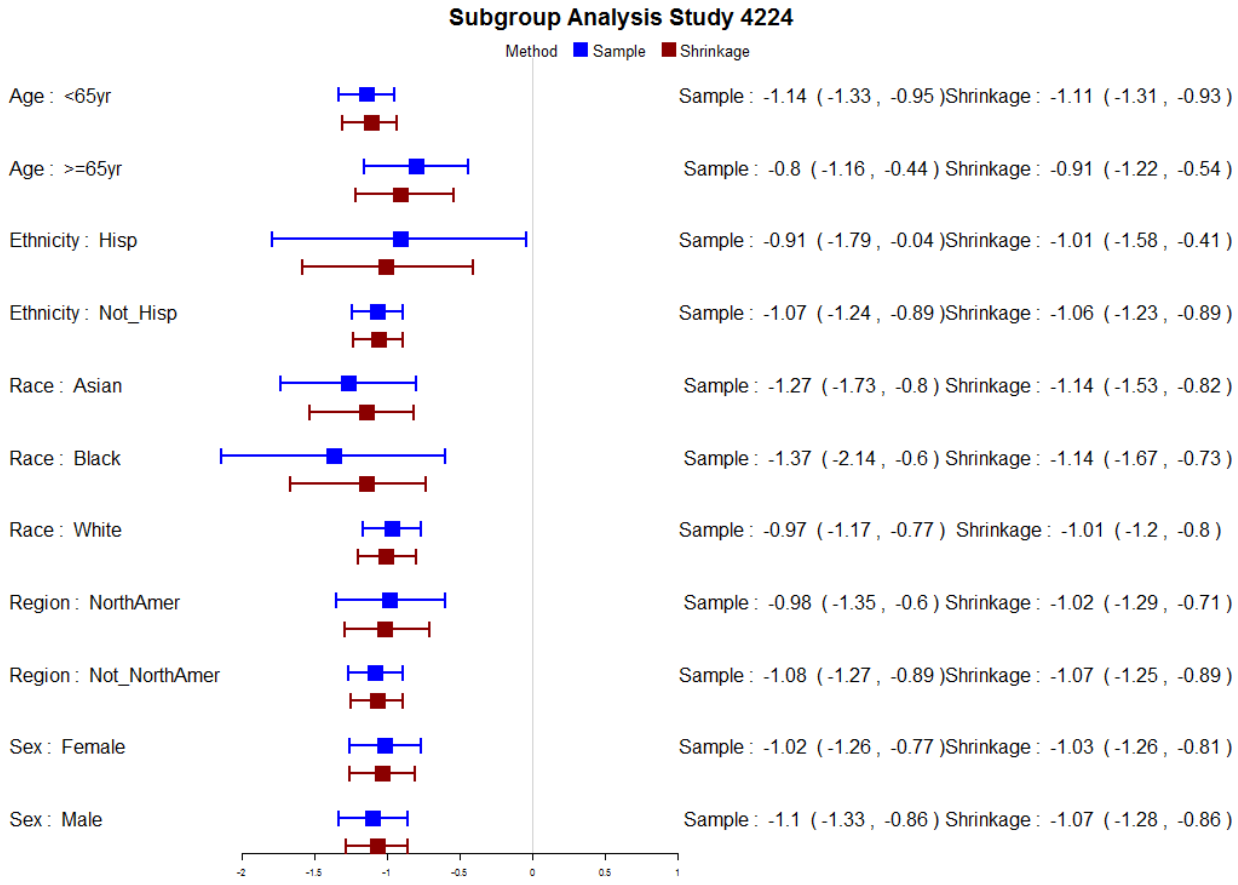
source: S14_4222 Subgroup Forest Plot.R, group for review 2019 06 27.sas, applicant code ana-a1c-4222-subgrp.sas

Figure 6. Subgroup Analyses, Semaglutide 14 mg vs Empagliflozin, Study 4223



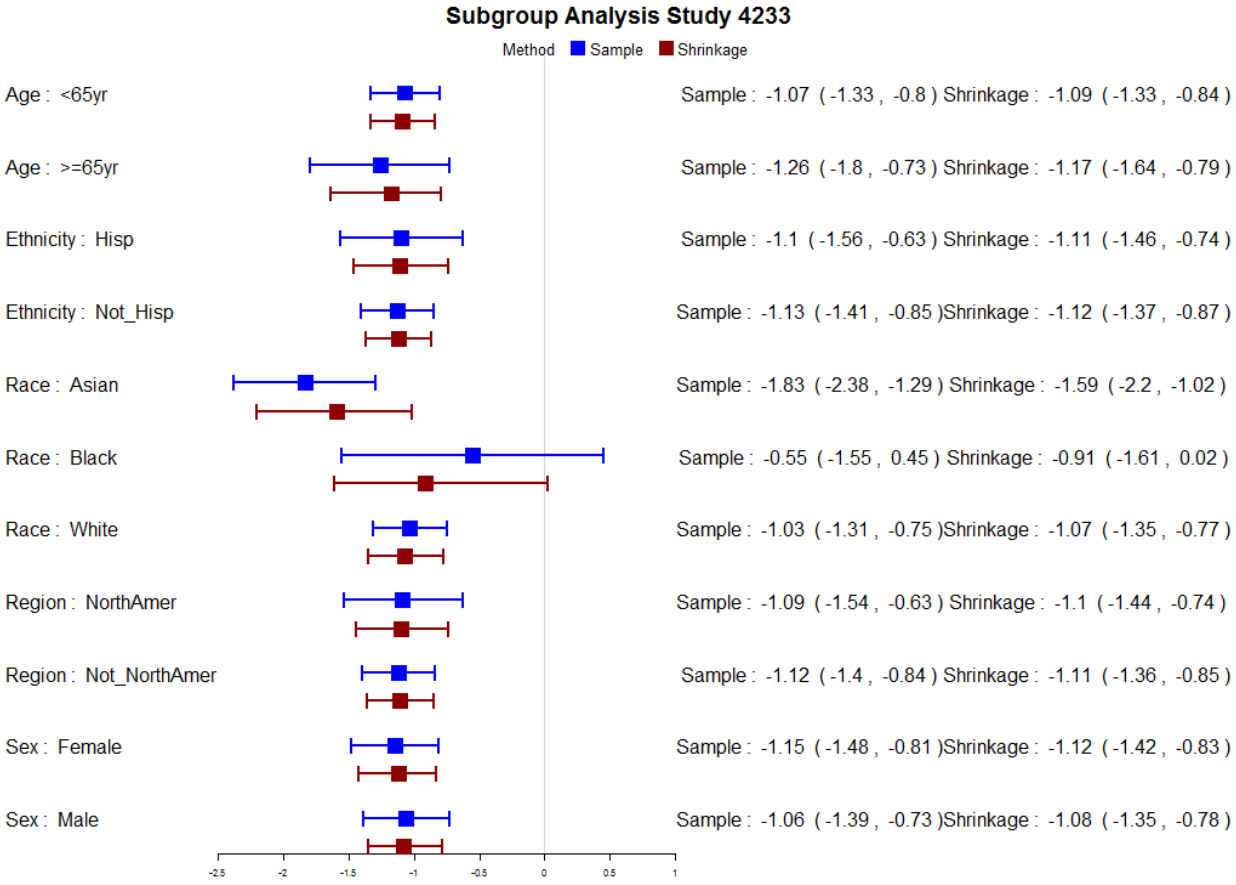
source: S14_4223 Subgroup Forest Plot.R, group for review 2019 06 27.sas, applicant code ana-a1c-4223-subgrp.sas

Figure 7. Subgroup Analyses, Semaglutide 14 mg vs Placebo, Study 4224



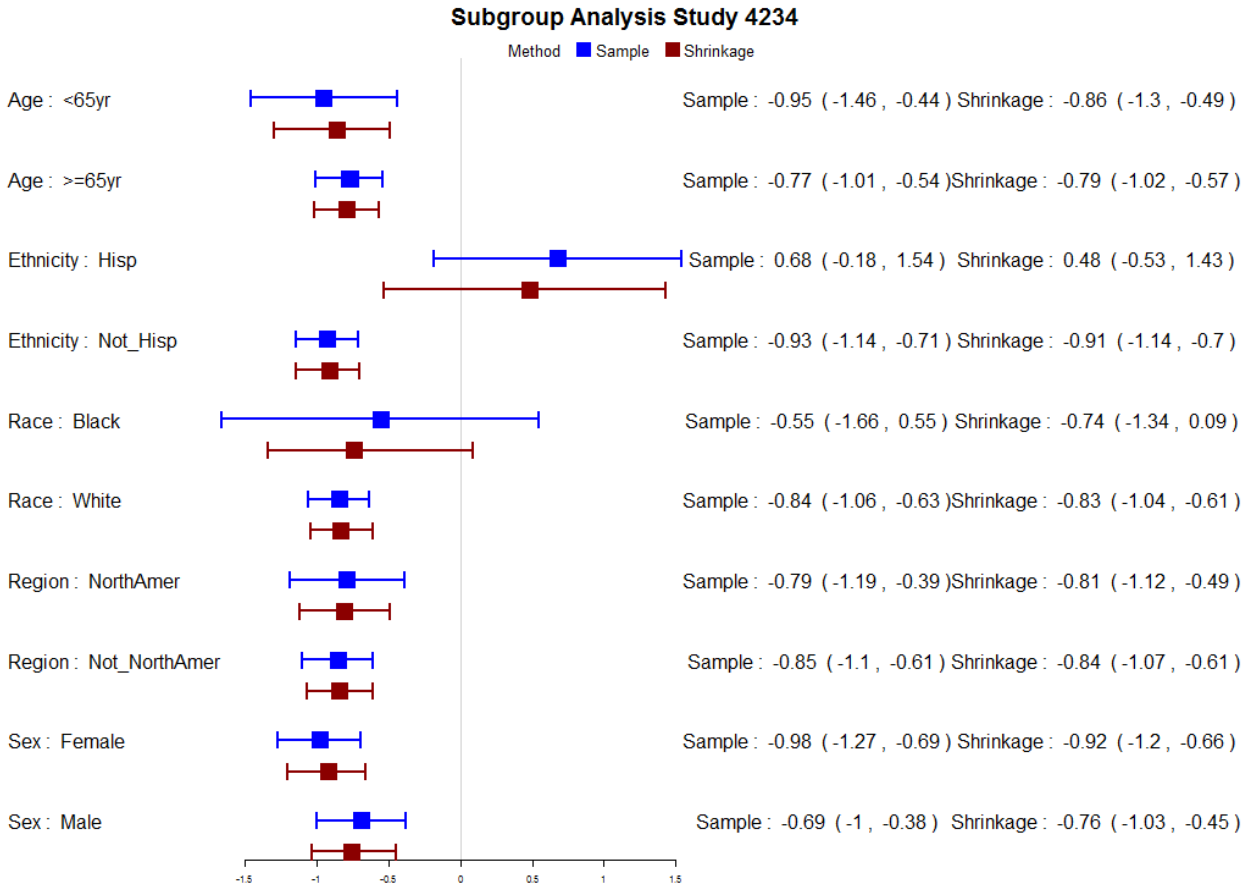
source: S14_4224 Subgroup Forest Plot.R, group for review 2019 06 27.sas, applicant code ana-a1c-4224-subgrp.sas

Figure 8. Subgroup Analyses, Semaglutide 14 mg vs Placebo, Study 4233



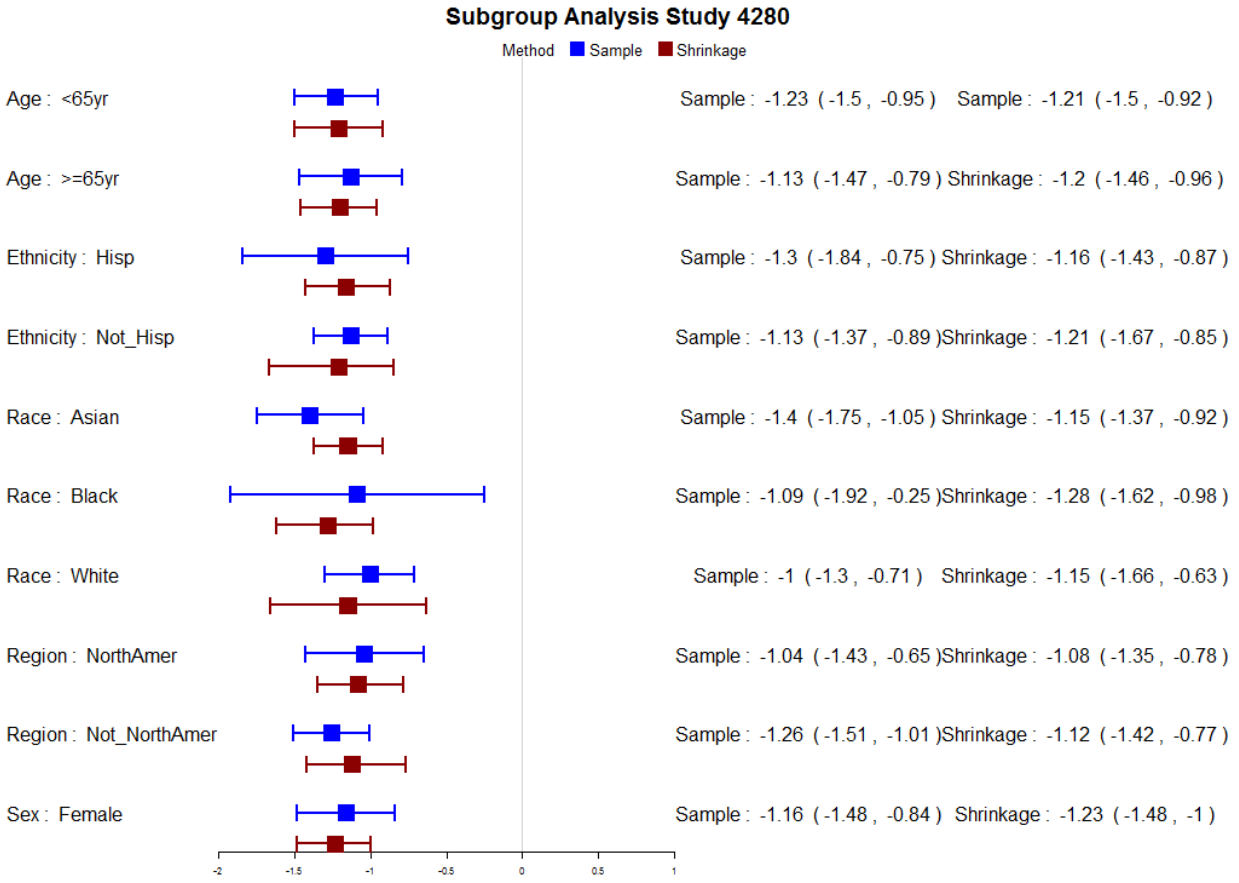
source: S14_4233 Subgroup Forest Plot.R, group for review 2019 06 27.sas, applicant code ana-a1c-4233-subgrp.sas

Figure 9. Subgroup Analyses, Semaglutide 14 mg vs Placebo, Study 4234



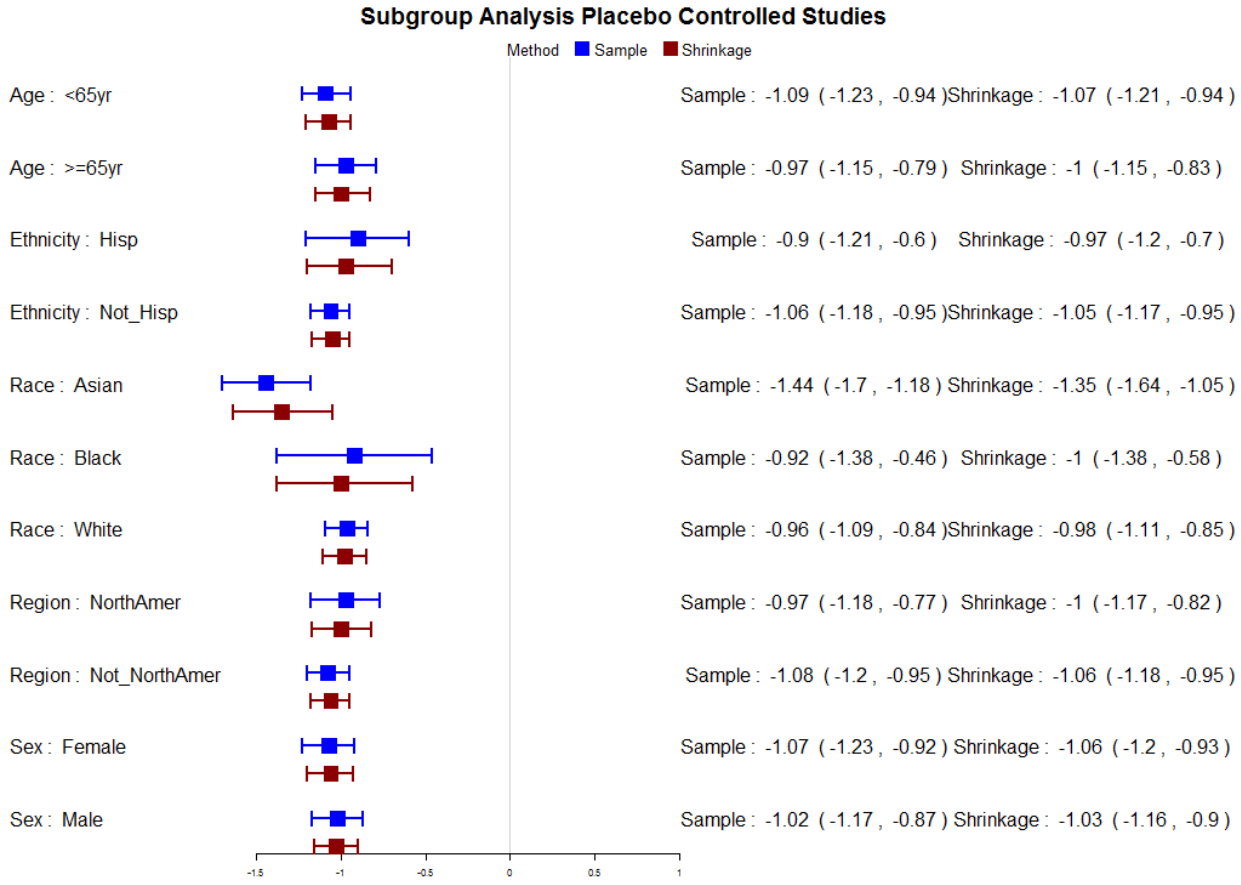
source: S14_4234 Subgroup Forest Plot.R, group for review 2019 06 27.sas, applicant code ana-a1c-4222-subgrp.sas

Figure 10. Subgroup Analyses, Semaglutide 14 mg vs Placebo, Study 4280



source: S14_4280 Subgroup Forest Plot.R, group for review 2019 06 27.sas, applicant code ana-a1c-4280-subgrp.sas

Figure 11. Subgroup Meta-analyses, Semaglutide 14 mg vs Placebo, Studies 4224, 4234, 4233, and 4280



source: S14_1845 Subgroup Forest Plot.R, group for review 2019 06 27.sas, applicant code ana-a1c-pbopool-subgrp.sas

Table 22. Sample Sizes for Subgroup Analyses on Change from Baseline HbA1c

Study	Group	Subgroup	S14	S7	S3	Pbo	Sita	Empa	Lira
S4222	Age	< 65yr	320	318	317		330		
		≥ 65yr	116	120	118		116		
	Ethnicity	Hisp	69	71	72		89		
		Not_Hisp	357	359	359		350		
	Race	Asian	58	66	56		58		
		Black	42	36	30		36		
		White	298	310	323		317		
	Region	NorthAmer	122	119	122		122		
		Not_NorthAmer	314	319	313		324		
	Sex	Female	202	209	192		222		
		Male	234	229	243		224		
S4223	Age	< 65yr	289					291	
		≥ 65yr	103					104	
	Ethnicity	Hisp	84					103	
		Not_Hisp	308					292	
	Race	Asian	27					21	
		Black	24					31	
		White	341					340	
	Region	NorthAmer	106					115	
		Not_NorthAmer	286					280	
	Sex	Female	196					194	
		Male	196					201	

Table 22. Sample Sizes for Subgroup Analyses on Change from Baseline HbA1c (continued)

Study	Group	Subgroup	S14	S7	S3	Pbo	Sita	Empa	Lira
S4224	Age	< 65 yr	226			103			210
		≥ 65 yr	52			31			62
	Ethnicity	Hisp	16			4			17
		Not_Hisp	262			130			255
	Race	Asian	39			17			35
		Black	11			7			8
		White	202			95			203
	Region	NorthAmer	74			25			70
		Not_NorthAmer	204			109			202
	Sex	Female	135			64			129
		Male	143			70			143
S4233	Age	< 65yr	131	123	129	136			
		≥ 65yr	29	37	38	32			
	Ethnicity	Hisp	38	25	48	46			
		Not_Hisp	116	124	112	117			
	Race	Asian	29	29	30	31			
		Black	9	9	6	9			
		White	117	121	128	123			
	Region	NorthAmer	40	44	48	54			
		Not_NorthAmer	120	116	119	114			
	Sex	Female	79	74	82	84			
		Male	81	86	85	84			

Table 22. Sample Sizes for Subgroup Analyses on Change from Baseline HbA1c (continued)

Study	Group	Subgroup	S14	S7	S3	Pbo	Sita	Empa	Lira
S4234	Age	< 65yr	22			38			
		≥ 65yr	132			117			
	Ethnicity	Hisp	7			12			
		Not_Hisp	147			143			
	Race	Asian	1						
		Black	4			9			
		White	149			146			
	Region	NorthAmer	39			49			
		Not_NorthAmer	115			106			
	Sex	Female	75			87			
		Male	79			68			
S4280	Age	< 65yr	104	115	102	108			
		≥ 65yr	69	59	74	68			
	Ethnicity	Hisp	29	21	18	24			
		Not_Hisp	136	143	146	144			
	Race	Asian	63	64	65	64			
		Black	10	10	11	13			
		White	90	90	86	92			
	Region	NorthAmer	56	59	60	54			
		Not_NorthAmer	117	115	116	122			
	Sex	Female	89	73	78	75			
		Male	84	101	98	101			

source: program group for review 2019 06 27.sas

Table 23. Sample Sizes for Subgroup Analyses on Change from Baseline HbA1c, Total for Placebo-controlled Studies 4224, 42333, 4234, and 4280

Group	Subgroup	S14	S7	S3	Pbo	Sita	Empa	Lira
Age	< 65yr	483	238	231	385			210
	≥ 65yr	282	96	112	248			62
Ethnicity	Hisp	90	46	66	86			17
	Not_Hisp	661	267	258	534			255
Race	Asian	132	93	95	112			35
	Black	34	19	17	38			8
	White	558	211	214	456			203
Region	NorthAmer	209	103	108	182			70
	Not_NorthAmer	556	231	235	451			202
Sex	Female	378	147	160	310			129
	Male	387	187	183	323			143

source: program group for review 2019 06 27.sas

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

There are no outstanding statistical issues associated with this submission.

5.2 Collective Evidence

The collective evidence from primary and secondary endpoints in these studies consistently supports effectiveness of oral semaglutide for the treatment of type 2 diabetes.

5.3 Conclusions and Recommendations

Six randomized trials demonstrate effectiveness of semaglutide 7 or 14 mg administered as oral tablets once daily for the improvement of glycemic control in patients with type 2 diabetes mellitus. Four of the trials, 4224, 4233, 4234, and 4280, were conducted with a placebo control. One of the trials, 4224, included both a placebo and an active (liraglutide injection 1.8 mg qd) control, and two trials, 4222 and 4223, were conducted using respective active controls empagliflozin po 25 mg qd and sitagliptin 100 mg po qd.

Compared to the control, five of the six trials showed statistically significant effects of 7 and/or 14 mg semaglutide for the primary endpoint, week 26 change from baseline HbA1c. In trial 4224, however, no significant difference was seen between the only treatment arms, semaglutide 14 mg and another drug in the same class, liraglutide 1.8 mg. Of the five remaining trials testing statistical efficacy for the key secondary endpoint, week 26 change from baseline body weight, five showed statistically significant effects of both 7 and 14 mg semaglutide against control, and one, trial 4223 showed superiority of the 14 mg, but not the 7 mg dose, to placebo. Supportive endpoints, proportion of patients with HbA1c less than 7%, and week 26 change from fasting plasma glucose, all trended in the direction expected if semaglutide is efficacious.

The 7 mg and 14 mg maintenance doses recommended on the proposed product label are consistent with the positive dose-response curve seen in these trials. In particular, the proposed label recommends initiating treatment with a 3 mg dose, escalating to 7 mg after four weeks and then, if necessary after an additional four weeks, escalating from 7 to 14 mg.

Compared to existing treatments, i.e. the active controls, semaglutide was not seen to increase the rate of clinically significant or severe instances of hypoglycemia. Rates of severe adverse events associated with nausea, vomiting, diarrhea, and constipation were significantly higher among patients treated with semaglutide 14 mg than among placebo or active controls.

Based on efficacy and safety analyses results, I recommend approval of oral semaglutide for glycemic control in patients with type 2 diabetes mellitus.

5.4 Labeling Recommendations

Consistent with conclusions from the current review, the review team may wish to consider the following recommendations:

1. The last row of Table 2 in the proposed label, for hypoglycemic events, should be replaced with results from Table 20 of this review. In particular, Table 2 of the proposed label excludes patients who were on treatment but took rescue medication. The label should, as much as possible, reflect actual clinical practice, where administration of semaglutide or other treatments does not preclude use of rescue medications.

2. As in Table 18 of the current review, the digit 7 in Table 8 of the proposed label, for change in FPG, should be replaced with '5.3.' The '7' on the current label appears to be due to a typographical error.

3. Available data was not optimal to evaluate potential clinical effects of the gastrointestinal absorption enhancer, salcaprozate sodium, on patients administered concurrent oral medications, such as thyroxine for hypothyroidism. If this is an outstanding clinical concern, it may be desirable to request additional data collection, or to include advice in the product label under limitations of use, warnings and precautions, or use in specific populations.

6 Appendix: Shrinkage Estimates for Subgroup Analyses

As an example, given the following input dataset containing results from the non-shrinkage analysis by sex,

Trt	Class	Class2	Estimate	Stderr
S14vsPbo	Sex	M	-1.061762	0.168244
S14vsPbo	Sex	F	-1.147317	0.171678

shrinkage estimates were provided by the following SAS program:

(b) (4)

This is a representation of an electronic record that was signed electronically. Following this are manifestations of any and all electronic signatures for this electronic record.

/s/

ROBERT ABUGOV
07/30/2019 07:31:16 AM

YUN WANG
07/30/2019 07:38:31 AM

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

Statistical Review and Evaluation

CARCINOGENICITY STUDY

IND/NDA Number:	IND 114464
Drug Name:	Salcaprozate sodium (SNAC)
Indication:	An adjunct to diet and exercise to improve glycemic control in patients with type 2 diabetes mellitus.
Applicant:	**Sponsor:** Novo Nordisk PO Box 846 800 Scudders Mill Rd Plainsboro, NJ 08536 US **Test facility:** Oral Analytical Development Novo Nordisk A/S, Novo Nordisk Park DK-2760 Måløv, Denmark
Documents Reviewed:	Study (transgenic mice and Sprague-Dawley Rats) reports submitted on March 31, 2016 via IND114464/0131; Electronic data submitted on June 24, 2016 via IND114464/0146
Review Priority:	Standard
Biometrics Division:	Division of Biometrics -6
Statistical Reviewer:	Zhuang Miao, Ph.D.
Concurring Reviewer:	Mohammad Atiar Rahman, Ph.D. Karl Lin, Ph.D.
Medical Division:	Division of Metabolism and Endocrinology Products
Reviewing Pharmacologist:	Elmore, Calvin (Lee), Ph.D. Antonipillai, Indra, Ph.D.
Project Manager:	
Keywords:	Carcinogenicity, Dose response

Table of Contents

1. Summary

In this submission the sponsor included reports of two animal carcinogenicity studies, one in Sprague-dawley rats and one in transgenic rasH2 mice. The studies were intended to assess the carcinogenic potential of Salcaprozate sodium (SNAC) when administered orally by gavage at appropriate drug levels for 104 weeks in rats and 26 weeks in mice.

Rat Study: Two separate experiments were conducted, one in males and one in females. In each of these two experiments there were three treated groups and one vehicle control group. Two hundred and eighty-eight Sprague Dawley rats of each sex were randomly assigned to the treated and vehicle control groups in equal size of 72 mice per group. The dose levels for treated groups were 75, 200, and 500 mg/kg/day. The rats in the vehicle control group received the vehicle (water for injection) at the same volume dose as the treated groups. The study was designed to continue for up to 104 weeks for both sexes, however in accordance with study termination criteria, all surviving rats were sacrificed during Week 105.

The mortality data analysis results showed a statistically significant dose response relationship in mortality across the treatment groups in female rats. The pairwise comparisons showed a statistically significant increase in mortality in high dose group compared to the vehicle control in female rats. The mortality data analysis results did not show a statistically significant dose response relationship in mortality across the treatment groups in male rats. The pairwise comparison did not show statistically significant increases in mortality in any treated group in either sex compared to their respective vehicle control.

Based on the criterion for multiple testing adjustment, the incidence of lipoma (skin and subcutis) (p-value=0.0246) and combined incidences of liphoma and liposarcoma (multiple organs)(p-value=0.0078)) in male rats were considered to have statistically significant positive dose response relationship.

The counts of the tumor data anlaysis from Table 9 and Table 10 are the same as the sponsor's counts of tumor data analysis.

Mouse Study: Two separate experiments were conducted, one in males and one in females. In each of these two experiments there were three treated groups and one vehicle control group. One hundred Tg.rasH2 mice of each sex were randomly assigned to the treated and vehicle control groups in equal size of 25 mice per group. The dose levels for treated groups were 30, 100, and 300 mg/kg/day. The mice in the vehicle control group received the vehicle (water for injection) only. The study was designed to continue for up to 26 weeks for both sexes, however in accordance with study termination criteria, all surviving mice were sacrificed during Week 27. The positive control group with 5 animals and treated with 37.5 mg/kg MNUC at the start of the study was excluded from analysis.

The mortality data analysis results did not show a statistically significant dose response relationship in mortality across the treatment groups in either sex. The pairwise comparison also did not show statistically significant increases in mortality in any treated group in either sex compared to their respective control.

The tumor data analysis results did not show statistically significant dose response relationship in tumor incidence in any observed tumor types across the treatment groups in either sex. The pairwise comparisons did not show statistically significant increases in incidence in any observed tumor type in any treated group in either sex when compared to their respective vehicle control.

The counts of the tumor data anlaysis from Table 15 and Table 16 are the same as the sponsor's counts of tumor data analysis.

2. Background

In this submission the sponsor included reports of two animal carcinogenicity studies, one in Sprague-dawley rats and one in transgenic rasH2 mice. These studies were intended to assess the carcinogenic potential of Salcaprozate sodium (SNAC) when administered orally by gavage at appropriate drug levels for 104 weeks in rats and 26 weeks in mice. Results of this review have been discussed with the reviewing pharmacologist Dr. Antonipillai. This review analyzed the SAS data sets of these studies received from the sponsor on June 24, 2016 via IND114464/0146.

In this review the phrase "dose response relationship" refers to the linear component of the effect of treatment, and not necessarily to a strictly increasing or decreasing mortality or tumor incidence rate as the dose increases.

3. Rat Study

Two separate experiments were conducted, one in males and one in females. In each of these two experiments there were three treated groups and one control group. Two hundred and eighty eight Sprague Dawley rats of each sex were randomly allocated to treated and control groups in equal size of 72 animals. The dose levels for treated groups were 75, 200, and 500 mg/kg/day for males and females. In this review these dose groups would be referred to as the low, medium, and high dose group, respectively. The one control group will be referred to as vehicle control. The control received the vehicle (water for injection) 0 mg/kg/day.

For females, dosing at 200 mg/kg/day (Group 3) was suspended during Week 103 and dosing at 500 mg/kg/day (Group 4) was suspended during Week 99. For the purposes of the statistical trend analysis the weighted average dose over the 104 week study period was calculated to be 197.8 mg/kg/day and 476.6 mg/kg/day respectively and these nominal dose levels were used in the sponsor's analysis. In the reviewer's analysis, 200 mg/kg/day and 500 mg/kg/day were used.

Table 1: Study Design in Rat Study

Protocol Group No.	Salcaprozate sodium (SNAC) Dose Levels (mg/kg/day)	Number of Animals Examined	
		Males	Females
1	0	72	72
2	75	72	72
3	200	72	72
4	500	72	72

3.1. Sponsor's analyses

3.1.1. Survival analysis

The number of animal deaths during the study, up to terminal sacrifice, were analysed by logrank tests for a trend across the groups (Mantel 1966, Peto 1974). Males and females were analysed separately. The numbers of animal deaths during the study were presented as life-tables and Kaplan-Meier survival curves (Kaplan and Meier 1958). These tables divide the study into time strata, where the strata are defined as those weeks during

which there were deaths; N is the number of survivors at the start of the time period, R is the number that died during that period. The following statistical tests were carried out:

1) a two-tailed test for a trend with dose level.
2) a two-tailed pairwise comparison test of each treatment group against the control group.

Where the test for trend was statistically significant, the highest dose group was excluded and the trend test repeated (using a one-tailed test in the direction identified with all groups included), until the test was no longer statistically significant.

As a check, tests for non-linearity (not presented) were carried out. In this study, the non-linearity tests were not statistically significant at the 1% level and thus the results of the trend tests are to be preferred.

Significance levels were calculated using χ^2 tests and adjusted with a continuity correction where there was only one degree of freedom. These p-values have been quoted in the logrank results tables.

Sponsor's findings: Sponsor's analysis showed the numbers (percent) of death were 43 (59.7%), 49 (68.1%), 44 (61.1%) and 38 (52.8%) in vehicle control, low, medium, and high dose groups, respectively in males and 43 (59.7%), 45 (62.5%), 53 (73.6%) and 55 (73.6%), respectively in females. The sponsor made the following conclusions:

Males
The trend test was not statistically significant when all groups were included in the analysis. None of the pairwise comparisons were statistically significant.

Females
The trend test, when all treated groups were included, was statistically significant (p=0.008). Upon exclusion Group 4, the trend test was still significant (p=0.038). Upon further exclusion Group 3, the trend test was no longer significant. The pairwise comparison of the control group with Group 4 was statistically significant (p=0.038).

3.1.2. Tumor data analysis

Tumours were categorised as benign or malignant.

For non-palpable tumours, each tumour was categorised as non-incidental if the tumour was a factor contributing towards the death of the animal, incidental otherwise. For statistical purposes, all animals that died after terminal sacrifice commenced (Week 105) were considered terminal and the tumours observed in these animals were categorised as incidental.

For palpable tumours, each tumour was classified as non-incidental if it was palpable before death and before the terminal sacrifice commenced, or, if it was a factor contributing towards the death of the animal. The tumour was classified as incidental, if the tumour was first found after death and was not a factor contributing towards the death of the animal, or if it was first found in or after the first week of the terminal sacrifice.

The analyses were carried out for benign, malignant and benign and malignant tumours combined. If an animal had a benign and a malignant non-palpable tumour then only the malignant tumour was included in the analysis of both tumours together. Also, if an animal had more than one palpable tumour of the same category, benign or malignant, then only the first observed tumour was included in the analysis.

Tumour types were selected for full statistical analysis where at least two tumor bearing animals were observed in treated groups for which all animals were examined.

For incidental tumours, the following fixed time intervals were used to adjust for differential mortality between the treatment groups: 1-52, 53-78, 79-92, 93-104 weeks and terminal sacrifice.

Log-rank methods were used to analyse the number of animals with tumours across treatment groups (Mantel 1966, Peto 1974, Peto et al 1980).

The log-rank analysis results tables show the total observed and expected number of tumours and the relative tumour rate across all time intervals. The observed and the expected frequencies were calculated for each time interval under a hypothesis of no difference between treatment groups. The observed and expected frequencies were then compared across the treatment groups. The following $\chi 2$ statistical tests were carried out:
1) a one-tailed test for a trend using nominal dose levels (weighted average dose levels where appropriate), with the control group
2) a one-tailed pairwise comparison test of each treatment group against the control group

Where the test for trend showed a p-value < 5%, the highest dose group was excluded and the trend test was repeated, using a one-tailed test until the test showed a p-value $\geqslant$5%. The significance levels were adjusted using a continuity correction where there was one degree of freedom.

As a check, tests for non-linearity were carried out (not presented). For malignant fibrosarcoma (skin) in males and for benign adenoma (mammary areas) in females the non-linearity tests were significant (p=0.004 and p=0.005 respectively), therefore the pairwise tests are to be preferred to the trend test. For all other analyses the non-linearity tests were not statistically significant at the 1% level, hence the trend tests are to be preferred.

Where there were fewer than ten observed tumor bearing animals across all treatment groups, exact one-tailed p-values were calculated using permutation tests for stratified contingency tables, to test for trend and for pairwise comparisons of each treatment group against the control group (SAS Institute Inc., 2002).

The spnsor followed the 2001 FDA guidance for the carcinogenicity studies in its adjustment for multiple testing.

Sponsor's findings: There were no statistically significant differences between the control and treated groups in male rats or female rats.

3.2. Reviewer's analyses

To verify sponsor's analyses and to perform additional analysis suggested by the reviewing pharmacologist, this reviewer independently performed survival and tumor data analyses. Data used in this reviewer's analyses were

provided by the sponsor electronically via IND114464/0146.

3.2.1. Survival analysis

The survival distributions of animals in all four treatment groups were estimated by the Kaplan-Meier product limit method. The dose response relationship and homogeneity of survival distributions were tested for vehicle control, low, medium and high dose groups using the Likelihood Ratio test and the Log-Rank test. The intercurrent mortality data are given in Tables 5 and 6 in the appendix for males and females, respectively. The Kaplan-Meier curves for survival rate are given in Figures 1 and 2 in the appendix for males and females, respectively. Results of the tests for dose response relationship and homogeneity of survivals, are given in Tables 7 and 8 in the appendix for males and females, respectively.

Reviewer's findings: This reviewer's analysis showed the numbers (percent) of death were 43 (59.72%), 50 (69.44%), 44 (61.11%), and 38 (52.78%) in male rats and 44 (61.11%), 46 (63.89%), 53 (73.61%), and 54 (75%) in female rats in the vehicle control, low, median and high dose groups, respectively.

The tests did not show a statistically significant dose responserelationship in mortality across vehicle control and treated groups for male rats. The tests did show a statistically significant dose responserelationship in mortality across VC and treated groups for female rats (p=0.0116). The pairwise comparisons did show a statistically significant mortality between vehicle control and high dose group female rats (p=0.0336).

Table 2: Intercurrent Mortality Comparison between Treated Groups and Vehicle Control with P-Values ⩽ 0.05, Female Rats

Test	Statistic	P_Value Dose Resp	P_Value VC vs. High
Dose-Response	Likelihood Ratio	0.0116	0.0336
Homogeneity	Log-Rank	0.0560	0.0312

Reviewer's comment: *For male rats, the sponsor's analysis showed 49 (68.1%) death, while this reviewer's analysis showed 50 (669.44%) death in low dose group. These differences are due to the facts that the sponsor considered one animal (animal 133 in low dose group) to be accidental death and treated animal 133 as censored observation but the reviewer treated animal 133 as death in the analysis.*

For female rats, the sponsor's analysis showed 43 (59.7%), 45 (62.5%) and 55 (73.6%) death, while this reviewer's analysis showed 44 (61.11%), 46 (63.89%) and 54 (75%) death in vehicle control, low and high dose groups, respectively. These differences are due to the facts that the sponsor considered three animals (animal 362 in vehicle contro group, animal 415 in low dose group and animal 572 in high dose group) to be accidental death and treated these three animals as censored observation but the reviewer treated these three animal as death in the analysis.

These discrepancies did not significantly affect the survival analysis results.

3.2.2. Tumor data analysis

The tumor data were analyzed for dose response relationships and pairwise comparison of each of the treated groups with control group. Both the dose response relationship tests and pairwise comparisons were performed using the Poly-K method described in the paper of Bailer and Portier (1988) and Bieler and Williams (1993). In

this method an animal that lives the full study period (w_{max}) or dies before the terminal sacrifice but develops the tumor type being tested gets a score of s_h =1. An animal that dies at week w_h without a tumor before the end of the study gets a score of $s_h = \left(\frac{w_h}{w_{max}}\right)^k < 1$. The adjusted group size is defined as Σs_h. As an interpretation, an animal with score s_h =1 can be considered as a whole animal while an animal with score $s_h < 1$ can be considered as a partial animal. The adjusted group size Σs_h is equal to N (the original group size) if all animals live up to the end of the study or if each animal that dies before the terminal sacrifice develops at least one tumor, otherwise the adjusted group size is less than N. These adjusted group sizes are then used for the dose response relationship (or the pairwise) tests using the Cochran-Armitage test. One critical point for Poly-k test is the choice of the appropriate value of k, which depends on the tumor incidence pattern with the increased dose. For long term 104 week standard rat and mouse studies, a value of k=3 is suggested in the literature. Hence, this reviewer used k=3 for the analysis of this data. For the calculation of p-values the exact permutation method was used. The tumor rates and the p-values for the positive dose response relationship tests and pairwise comparisons are listed in Tables 9 and 10 in the appendix for male and female rats, respectively.

Adjustment for multiple testing: For the adjustment of multiple testing of dose response relationship for a submission with one chronic rat study and one transgenic mouse studies, the more recently revised draft (January, 2013) FDA guidance for the carcinogenicity studies suggests the use of test levels a=0.005 for common tumors and a=0.025 for rare tumors for the chronic rat study, and a significance level and a=0.05 for both common and rare tumors for the transgenic mouse study. For pairwise comparisons, the same guidance document suggests the use of test levels a=0.01 for common tumors and a=0.05 for rare tumors for the chronic rat study, and also a significance level a=0.05 for both common and rare tumors for the transgenic mouse study.

It should be noted that the FDA guidance for multiple testing for dose response relationship is based on a publication by Lin and Rahman (1998). In this work the authors investigated the use of this rule for Peto analysis. However, in a later work Rahman and Lin (2008) showed that this rule for multiple testing for dose response relationship is also suitable for Poly-K tests.

Reviewer's findings: Following tumor types showed p-values less than or equal to 0.05 either for dose response relationship and/or pairwise comparisons of control and treated groups.

Table 3: Tumor Types with P-Values ⩽ 0.05 for Dose Response Relationship or Pairwise Comparisons -Male Rats

Organ Name	Tumor Name	0 mg|kg|day Vehicle Control N=72	75 mg|kg|day Low N=72	200 mg|kg|day Med N=72	500 mg|kg|day High N=72	P_Value Dose Resp	P_Value VC vs. L	P_Value VC vs. M	P_Value VC vs. H
Male									
ADRENALS	**C_Pheoc_M+B**	**8**	**17**	**10**	**15**	**0.2797**	**0.0219**	**0.3285**	**0.1125**
HEMOPOIETIC	LYMPHOMA	0	0	3	3	0.0416	.	0.1142	0.1394

Organ Name	Tumor Name	0 mg|kg|day Vehicle Control N=72	75 mg|kg|day Low N=72	200 mg|kg|day Med N=72	500 mg|kg|day High N=72	P_Value Dose Resp	P_Value VC vs. L	P_Value VC vs. M	P_Value VC vs. H
SYSTEM									
Multiple Organs	**C_Lipomas+Liposarcomas**	**0**	**0**	**1**	**4**	**0.0078***	**.**	**0.4851**	**0.0685**
SKIN AND SUBCUTIS	**C_Fib+Fib dermoid+Fibrosa**	**7**	**11**	**8**	**16**	**0.0441**	**0.1814**	**0.4353**	**0.0455**
	C_Fibroma+Fibr dermoid	**6**	**5**	**8**	**15**	**0.0080**	**0.6780**	**0.3290**	**0.0384**
	FIBROSARCOMA	1	6	0	1	0.8897	0.0473	1.0000	0.7705
	LIPOMA	0	0	1	3	0.0246*	.	0.4851	0.1358
Female									
KIDNEYS	**C_Lipoma+Re Liposa**	**0**	**0**	**0**	**2**	**0.0475**	**.**	**.**	**0.2013**
MAMMARY	**C_Adenocarci+Fibr+Ade**	**46**	**45**	**42**	**52**	**0.0305**	**0.6544**	**0.6655**	**0.0727**

Based on the above criterion for multiple testing adjustment, the incidences of lipoma (skin and subcutis) (p-value=0.0246) and combined incidences of liphoma and liposarcoma (multiple organs)(p-value=0.0078) in male rats were considered to have statistically significant positive dose response relationship. The pairwise comparison of treated group with the vehicle control did not reach the statistical significant levels. The detailed information of combined incidences of liphoma and liposarcoma(multiple organs) can be found in table 9.

4. Mouse Study

Two separate experiments were conducted, one in males and one in females. In each of these two experiments there were three treated groups and one vehicle control group. One hundred Tg.rasH2 mice of each sex were randomly assigned to the treated and vehicle control groups in equal size of 25 mice per group. The dose levels for treated groups were 30, 100 and 300 mg/kg/day. In this review these dose groups would be referred to as the low, medium, and high dose groups, respectively. The mice in the vehicle control group received the vehicle (water for injection) only. The study was designed to continue for up to 26 weeks for both sexes. All surviving mice were sacrificed during Week 27. The positive control group with 5 animals and treated with 37.5 mg/kg MNUC at the start of the study was excluded from analysis.

Table 4: Study Design in Mouse Study

Protocol Group No.	Salcaprozate sodium (SNAC) Dose Levels (mg/kg/day)	Number of Animals Examined	
		Males	Females
1	0	25	25
2	30	25	25
3	100	25	25
4	300	25	25

4.1. Sponsor's analyses

4.1.1. Survival analysis

The sponsor used the same statistical methods as for the rats study.

Sponsor's findings: The sponsor analysis did not show a statistically significant dose response relationship in mortality among the drug-treated groups and vehicle control group in either sex.

Males
The trend test was not statistically significant when all groups were included in the analysis. None of the pairwise comparisons were statistically significant.

Females
The trend test was not statistically significant when all groups were included in the analysis. None of the pairwise comparisons were statistically significant.

4.1.2. Tumor data analysis

The sponsor used the same statistical methods as for the rats study.

Sponsor's findings: The sponsor's analyses did not show a statistically significant dose response relationship among the treated and vehicle groups in any of the observed tumor types in male or female mice.

4.2. Reviewer's analyses

To verify the sponsor's analyses and to perform additional analyses suggested by the reviewing pharmacologist, the reviewer independently performed survival and tumor data analyses. Data used in this reviewer's analyses were provided by the sponsor electronically in IND114464/0146. The significance level for all statistical tests was set at 0.05.

4.2.1. Survival analysis

The survival distributions of three treated groups, one vehical control group and one positive control group were estimated using the Kaplan-Meier product limit method. The dose response relationship in survival was tested using the likelihood ratio test and the homogeneity of survival distributions was tested using the log-rank test. The Kaplan-Meier curves for survival rates are given in Figures 3 and 4 in the appendix for male and female mice, respectively. The intercurrent mortality data are given in Tables 11 and 12 in the appendix for male and female mice, respectively. Results of the tests for dose response relationship and homogeneity of survivals among the vehicle control and three treated groups are given in Tables 13 and 14 in the appendix for male and female mice, respectively.

Reviewer's findings: This reviewer's analysis showed 12.0%, 0%, 0%, and 4.0% mortalities in male mice, and 4.0%, 0.0%, 4.0%, and 4.0% mortalities in female mice in vehicol control, low, medium, and high dose groups, respectively. The tests did not show a statistically significant dose response relationship in mortality across the treated groups and vehicle control in either sex. The pairwise comparison also did not show

statistically significant decreases in mortality in any treated groups in either sex when compared to their vehicle control.

4.2.2. Tumor data analysis

The tumor data were analyzed for dose response relationships and pairwise comparison of each of the treated groups with vehicle control group. Both the dose response relationship tests and pairwise comparisons were performed using the Poly-K method described in the paper of Bailer and Portier (1988) and Bieler and Williams (1993).

In this method an animal that lives through the full study period ($w_{\max}$) or dies before the terminal sacrifice but develops the tumor type being tested gets a score of s_h =1. An animal that dies at week w_h without a tumor before the end of the study gets a score of $s_h = \left(\frac{w_h}{w_{\max}} \right)^k < 1$. The adjusted group size is defined as Σs_h. As an interpretation, an animal with score s_h =1 can be considered as a whole animal while an animal with score $s_h < 1$ can be considered as a partial animal. The adjusted group size Σs_h is equal to N (the original group size) if all animals live up to the end of the study or if each animal that dies before the terminal sacrifice develops the tumor type being tested, otherwise the adjusted group size is less than N. These adjusted group sizes are then used for the dose response relationship (or the pairwise) tests using the Cochran-Armitage test.

One critical point for Poly-k test is the choice of the appropriate value of k, which depends on the tumor incidence pattern with the increased dose. For long term 104 week standard rat and mouse studies, a value of k=3 is suggested in the literature. The present study is a 26 week study. For this kind of medium or short term study no such suggested value of k in the literature is known to this reviewer. Following the suggested value for long term studies, this reviewer analyzed the tumor data using k=3. For the calculation of p-values the exact permutation method was used.

The tumor rates and the p-values for the positive dose response relationship tests and pairwise comparisons among vehicle control and three treated groups are listed in Tables 15 and 16 in the appendix for male and female mice, respectively.

Adjustment for multiple testing: For the adjustment of multiple testing of dose response relationship for a submission with one chronic rat study and one transgenic mouse studies, the more recently revised draft (January, 2013) FDA guidance for the carcinogenicity studies suggests the use of test levels a=0.005 for common tumors and a=0.025 for rare tumors for the chronic rat study, and a significance level and a=0.05 for both common and rare tumors for the transgenic mouse study. For pairwise comparisons, the same guidance document suggests the use of test levels a=0.01 for common tumors and a=0.05 for rare tumors for the chronic rat study, and also a significance level a=0.05 for both common and rare tumors for the transgenic mouse study.

It should be noted that the FDA guidance for multiple testing for dose response relationship is based on a publication by Lin and Rahman (1998). In this work the authors investigated the use of this rule for Peto analysis. However, in a later work Rahman and Lin (2008) showed that this rule for multiple testing for dose response relationship is also suitable for Poly-K tests.

Reviewer's findings: Based on the criteria of adjustment for multiple testing discussed in the mouse data analysis section, no tumor types have statistically significant dose response relationship in male or female mice. The pairwise comparisons did not show statistically significant increases in incidence in any observed tumor type in any treated group in either sex when compared to their respective control.

5. Conclusion

In this submission the sponsor included reports of two animal carcinogenicity studies, one in Sprague-dawley rats and one in transgenic rasH2 mice. The studies were intended to assess the carcinogenic potential of Salcaprozate sodium (SNAC) when administered orally by gavage at appropriate drug levels for 104 weeks in rats and 26 weeks in mice.

Rat Study: Two separate experiments were conducted, one in males and one in females. In each of these two experiments there were three treated groups and one vehicle control group. Two hundred and eighty-eight Sprague Dawley rats of each sex were randomly assigned to the treated and vehicle control groups in equal size of 72 mice per group. The dose levels for treated groups were 75, 200, and 500 mg/kg/day. The rats in the vehicle control group received the vehicle (water for injection) at the same volume dose as the treated groups. The study was designed to continue for up to 104 weeks for both sexes, however in accordance with study termination criteria, all surviving rats were sacrificed during Week 105.

The mortality data analysis results showed a statistically significant dose response relationship in mortality across the treatment groups in female rats. The pairwise comparisons showed a statistically significant increase in mortality in high dose group compared to their control in female rats. The mortality data analysis results did not show a statistically significant dose response relationship in mortality across the treatment groups in male rats. The pairwise comparison did not show statistically significant increases in mortality in any treated group in either sex compared to their respective control.

Based on the above criterion for multiple testing adjustment, the incidences of)lipoma (skin and subcutis) (p-value=0.0246) and combined incidences of liphoma and liposarcoma (multiple organs)(p-value=0.0078)) in male rats were considered to have statistically significant positive dose response relationship.

The counts of the tumor data anlaysis from Table 9 and Table 11 are the same as the sponsor's counts of tumor data analysis.

Mouse Study: Two separate experiments were conducted, one in males and one in females. In each of these two experiments there were three treated groups and one vehicle control group. One hundred Tg.rasH2 mice of each sex were randomly assigned to the treated and vehicle control groups in equal size of 25 mice per group. The dose levels for treated groups were 30, 100, and 300 mg/kg/day. The mice in the vehicle control group received the vehicle (water for injection) at the same volume dose as the treated groups. The study was designed to continue for up to 26 weeks for both sexes, however in accordance with study termination criteria, all surviving mice were sacrificed during Week 27. The positive control group with 5 animals and treated with 37.5 mg/kg MNUC at the start of the study was excluded from analysis.

The mortality data analysis results did not show a statistically significant dose response relationship in mortality across the treatment groups in either sex. The pairwise comparison also did not show statistically significant increases in mortality in any treated group in either sex compared to their respective control.

The tumor data analysis results did not show statistically significant dose response relationship in tumor incidence in any observed tumor types across the treatment groups in either sex. The pairwise comparisons did not show statistically significant increases in incidence in any observed tumor type in any treated group in either sex when compared to their respective vehicle control.

The counts of the tumor data anlaysis from Table 15 and Table 16 are the same as the sponsor's counts of tumor data analysis.

Zhuang Miao, Ph.D.
Mathematical Statistician

Concur:
Mohammad Atiar Rahman, Ph.D., Acting Deputy Director
Mathematical Statistician, Biometrics-6

Karl Lin, Ph.D.
Mathematical Statistician, Team Leader, Biometrics-6

cc:
Archival NDA 208552
Yi Tsong, Ph.D.
Antonipillai, Indra, Ph.D.
Elmore, Calvin (Lee), Ph.D.

6. Appendix

Table 5: Intercurrent Mortality Rate -Male Rats

	0 mg|kg|day		75 mg|kg|day		200 mg|kg|day		500 mg|kg|day	
Week	**No. of Death**	**Cum. %**	**No. of Death**	**Cum. %**	**No. of Death**	**Cum. %**	**No. of Death**	**Cum. %**
0 - 52	3	4.17	6	8.33	6	8.33	4	5.56
53 - 78	14	23.61	15	29.17	14	27.78	9	18.06
79 - 91	10	37.50	13	47.22	11	43.06	6	26.39
92 - 104	16	59.72	16	69.44	13	61.11	19	52.78
Ter. Sac.	29	40.28	22	30.56	28	38.89	34	47.22

Cum. %: Cumulative percentage except for Ter. Sac.

Table 6: Intercurrent Mortality Rate -Female Rats

	0 mg|kg|day		75 mg|kg|day		200 mg|kg|day		500 mg|kg|day	
Week	**No. of Death**	**Cum. %**	**No. of Death**	**Cum. %**	**No. of Death**	**Cum. %**	**No. of Death**	**Cum. %**
0 - 52	2	2.78	1	1.39	4	5.56	4	5.56
53 - 78	16	25.00	16	23.61	14	25.00	27	43.06
79 - 91	10	38.89	11	38.89	19	51.39	10	56.94
92 - 104	16	61.11	18	63.89	16	73.61	13	75.00
Ter. Sac.	28	38.89	26	36.11	19	26.39	18	25.00

Cum. %: Cumulative percentage except for Ter. Sac.

Table 7: Intercurrent Mortality Comparison between Treated Groups and Vehicle Control -Male Rats

Test	Statistic	P_Value Dose Resp	P_Value VC vs. Low	P_Value VC vs. Medium	P_Value VC vs. High
Dose-Response	Likelihood Ratio	0.1054	0.2349	0.7019	0.3276
Homogeneity	Log-Rank	0.1744	0.2299	0.6997	0.3228

Table 8: Intercurrent Mortality Comparison between Treated Groups and Vehicle Control -Female Rats

Test	Statistic	P_Value Dose Resp	P_Value VC vs. Low	P_Value VC vs. Med	P_Value VC vs. High
Dose-Response	Likelihood Ratio	0. 0116	0.9559	0.1299	0.0336
Homogeneity	Log-Rank	0.0560	0.9554	0.1257	0.0312

Table 9: Tumor Rates and P-Values for Dose Response Relationship and Pairwise Comparisons -Male Rats

Organ Name	Tumor Name	0 mg|kg|day Vehicle Control N=72	75 mg|kg|day Low N=72	200 mg|kg|day Med N=72	500 mg|kg|day High N=72	P_Value Dose Resp	P_Value VC vs. L	P_Value VC vs. M	P_Value VC vs. H
ABDOMINAL CAVITY	HEMANGIOSARCOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	PARAGANGLIOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
ADRENALS	ADENOMA, CORTICAL	0	1	0	1	0.3340	0.4747	.	0.5185
	CARCINOMA, CORTICAL	0	1	0	0	0.7463	0.4800	.	.
	C_Cortical_Ade+Carci	**0**	**2**	**0**	**1**	**0.4669**	**0.2279**	**.**	**0.5185**
	C_Pheoc_M+B	**8**	**17**	**10**	**15**	**0.2797**	**0.0219**	**0.3285**	**0.1125**
	LEIOMYOSARCOMA	0	0	0	1	0.2745	.	.	0.5185
	PHEOCHROMOCYTOMA	8	15	10	10	0.6770	0.0568	0.3285	0.4655
	PHEOCHROMOCYTOMA, MALIGNANT	1	3	1	5	0.0781	0.2714	0.7374	0.1208
All Bone	**C_Chondromas+Osteosarcomas+Ost**	**0**	**1**	**1**	**1**	**0.3214**	**0.4800**	**0.4851**	**0.5185**
BONE (OTHER)	OSTEOMA	0	0	0	1	0.2745	.	.	0.5185
	OSTEOSARCOMA	0	1	0	0	0.7463	0.4800	.	.
BONE, STERNUM	OSTEOSARCOMA	0	0	1	0	0.5147	.	0.4851	.
BRAIN	ADENOMA, PINEAL GLAND	0	1	0	0	0.7463	0.4800	.	.
	ASTROCYTOMA	2	0	2	0	0.8479	1.0000	0.6685	1.0000
	GLIOMA, MIXED	0	0	1	0	0.5147	.	0.4851	.
	OLIGODENDROGLIOMA	0	1	0	0	0.7463	0.4800	.	.
	SCHWANNOMA, MALIGNANT	0	0	1	0	0.5171	.	0.4902	.
	TUMOR, GRANULAR CELL, BENIGN	2	0	2	0	0.8479	1.0000	0.6685	1.0000
EXTREMITIES	PAPILLOMA, SQUAMOUS CELL	0	1	0	0	0.7463	0.4800	.	.
EYELIDS	FIBROSARCOMA	0	1	0	0	0.7451	0.4747	.	.
HEAD	CARCINOMA, SQUAMOUS CELL	1	0	0	0	1.0000	1.0000	1.0000	1.0000
HEART	SCHWANNOMA, MALIGNANT	0	0	1	0	0.5171	.	0.4902	.
HEMOPOIETIC SYSTEM	LEUKEMIA, MYELOID CELL	0	0	2	0	0.5313	.	0.2378	.
	LYMPHOMA	0	0	3	3	0.0416	.	0.1142	0.1394
	SARCOMA, HISTIOCYTIC	1	0	1	2	0.2008	1.0000	0.7374	0.5347
JEJUNUM	ADENOCARCINOMA	0	1	0	0	0.7463	0.4800	.	.
KIDNEYS	CARCINOMA, TUBULAR	1	0	0	0	1.0000	1.0000	1.0000	1.0000

Organ Name	Tumor Name	0 mg|kg|day Vehicle Control N=72	75 mg|kg|day Low N=72	200 mg|kg|day Med N=72	500 mg|kg|day High N=72	P_Value Dose Resp	P_Value VC vs. L	P_Value VC vs. M	P_Value VC vs. H
	C_Lipoma+Renal Liposa	**0**	**0**	**0**	**2**	**0.0744**	**.**	**.**	**0.2665**
	LIPOMA	0	0	0	1	0.2745	.	.	0.5185
	LIPOSARCOMA, RENAL	0	0	0	1	0.2745	.	.	0.5185
	SARCOMA, RENAL	0	0	1	0	0.5147	.	0.4851	.
LIVER	ADENOMA, HEPATOCELLULAR	1	3	1	1	0.7232	0.2714	0.7374	0.7705
	CARCINOMA, HEPATOCELLULAR	0	0	1	0	0.5147	.	0.4851	.
	C_Hepatocellular Ade+Carci	**1**	**3**	**2**	**1**	**0.7148**	**0.2714**	**0.4775**	**0.7705**
	HEMANGIOSARCOMA	0	0	1	0	0.5147	.	0.4851	.
MAMMARY	ADENOCARCINOMA	1	0	0	1	0.4746	1.0000	1.0000	0.7705
	C_Adenocarci+Fibroade	**1**	**2**	**0**	**2**	**0.3944**	**0.4618**	**1.0000**	**0.5280**
	FIBROADENOMA	0	2	0	1	0.4689	0.2228	.	0.5185
Multiple Organs	**C_Lipomas+Liposarcomas**	**0**	**0**	**1**	**4**	**0.0078***	**.**	**0.4851**	**0.0685**
	C_hemangiosar+heman	**2**	**2**	**1**	**2**	**0.5561**	**0.6604**	**0.8674**	**0.7185**
PANCREAS	ADENOMA, ACINAR CELL	0	0	1	1	0.2069	.	0.4851	0.5185
	ADENOMA, ISLET CELL	7	10	4	7	0.7285	0.2527	0.8803	0.6685
	CARCINOMA, ISLET CELL	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	C_Aci Cell Aden+Is Cell Aden	**7**	**10**	**5**	**8**	**0.6138**	**0.2527**	**0.7911**	**0.5622**
	C_I_cell_adeno+carci	**8**	**10**	**4**	**7**	**0.7907**	**0.3448**	**0.9249**	**0.7615**
PITUITARY	ADENOMA, PARS DISTALIS	36	32	33	31	0.8139	0.6928	0.5812	0.8602
	ADENOMA, PARS INTERMEDIA	0	0	0	1	0.2745	.	.	0.5185
	CARCINOMA, PARS DISTALIS	0	0	1	1	0.2069	.	0.4851	0.5185
	C_Dis pars_adeno+carci	**36**	**32**	**34**	**32**	**0.7872**	**0.6928**	**0.5046**	**0.8403**
PROSTATE	CARCINOSARCOMA	0	1	0	0	0.7463	0.4800	.	.
SKELETAL MUSCLE (O	FIBROSARCOMA	0	0	1	0	0.5147	.	0.4851	.
	FIBROSARCOMA, PLEOMORPHIC	0	1	0	0	0.7451	0.4747	.	.
	HEMANGIOSARCOMA	0	1	0	0	0.7463	0.4800	.	.
SKIN AND SUBCUTIS	ADENOMA, SEBACEOUS CELL	2	2	0	0	0.9805	0.6445	1.0000	1.0000
	CARCINOMA, BASAL CELL	0	0	1	0	0.5147	.	0.4851	.
	C_Fib+Fib dermoid+Fibrosa	**7**	**11**	**8**	**16**	**0.0441**	**0.1814**	**0.4353**	**0.0455**
	C_Fibroma+Fibr dermoid	**6**	**5**	**8**	**15**	**0.0080**	**0.6780**	**0.3290**	**0.0384**
	C_Squ Cell Papil+Keratoa	**8**	**8**	**5**	**6**	**0.8223**	**0.5559**	**0.8587**	**0.8516**
	FIBROMA	3	3	2	7	0.0702	0.6228	0.7933	0.1836
	FIBROMA (DERMOID)	4	2	6	8	0.0648	0.8732	0.3339	0.2174

Organ Name	Tumor Name	0 mg|kg|day Vehicle Control N=72	75 mg|kg|day Low N=72	200 mg|kg|day Med N=72	500 mg|kg|day High N=72	P_Value Dose Resp	P_Value VC vs. L	P_Value VC vs. M	P_Value VC vs. H
	FIBROSARCOMA	1	6	0	1	0.8897	0.0473	1.0000	0.7705
	HISTIOCYTOMA, MALIGNANT, FIBRO	0	0	0	1	0.2745	.	.	0.5185
	KERATOACANTHOMA	6	4	5	4	0.7402	0.8154	0.7019	0.8686
	LIPOMA	0	0	1	3	0.0246*	.	0.4851	0.1358
	OSTEOSARCOMA	0	1	0	0	0.7451	0.4747	.	.
	PAPILLOMA, SQUAMOUS CELL	3	4	0	2	0.8140	0.4549	1.0000	0.8407
	SARCOMA NOS	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	TUMOR, HAIR FOLLICLE	1	0	1	1	0.4254	1.0000	0.7374	0.7705
SPINAL CORD CERVIC	ASTROCYTOMA	0	0	0	1	0.2745	.	.	0.5185
SPLEEN	FIBROSARCOMA	0	1	0	0	0.7463	0.4800	.	.
	HEMANGIOMA	0	0	0	1	0.2745	.	.	0.5185
	HEMANGIOSARCOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	SARCOMA NOS	0	0	1	0	0.5147	.	0.4851	.
STOMACH	ADENOMA	0	0	0	1	0.2745	.	.	0.5185
	CARCINOMA, SQUAMOUS CELL	0	1	0	0	0.7451	0.4747	.	.
TAIL	FIBROSARCOMA	0	0	0	1	0.2745	.	.	0.5185
TESTES	ADENOMA, INTERSTITIAL (LEYDIG)	0	0	0	2	0.0744	.	.	0.2665
	C_Heman+Hemangiosar	**0**	**1**	**0**	**1**	**0.3334**	**0.4800**	**.**	**0.5185**
	HEMANGIOMA	0	1	0	0	0.7463	0.4800	.	.
	HEMANGIOSARCOMA	0	0	0	1	0.2745	.	.	0.5185
THYMUS	THYMOMA, EPITHELIAL	0	0	1	0	0.5147	.	0.4851	.
		1	0	0	0	1.0000	1.0000	1.0000	1.0000
THYROIDS	ADENOMA, C-CELL	5	5	7	8	0.2226	0.5646	0.3393	0.3274
	ADENOMA, C-CELL, GANGLIONEUROM	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	ADENOMA, FOLLICULAR CELL	0	2	1	0	0.7550	0.2228	0.4851	.
	CARCINOMA, C-CELL	1	0	1	0	0.7657	1.0000	0.7374	1.0000
	C_C-cell Ade+Carci	**7**	**5**	**8**	**8**	**0.3753**	**0.7693**	**0.4675**	**0.5610**
ZYMBAL'S GLAND	PAPILLOMA, SQUAMOUS CELL	0	0	0	2	0.0763	.	.	0.2712

Table 10: Tumor Rates and P-Values for Dose Response Relationship and Pairwise Comparisons -Female Rats

Organ Name	Tumor Name	0 mg|kg|day Vehicle Control N=72	75 mg|kg|day Low N=72	200 mg|kg|day Med N=72	500 mg|kg|day High N=72	P_Value Dose Resp	P_Value VC vs. L	P_Value VC vs. M	P_Value VC vs. H
ABDOMINAL CAVITY	LEIOMYOSARCOMA	0	0	1	0	0.4635	.	0.4796	.
ADRENALS	ADENOMA, CORTICAL	0	2	1	1	0.3535	0.2524	0.4742	0.4516
	CARCINOMA, CORTICAL	0	0	0	1	0.2199	.	.	0.4516
	C_Pheochr_M+B	**2**	**0**	**3**	**2**	**0.2346**	**1.0000**	**0.4507**	**0.6148**
	GANGLIONEUROMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	PHEOCHROMOCYTOMA	1	0	3	2	0.1293	1.0000	0.2705	0.4269
	PHEOCHROMOCYTOMA, MALIGNANT	1	0	0	0	1.0000	1.0000	1.0000	1.0000
BRAIN	ASTROCYTOMA	1	1	1	1	0.4528	0.7573	0.7317	0.7020
	MENINGIOMA	0	1	0	0	0.7330	0.5049	.	.
	TUMOR, GRANULAR CELL, BENIGN	0	1	1	1	0.2467	0.5096	0.4742	0.4516
DUODENUM	LEIOMYOMA	0	1	0	0	0.7330	0.5049	.	.
HEAD	FIBROSARCOMA	0	0	0	1	0.2199	.	.	0.4516
HEMOPOIETIC SYSTEM	LEUKEMIA, MYELOID CELL	0	0	1	0	0.4607	.	0.4742	.
	SARCOMA, HISTIOCYTIC	1	0	0	0	1.0000	1.0000	1.0000	1.0000
JEJUNUM	LEIOMYOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
KIDNEYS	**C_Lipoma+Re Liposa**	**0**	**0**	**0**	**2**	**0.0475**	**.**	**.**	**0.2013**
	LIPOMA	0	0	0	1	0.2199	.	.	0.4516
	LIPOSARCOMA, RENAL	0	0	0	1	0.2199	.	.	0.4516
LIVER	ADENOMA, HEPATOCELLULAR	2	1	0	0	0.9818	0.8822	1.0000	1.0000
LUNGS AND BRONCHI	ADENOCARCINOMA, BRONCHIOLOALVE	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	ADENOMA, BRONCHIOLOALVEOLAR	0	0	0	1	0.2199	.	.	0.4516
MAMMARY	ADENOCARCINOMA	22	15	21	21	0.1567	0.9302	0.5019	0.4022
	ADENOCARCINOMA ARISING IN FIBR	1	2	2	1	0.5213	0.5074	0.4691	0.7020
	ADENOMA	8	2	0	4	0.6911	0.9913	1.0000	0.8786
	C_Adenocarci+Fibr+Ade	**46**	**45**	**42**	**52**	**0.0305**	**0.6544**	**0.6655**	**0.0727**
	C_Adenomas+Carcinomas	**23**	**16**	**21**	**21**	**0.2158**	**0.9266**	**0.5758**	**0.4744**
	FIBROADENOMA	35	38	34	38	0.1985	0.3976	0.4716	0.2301
Multiple	**C_Lipomas+Liposarcomas**	**2**	**0**	**0**	**2**	**0.2225**	**1.0000**	**1.0000**	**0.6148**

Organ Name	Tumor Name	0 mg|kg|day Vehicle Control N=72	75 mg|kg|day Low N=72	200 mg|kg|day Med N=72	500 mg|kg|day High N=72	P_Value Dose Resp	P_Value VC vs. L	P_Value VC vs. M	P_Value VC vs. H
Organs									
OVARIES	ADENOMA, SERTOLIFORM, TUBULAR	1	1	0	1	0.4938	0.7573	1.0000	0.7020
	TUMOR, GRANULOSA CELL, BENIGN	0	0	1	0	0.4607	.	0.4742	.
PANCREAS	ADENOMA, ISLET CELL	4	5	2	1	0.9272	0.5124	0.8724	0.9548
	CARCINOMA, ISLET CELL	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	C_I_cell_adeno+carci	**4**	**5**	**2**	**1**	**0.9272**	**0.5124**	**0.8724**	**0.9548**
PITUITARY	ADENOMA, PARS DISTALIS	51	49	46	42	0.7178	0.5553	0.7808	0.7477
	CARCINOMA, PARS DISTALIS	**6**	**10**	**10**	**5**	**0.5980**	**0.2326**	**0.1616**	**0.6045**
	C_Dis pars_adeno+carci	**57**	**59**	**56**	**47**	**0.8262**	**0.1740**	**0.4477**	**0.7414**
SKELETAL MUSCLE (O	LIPOSARCOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
SKIN AND SUBCUTIS	CARCINOMA, BASAL CELL	0	0	1	0	0.4607	.	0.4742	.
	CARCINOMA, SQUAMOUS CELL	0	0	1	0	0.4635	.	0.4796	.
	C_Fib+Fib dermoid+Fibrosa	**3**	**5**	**5**	**2**	**0.6615**	**0.3796**	**0.3124**	**0.7528**
	C_Fibroma+Fibr dermoid	**3**	**2**	**3**	**1**	**0.7436**	**0.8247**	**0.6214**	**0.9144**
	FIBROMA	3	2	3	0	0.9180	0.8247	0.6214	1.0000
	FIBROMA (DERMOID)	0	0	0	1	0.2199	.	.	0.4516
	FIBROSARCOMA	0	3	2	1	0.4416	0.1286	0.2274	0.4516
	LIPOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	OSTEOSARCOMA	0	0	0	1	0.2199	.	.	0.4516
	PAPILLOMA, SQUAMOUS CELL	0	0	1	0	0.4607	.	0.4742	.
	TUMOR, BASAL CELL, BENIGN	0	0	0	1	0.2199	.	.	0.4516
SPINAL CORD THORAC	ASTROCYTOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
THORACIC CAVITY	FIBROSARCOMA	0	1	0	0	0.7330	0.5049	.	.
THYROIDS	ADENOMA, C-CELL	3	8	4	2	0.8049	0.1129	0.4419	0.7528
	ADENOMA, C-CELL, GANGLIONEUROM	0	1	0	0	0.7330	0.5049	.	.
	ADENOMA, FOLLICULAR CELL	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	CARCINOMA, C-CELL	2	1	1	0	0.9049	0.8822	0.8588	1.0000
	CARCINOMA, FOLLICULAR	0	1	0	0	0.7330	0.5049	.	.

Organ Name	Tumor Name	0 mg|kg|day Vehicle Control N=72	75 mg|kg|day Low N=72	200 mg|kg|day Med N=72	500 mg|kg|day High N=72	P_Value Dose Resp	P_Value VC vs. L	P_Value VC vs. M	P_Value VC vs. H
	CELL								
	C_C-cell Adenomas+Carcinomas	**5**	**9**	**5**	**2**	**0.9132**	**0.2170**	**0.5620**	**0.9079**
	C_Follicular cell Adenomas+Car	**1**	**1**	**0**	**0**	**0.9297**	**0.7573**	**1.0000**	**1.0000**
UTERINE CERVIX	ADENOCARCINOMA, ENDOMETRIAL	0	0	1	0	0.4635	.	0.4796	.
	POLYP, ENDOMETRIAL	1	0	0	2	0.1265	1.0000	1.0000	0.4356
	SCHWANNOMA, MALIGNANT	0	1	0	0	0.7344	0.5096	.	.
	TUMOR, GRANULAR CELL, BENIGN	1	0	0	0	1.0000	1.0000	1.0000	1.0000
UTERUS	ADENOCARCINOMA, ENDOMETRIAL	1	3	2	2	0.3536	0.3162	0.4610	0.4269
	ADENOMA, ENDOMETRIAL	0	0	1	0	0.4607	.	0.4742	.
	C_Adenomas+Adenocarcinomas	**1**	**3**	**2**	**2**	**0.3536**	**0.3162**	**0.4610**	**0.4269**
	C_Polyp_Gland+Endom	**9**	**7**	**2**	**5**	**0.7670**	**0.7796**	**0.9923**	**0.8277**
	LEIOMYOSARCOMA	0	1	0	0	0.7330	0.5049	.	.
	POLYP, ENDOMETRIAL	8	7	2	4	0.8298	0.6969	0.9857	0.8623
	POLYP, GLANDULAR	1	0	0	1	0.3923	1.0000	1.0000	0.7020
	SCHWANNOMA, MALIGNANT	0	0	1	0	0.4635	.	0.4796	.
VAGINA	PAPILLOMA, SQUAMOUS CELL	0	0	1	0	0.4607	.	0.4742	.
	TUMOR, GRANULAR CELL, BENIGN	1	2	1	0	0.8448	0.5074	0.7317	1.0000

Table 11: Intercurrent Mortality Rate -Male Mice

	0 mg|kg|day		30 mg|kg|day		100 mg|kg|day		300 mg|kg|day	
Week	**No. of Death**	**Cum. %**	**No. of Death**	**Cum. %**	**No. of Death**	**Cum. %**	**No. of Death**	**Cum. %**
0 - 13	1	4.00	.	.	.	.	1	4.00
14 - 26	2	12.00	.	.	.	.	.	.
Ter. Sac.	22	88.00	25	100.00	25	100.00	24	96.00

Cum. %: Cumulative percentage except for Ter. Sac.

Table 12: Intercurrent Mortality Rate -Female Mice

	0 mg|kg|day		30 mg|kg|day		100 mg|kg|day		300 mg|kg|day	
Week	**No. of Death**	**Cum. %**	**No. of Death**	**Cum. %**	**No. of Death**	**Cum. %**	**No. of Death**	**Cum. %**
0 - 13	.	.	.	.	.	.	.	.
14 - 26	1	4.00	.	.	1	4.00	1	4.00
Ter. Sac.	24	96.00	25	100.00	24	96.00	24	96.00

Cum. %: Cumulative percentage except for Ter. Sac.

Table 13: Intercurrent Mortality Comparison between Treated Groups and Vehicle Control -Male Mice

Test	Statistic	P_Value Dose Resp	P_Value VC vs. Low	P_Value VC vs. Med	P_Value VC vs. High
Dose-Response	Likelihood Ratio	0.5646	0.0384	0.0384	0.3061
Homogeneity	Log-Rank	0.1042	0.0770	0.0770	0.3170

Table 14: Intercurrent Mortality Comparison between Treated Groups and Vehicle Control -Female Mice

Test	Statistic	P_Value Dose Resp	P_Value VC vs. Low	P_Value VC vs. Med	P_Value VC vs. High
Dose-Response	Likelihood Ratio	0.7671	0.2390	0.9885	0.9885
Homogeneity	Log-Rank	0.7978	0.3173	0.9885	0.9885

Table 15: Tumor Rates and P-Values for Dose Response Relationship and Pairwise Comparisons -Male Mice

Organ Name	Tumor Name	0 mg|kg|day Vehicle Cont N=25	30 mg|kg|day Low N=25	100 mg|kg|day Med N=25	300 mg|kg|day High N=25	P_Value Dose Resp	P_Value VC vs. L	P_Value VC vs. M	P_Value VC vs. H
BONE, FEMUR INCLUD	HEMANGIOMA	0	0	1	0	0.5052	.	0.5208	.
HARDERIAN GLANDS	ADENOMA	0	0	1	0	0.5052	.	0.5208	.
HEAD	CARCINOMA	1	0	0	0	1.0000	.	.	.

Organ Name	Tumor Name	0 mg|kg|day Vehicle Cont N=25	30 mg|kg|day Low N=25	100 mg|kg|day Med N=25	300 mg|kg|day High N=25	P_Value Dose Resp	P_Value VC vs. L	P_Value VC vs. M	P_Value VC vs. H
KIDNEYS	ADENOMA, TUBULAR	0	0	0	1	0.2474	.	.	0.5106
LIVER	HEMANGIOMA	0	0	1	0	0.5052	.	0.5208	.
LUNGS AND BRONCHI	ADENOCARCINOMA, BRONCHIOLOALVE	1	1	0	0	0.9457	0.7757	1.0000	1.0000
	ADENOMA, BRONCHIOLOALVEOLAR	3	1	1	0	0.9708	0.9545	0.9545	1.0000
	C_alveolar_carci+ade	**4**	**2**	**1**	**0**	**0.9928**	**0.9232**	**0.9803**	**1.0000**
Multiple Organs	**C_hemangiosarcoma_hemangioma**	**2**	**2**	**3**	**1**	**0.7369**	**0.7270**	**0.5408**	**0.8908**
SPLEEN	HEMANGIOMA	2	1	1	0	0.9311	0.8976	0.8976	1.0000
	HEMANGIOSARCOMA	0	1	1	0	0.6342	0.5208	0.5208	.
	C_hemangiosarcoma_hemangioma	**2**	**2**	**2**	**0**	**0.9281**	**0.7270**	**0.7270**	**1.0000**
STOMACH	CARCINOMA, SQUAMOUS CELL	0	0	0	1	0.2474	.	.	0.5106
TESTES	HEMANGIOMA	0	0	0	1	0.2474	.	.	0.5106

Table 16: Tumor Rates and P-Values for Dose Response Relationship and Pairwise Comparisons -Female Mice

Organ Name	Tumor Name	0 mg|kg|day Vehicle Control N=25	30 mg|kg|day Low N=25	100 mg|kg|day Med N=25	300 mg|kg|day High N=25	P_Value Dose Resp	P_Value VC vs. L	P_Value VC vs. M	P_Value VC vs. H
HARDERIAN GLANDS	ADENOCARCINOMA	0	0	1	0	0.5000	.	0.5102	.
	ADENOMA	1	0	0	1	0.4317	1.0000	1.0000	0.7553
Harderian glands	**C_adenocarcinoma+adenoma**	**1**	**0**	**1**	**1**	**0.3882**	**.**	**.**	**.**
LUNGS AND BRONCHI	ADENOCARCINOMA, BRONCHIOLOALVE	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	ADENOMA, BRONCHIOLOALVEOLAR	1	0	2	1	0.3637	1.0000	0.5156	0.7553
LYMPH NODE, MESENT	HEMANGIOSARCOMA	0	0	1	0	0.5000	.	0.5102	.
Lungs and bronchi	**C_Bronchioloalveolar_adenocarc**	**2**	**0**	**2**	**1**	**0.5259**	**.**	**.**	**.**
Multiple Organs	**C_hemangiosarcoma_hemangioma**	**1**	**0**	**1**	**1**	**0.3930**	**1.0000**	**0.7551**	**0.7551**

Organ Name	Tumor Name	0 mg|kg|day Vehicle Control N=25	30 mg|kg|day Low N=25	100 mg|kg|day Med N=25	300 mg|kg|day High N=25	P_Value Dose Resp	P_Value VC vs. L	P_Value VC vs. M	P_Value VC vs. H
SKELETAL MUSCLE	HEMANGIOSARCOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
SPLEEN	HEMANGIOSARCOMA	0	0	0	1	0.2525	.	.	0.5102
STOMACH	PAPILLOMA, SQUAMOUS CELL	0	0	1	0	0.5000	.	0.5102	.
UTERUS	POLYP, ENDOMETRIAL	0	0	0	1	0.2449	.	.	0.5000

Figure 1: Kaplan-Meier Survival Functions for Male Rats

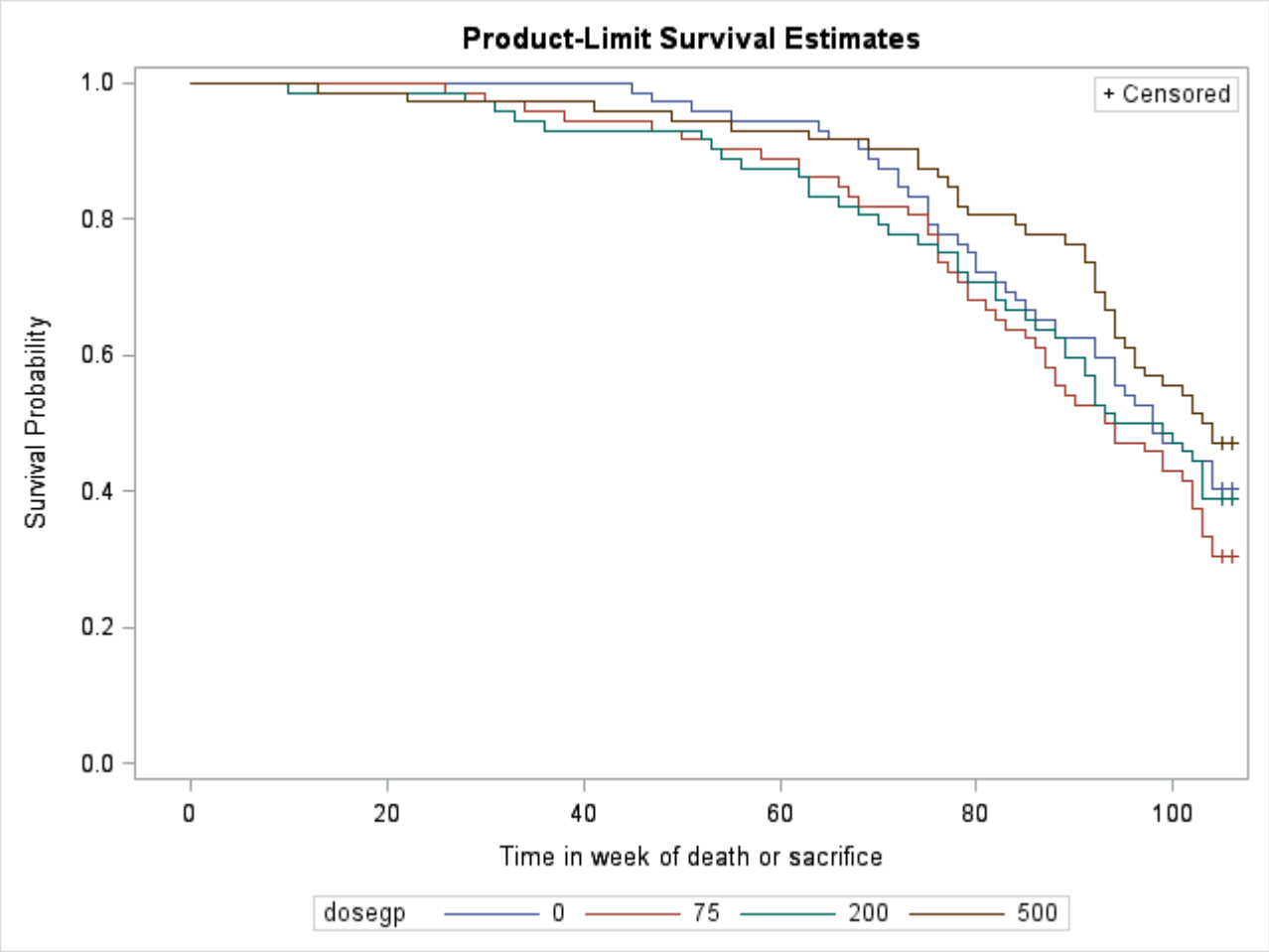


Figure 2: Kaplan-Meier Survival Functions for Female Rats

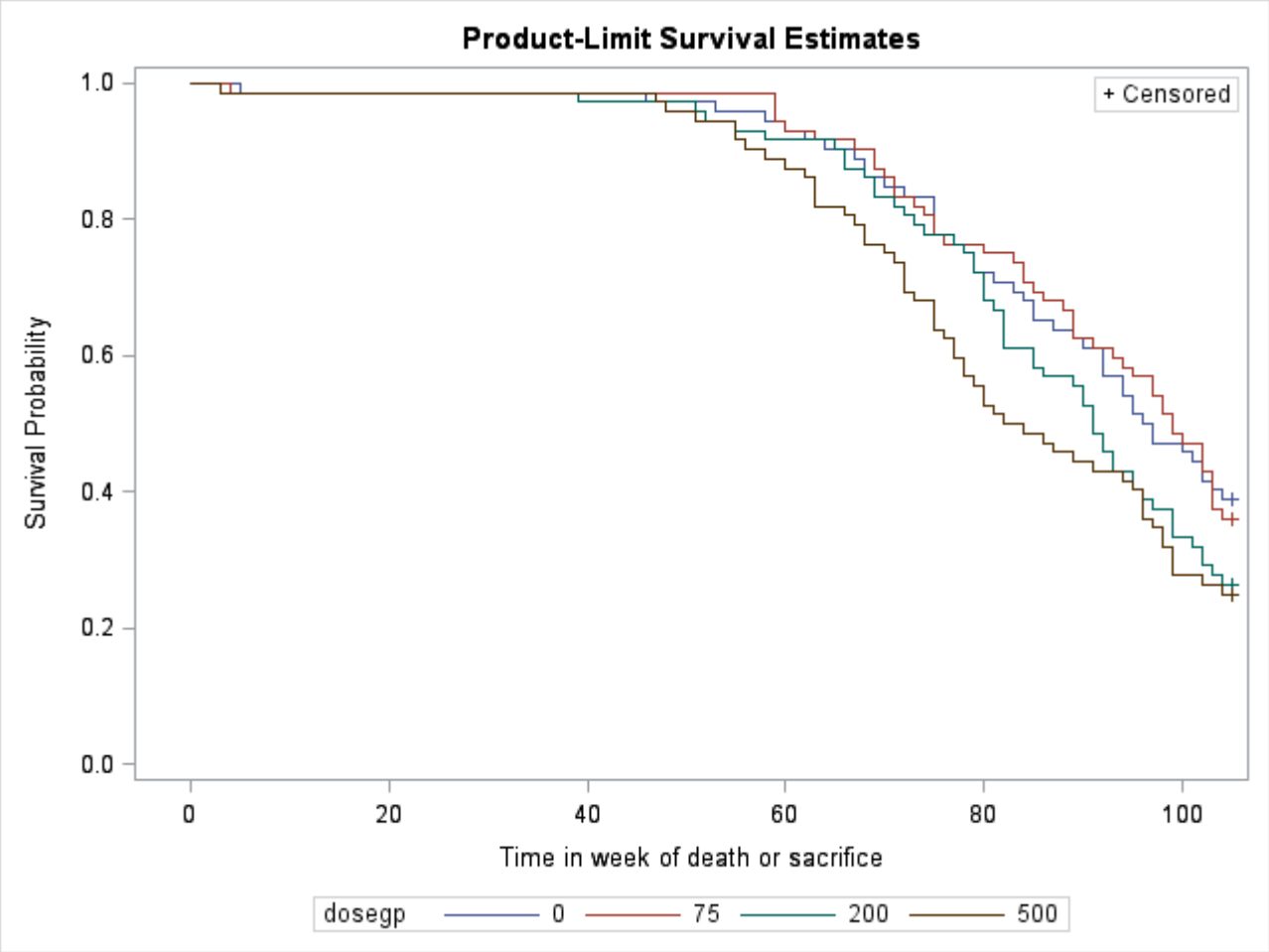


Figure 3: Kaplan-Meier Survival Functions for Male Mice

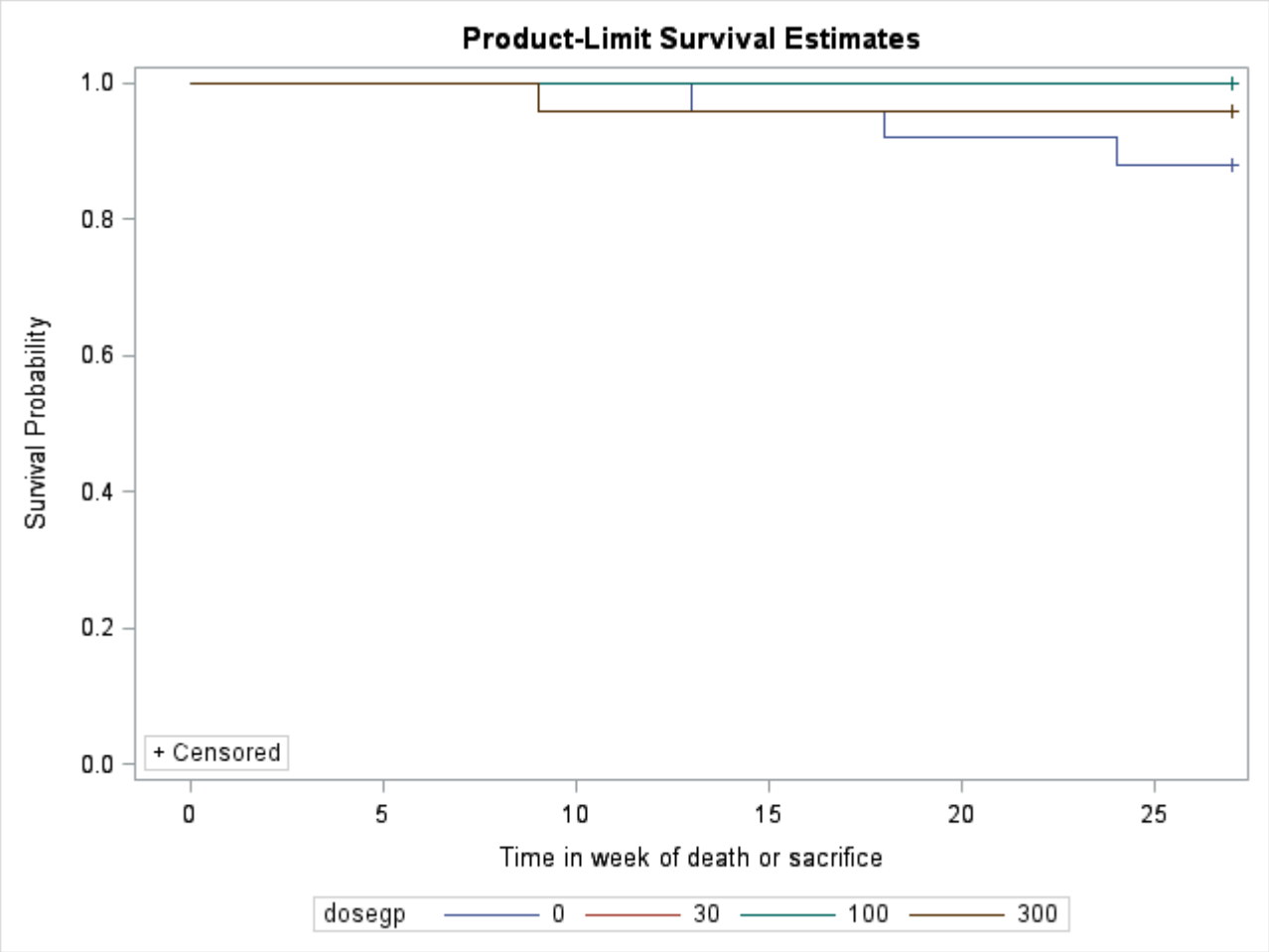
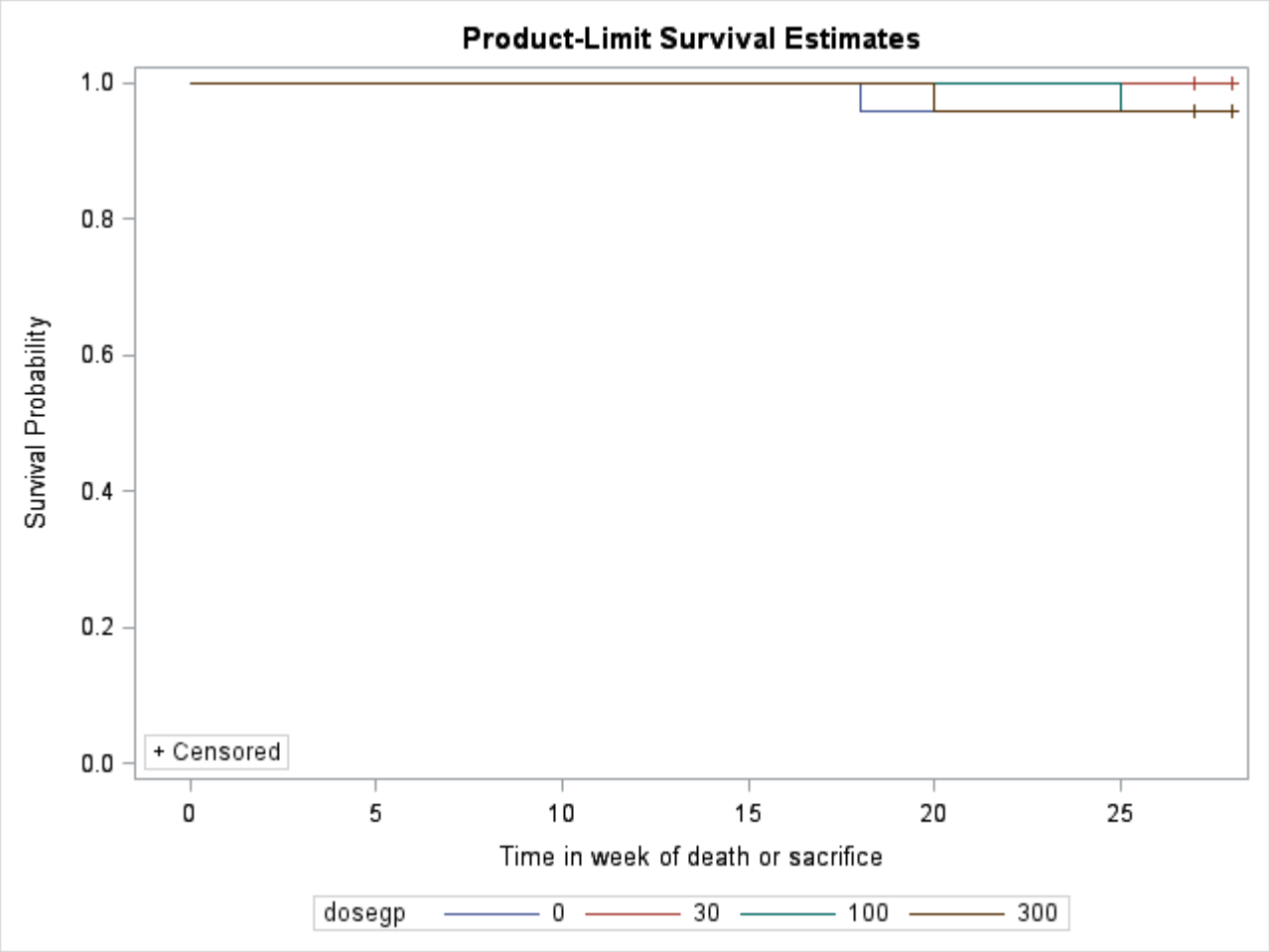


Figure 4: Kaplan-Meier Survival Functions for female Mice



7. References

- Kaplan EL and Meier P (1958) Nonparametric estimation from incomplete observations. J. Am. Statist. Assoc., 53, 457-481.
- Mantel N (1966) Evaluation of survival data and two new rank order statistics arising in its consideration. Cancer Chemotherapy Reports, 50, 163-170.
- Peto R (1974) Guidelines on the analysis of tumour rates and death rates in experimental animals. British J. Cancer, 29, 101-105.
- Lin KK (2000) Carcinogenicity Studies of Pharmaceuticals. In: Encyclopedia of Biopharmaceutical Statistics, ed. Shein-Chung Chow, Marcel Dekker, New York.
- Peto R et al. (1980) Guidelines for simple, sensitive significance tests for carcinogenic effects in long-term animal experiments. In: Long term and Short term Screen Assays for Carcinogens: A Critical Appraisal. IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans. Supplement 2, pp.311-426. WHO International Agency for Research on Cancer, Lyon.
- SAS Institute (2002) SAS OnlineDoc® Version Nine. SAS Institute Inc., Cary, NC, USA.
- Peto, R., M.C. Pike, N.E. Day, R.G. Gray, P.N. Lee, S. Parish, J. Peto, Richards, and J.Wahrendorf, "Guidelines for sample sensitive significance test for carcinogenic effects in long-term animal experiments", Long term and short term screening assays for carcinogens: A critical appraisal, International agency for research against cancer monographs, *Annex to supplement, World Health Organization, Geneva,* 311-426, 1980.
- Bailer AJ, Portier CJ (1988). "Effects of treatment-induced mortality and tumor-induced mortality on tests for carcinogenicity in small samples." *Biometrics*, 44, 417-431.
- Bieler, G. S. and Williams, R. L. (1993). "Ratio estimates, the delta method, and quantal response tests for increased carcinogenicity". *Biometrics* 49, 793-801.
- Tarone RE, "Test for trend in life table analysis", *Biometrika* 1975, 62: 679-82
- Lin K.K. and Rahman M.A.," Overall false positive rates in tests for linear trend in tumor incidence in animal carcinogenicity studies of new drugs", *Journal of Biopharmaceutical Statistics,* 8(1), 1-15, 1998.
- Rahman, A.M., and K.K. Lin (2008), "A Comparison of False Positive Rates of Peto and Poly-3 methods for Long-Term Carcinogenicity Data Analysis Using Multiple Comparison Adjustment Method Suggested by Lin and Rahman", Journal of Biopharmaceutical Statistics, 18:5, 849-858.
- Haseman, J, "A re-examination of false-positive rates for carcinogenesis studies", *Fundamental and Applied Toxicology*, 3: 334-339, 1983.
- Guidance for Industry. Statistical Aspects of the Design, Analysis, and Interpretation of Chronic Rodent Carcinogenicity Statues of Pharmaceuticals (Draft Guidance). U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), May 2001.

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

/s/

ZHUANG MIAO
12/15/2016

KARL K LIN
12/15/2016
Concur with review