CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

204629Orig1s000

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 #:	204629
Drug Name:	Empagliflozin 25 mg tablets
Indication:	Adjunct to diet and exercise to improve glycemic control in adults with type 2 diabetes mellitus
Applicant:	Boehringer Ingelheim Pharmaceuticals, Inc.
Date(s):	5 March 2013 (Submission Date)
	5 March 2014 (PDUFA Goal Date)
Review Priority:	Standard
Biometrics Division:	Division of Biometrics VII (DBVII)
Statistical Reviewer:	Janelle K. Charles, PhD.
Concurring Reviewers:	Mat Soukup, PhD.
	Statistics Team Leader, DBVII
	Mark Levenson, PhD.
	Deputy Division Director, DBVII
Medical Division:	Division of Metabolism and Endocrinology Products (DMEP)
Clinical Team:	William Chong, MD
	Medical Officer, DMEP
	Karen M. Mahoney, MD
	Diabetes Team Leader, DMEP
Project Manager:	Patricia Madara, DMEP

Keywords: type II diabetes, cardiovascular safety, meta-analysis, Cox proportional hazards

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1 EXECUTIVE SUMMARY

This is a statistical safety review of a cardiovascular (CV) meta-analysis report submitted on March 5, 2013 by Boehringer Ingelheim, the Applicant for the New Drug Application (NDA 204629) for empagliflozin tablets. The Applicant is seeking an indication for empagliflozin as an "adjunct to diet and exercise to improve glycemic control in adults with type 2 diabetes mellitus" (T2DM). The development program assessed the efficacy and safety of empagliflozin 10 mg and 25 mg tablets; however, the Applicant is proposing to commercialize the 25 mg dose only. The primary objective of the meta-analysis was to demonstrate that the upper bound of the confidence interval for the hazard ratio (empagliflozin to comparators) was smaller than the pre-market risk margin of 1.8 as stipulated in the FDA Diabetes Guidance for assessing CV safety in new anti-diabetic products.

The CV meta-analysis included 8 Phase II and III trials in the empagliflozin development program, including data from an ongoing dedicated CV outcomes trial, trial 1245.25. In accordance with the FDA Guidance recommendations the trials included in the meta-analysis enrolled patients with advanced CV disease, elderly patients, and patients with some degree of renal impairment. The meta-analysis was conducted in accordance with the meta-analysis plan that was finalized July 2012 and agreed upon with the FDA. There were some trials that were complete before finalization of the meta-analysis plan; see Section 3.2.1.1 for trial completion dates. The agreed upon population of interest for the meta-analysis comprised all randomized patients who received at least one dose of study medication. The main groups compared were empagliflozin (pooled 10 mg and 25 mg doses) and comparator (pooled placebo and active drugs). The meta-analysis was not designed to assess CV safety of the individual empagliflozin doses. The primary safety endpoint of this meta-analysis was **MACE+**, a composite endpoint comprising CV death, non-fatal myocardial infarction (MI), non-fatal stroke, or hospitalization for unstable angina. A key secondary endpoint was **MACE**, a composite endpoint comprising CV death, non-fatal MI, or non-fatal stoke. All events included in the meta-analysis were based on positively adjudicated events determined by a blinded Clinical Event Committee that used standardized definitions. The pre-specified primary statistical analysis used a Cox proportional hazards model, stratified by trial.

In addition to the evaluation of the pre-marketing risk margin of 1.8, the Applicant pre-specified an interim meta-analysis to test the post-marketing risk margin of 1.3 for MACE+ and MACE at the time point that the 1.8 risk margin was ruled out[1]. Additionally, at this time point, the Applicant pre-specified to use the interim data from trial 1245.25 alone to assess the 1.3 risk margin. While the Type I error is controlled separately within each data source (i.e. the meta-analysis and trial 1245.25 alone), in April 2012[2] the Agency communicated that using both data sources to evaluate the 1.3 risk margin does not control the Type I error across the two data sets.

[1] Note that the final meta-analysis to evaluate the post-marketing risk margin of 1.3 was to occur after the completion of trial 1245.25 which was ongoing at the time of NDA submission. In addition, trial 1245.25 was designed to evaluate and rule out the 1.3 risk margin on its own (i.e. separate from the meta-analysis of the other trials).
[2] Refer to preliminary meeting responses dated April 9, 2012.

No changes to the evaluation of the 1.3 risk margin were made that address this multiplicity issue across the two data sources. Since the multiplicity across the two data sources has not been addressed, we recommend that conclusions about the CV safety of empagliflozin with respect to the 1.3 risk margin be based on findings from the dedicated cardiovascular outcome trial alone, trial 1245.25, as this is consistent with advice provided to other sponsors of products intended to treat Type 2 diabetes. The findings from the interim analysis of trial 1245.25 alone for assessing the 1.3 risk are discussed in this review.

For the meta-analysis, a total of 6206 patients were randomized to an empagliflozin dose of which (b) (4) experienced a MACE+ event, whereas a total of 3830 patients were randomized to a comparator of which (b) (4) experienced a MACE+ event. More than 70% of the MACE+ included in the meta-analysis occurred in trial 1245.25; specifically (b) (4) in empagliflozin patients and (b) (4) in placebo patients. The estimated hazard ratio for MACE+ across all trials included in the meta-analysis was (b) (4) with 95% CI ((b) (4)); the meta-analysis of MACE yielded consistent results, see Table 1. (b) (4) Analyses were also conducted to compare the risk of MACE+ within each empagliflozin dose relative to all comparators. For the 10 mg dose, the HR estimate of MACE+ was (b) (4) with 95% CI (b) (4)). For the 25 mg dose, which is the proposed marketing dose of empagliflozin, the HR estimate was (b) (4) with 95% CI (b) (4) Thus, it can be concluded that the pre-specified meta-analysis to evaluate cardiovascular safety ruled out the pre-marketing risk margin of 1.8.

Table 1 Summary of Meta-analysis Results of MACE+ and MACE for 1.8 Risk Margin

Outcome	Number of Patients with Events		HR
	Empagliflozin[1] (N=6206)	Comparator[2] (N=3830)	(95% CI)
MACE+	(b) (4)		
MACE			

[1]Pooled 10 mg and 25 mg empagliflozin doses
[2]Pooled active and placebo comparators
Source: Created by the reviewer using dataset "adttecv.xpt"

The primary endpoint in trial 1245.25 was MACE, as defined previously. The trial followed a hierarchical testing strategy as outlined in Appendix II of this review and the Type I error was controlled at the one-sided α=0.025 level using the Haybittle-Peto method to account for the interim analysis that was planned to occur at the time of the evaluation of the 1.8 pre-marketing risk margin (α=0.0001 at the interim analysis and α=0.0249 for the final analysis). Based on this alpha spending function, the evaluation of the 1.3 post-marketing risk margin using the interim data utilizes a two-sided 99.98% confidence level. Using interim data from trial 1245.25, (b) (4)

MACE were reported in 3046 patients randomized to empagliflozin ((b)(4)%), and (b)(4) MACE were reported in 1513 patients randomized to placebo ((b)(4)%). The estimated HR for MACE was (b)(4) with corresponding 99.98% CI (b)(4)). Therefore, based on the interim results from trial 1245.25, (b)(4) Note that the trial is still ongoing with the final evaluation planned to occur after (b)(4) MACE have been positively adjudicated[3].

2 INTRODUCTION

2.1 Product Description and Regulatory Background

Empagliflozin is a novel, orally available, potent selective inhibitor of SGLT-2 that decreases renal reabsorption of glucose and thereby increases the urinary glucose excretion and lowers plasma glucose levels. The Applicant, Boehringer Ingelheim (BI), submitted a New Drug Application, NDA 204629, for empagliflozin tablets on March 5, 2013 (PDUFA Goal Date: March 5, 2014) to be indicated as an "adjunct to diet and exercise to improve glycemic control in adults with type 2 diabetes mellitus" (T2DM). The development program studied efficacy and safety of empagliflozin 10 mg and 25 mg tablets; however, BI is proposing to commercialize the 25 mg dose only. The NDA submission strategy and development program for empagliflozin tablets were discussed with the Division of Metabolism and Endocrinology Products (DMEP) and reflected in End of Phase II Advice (December 8, 2009), End of Phase II Meeting Minutes (June 3, 2010), Written Responses to Phase III analysis plans (May 22, 2012) and pre-NDA Meeting Minutes (December 17, 2012).

In accordance with the FDA Guidance for Industry *Diabetes Mellitus-Evaluating Cardiovascular Risk in New Antidiabetic Therapies to Treat Type 2 Diabetes* (issued December 2008), the Applicant performed a meta-analysis of clinical trials in their development program to assess whether empagliflozin is associated with an unacceptable increased risk of cardiovascular (CV) disease. In alignment with the guidance, the pre-marketing objective of the meta-analysis was to rule out an excess risk of 80%, as demonstrated by an upper bound of the two-sided 95% confidence interval (CI) less than 1.8 for the hazard ratio. The CV meta-analysis that was designed to rule out the 1.8 risk margin is the subject of this statistical safety review. The analysis plan for assessing the 1.8 risk margin, finalized July 2012, was reviewed and agreed upon by the FDA[4]. All CV events were prospectively adjudicated by an independent Clinical Event Committee (CEC) for inclusion in the statistical analyses. The CEC was blinded to treatment allocation information for all patients. Interim data from an ongoing dedicated CV outcomes trial, 1245.25, designed to rule out the 1.3 post-marketing risk margin was included in the CV meta-analysis and will also be evaluated in this review. The interim analysis for this trial

[3] Note that the Agency advised the sponsor on June 24, 2011 that the proposed interim analysis approach to evaluate the 1.3 risk margin was conservative as a large number of events would accrue between the interim and final analysis. It was advised that the sponsor consider adding an interim analysis prior to the final analysis with modifications to the alpha spending function.
[4] Refer to statistical review by Dr. Todd Sahlroot (DBII) signed off in DARRTS on August 31, 2012.

and the CV meta-analysis were performed by a team independent from the 1245.25 trial team, so that the trial team remains blinded to the results of the ongoing trial. A discussion regarding the firewall procedures implemented by the Applicant to protect the integrity of the trial data is provided in Section 3.1.

Note: The Applicant had planned to submit an interim CV meta-analysis report when the NDA was submitted. However, at the database lock for the meta-analysis, the number of patients with CV events exceeded the number that had been planned for the final meta-analysis to rule out 1.8. Therefore, the submitted report was considered the final meta-analysis report to evaluate the 1.8 risk margin.

2.2 Clinical Trial Overview

The CV meta-analysis was based on data from completed randomized double-blind phase II and phase III trials with treatment duration at least 12 weeks as well as data from randomized double-blind phase III trials that were ongoing at the time of the CV meta-analysis but had pre-planned interim analysis, see Table 2 and Table 3 extracted from the study report. All trials included in the meta-analysis were conducted in compliance with their respective trial protocols. There were three completed phase III trials included (1245.19, 1245.20, 1245.23) that were each of duration 24 weeks. These trials were similar with respect to procedures and basic design (double-blind, placebo-controlled, parallel-group) but differed in patient population and background therapies. Trial 1245.23 was considered 2 separate pivotal trials (one trial with metformin background therapy and the other trial with metformin plus sulfonylurea background therapy) under one study number. The other two completed trials (1245.33 and 1245.36) were each at least 52 weeks in duration.

There were two ongoing trials that were included in the meta-analysis at the time of the NDA submission. Trial 1245.25 was an event-driven placebo-controlled cardiovascular outcomes trial of patients at increased risk of CV disease and trial 1245.28 was a phase III trial comparing empagliflozin to glimepiride. Section 3.2.1 provides detailed discussion of the study designs and respective patient populations.

Table 2 Completed Phase II and Phase III Trials Included in CV Meta-analysis

Trial ID Clinical phase Reference	Objective / Trial design	Parallel treatment groups (excluding open label arms)/ Treatment duration	Number of randomised patients (excluding open label arms)
1245.19 Phase III [U12-1516-02]	Empa with pioglitazone background therapy Randomised, double-blind, placebo-controlled	Placebo Empa 10 mg Empa 25 mg 24 weeks	499
1245.20 Phase III [U12-1517-01]	Empa as monotherapy Randomised, double-blind, placebo- and active-controlled, uncontrolled open-label group for poorly controlled patients (HbA_{1c} >10%)	Placebo Empa 10 mg Empa 25 mg Sitagliptin 100 mg 24 weeks	899
1245.23 Phase III [U12-1518-01]	Empa with metformin or metformin plus sulfonylurea background therapy (two separate studies). Randomised, double-blind, placebo-controlled; uncontrolled open-label group for poorly controlled patients (HbA_{1c} >10%)	Placebo Empa 10 mg Empa 25 mg 24 weeks	638 (metformin) 669 (metformin plus sulfonylurea)
1245.36 Phase III	Empa in patients with various degrees of renal impairment Randomised, double-blind, placebo-controlled	Placebo Empa 10 mg Empa 25 mg 52 weeks	741
1245.33 Phase II	Empa with basal insulin ± metformin ± sulfonylurea background therapy Randomised, double-blind, placebo-controlled	Placebo Empa 10 mg Empa 25 mg 78 weeks	494

Source: Applicant's meta-analysis study report Table 6.2.1:1 (page 26)

Table 3 Ongoing Phase III and Extension Trial Included in the CV Meta-analysis

Trial ID Clinial phase Reference	Title	Parallel treatment groups / Treatment duration	Number of randomised patients
1245.25 Phase III -	Cardiovascular safety/outcome trial Randomised, double-blind, placebo-controlled	Placebo Empa 10 mg Empa 25 mg Event-driven (estimated: 8 years)	7000 planned to be randomised overall. 4500 patients planned for the interim analysis
1245.28 Phase III (interim CTR)	Empa versus glimepiride; with metformin background therapy Randomised, double-blind, active-controlled	Empa 25 mg Glimepiride (maximal tolerated dose between 1 mg and 4 mg) 208 weeks	1549
1245.31 Phase III [U12-1521-01] (interim CTR)	Long-term safety and efficacy of empa [extension of 1245.19, 20, and 23] Randomised, double-blind, placebo- and active-controlled	Placebo; Empa 10 mg Empa 25 mg Sitagliptin 100 mg Overall: minimum of 76 weeks (incl. 24 weeks of preceding trial) The interim analysis was performed after the DBL of the last preceding trial (1245.19)	Continued in extension trial: 305 (1245.19/pioglitazone) 615 (1245.20/drug-naive) 463 (1245.23/metformin) 473 (1245.23/metformin plus sulfonylurea)

Source: Applicant's meta-analysis report Table 6.2.1:2 (page 27)

Note that 1245.31 was the extension phase of completed phase III trials 1245.19, 1245.20, and 1245.23. Data from the extension was incorporated into the analyses of the respective main trials; there was no separate analysis of the extension trial only. The data cut-off dates for trials 1245.25, 1245.28, and, 1245.31 were June 22, 2012, July 31, 2012 and May 29, 2012 respectively.

2.3 Data Sources

The NDA was submitted electronically and included individual datasets for each of the trials included in the CV meta-analysis. All data tabulation datasets were provided in CDISC Study Data Tabulation Model and all analysis datasets were provided in CDISC Analysis Dataset Model format, except for those analyses datasets noted in the submitted Reviewer's Guide. This Reviewer's Guide provides clarification on the structure of the information provided in the application as well as information about dataset formats. Data definition files for the respective datasets were also included in the application.

In the original submission, the Applicant submitted time to CV event analysis datasets for each of the individual trials included in the meta-analysis, but an integrated CV dataset that compiled information from all trials to be utilized in the meta-analysis was not submitted. Due to the formats of the individual trial data and the extent of the number of trials, compilation of the individual trial data into an integrated database would require significant resources that would not be subject to the same quality control measures implemented by the Applicant. Therefore, on March 25, 2013, integrated datasets that were to be structured based on specified format preferences were requested to be submitted by the Applicant. In response, the integrated datasets were submitted to the application on April 12, 2013 and can be found at

EDR location: \\Cdsesub1\evsprod\NDA204629\0001\m5\datasets.

The following integrated datasets were used to perform statistical analyses in this review:

- "adsl.xpt" which contains the demographic and disposition data
- "adceccv.xpt" which contains the cardiovascular adjudication results
- "adttecv.xpt" which contains the time to event analysis variables.

A discussion of the data quality is provided in Section 3.1 of this review.

3 STATISTICAL EVALUATION

This is a statistical safety review that focuses on the cardiovascular safety meta-analysis for empagliflozin. An assessment of CV safety based on interim data from the dedicated CV outcomes trial, trial 1245.25, is also provided. Please refer to separate statistical review by Dr. Dongmei Liu for overall efficacy and safety evaluation.

3.1 Data and Analysis Quality

The CV meta-analysis to evaluate the pre-marketing risk margin of 1.8 was conducted in accordance with the meta-analysis plan, which was reviewed and agreed upon with the FDA. The Applicant stated in the study report that all trials included in the CV meta-analysis were conducted in compliance with their respective protocols, in accordance with ICH Good Clinical Practice, applicable regulatory requirements, relevant local guidelines, and standard operating procedures of BI.

The integrated time to event analysis dataset, "adttecv.xpt", provided the censoring variables and time to CV event or censoring so that the reviewer was able to replicate the Applicant's primary and secondary analyses results using the variables coded in this dataset. However, the specific event date or censoring date was not provided in the dataset. Therefore, the reviewer had difficulties in deriving the time at risk needed to estimate the hazard ratios for CV events. For patients with positively adjudicated events, the reviewer was able to recover the event dates from the CEC dataset, "adceccv.xpt", and the resulting time to event matched the value presented in the Applicant's integrated time to event analysis dataset. For the censored patients, the reviewer attempted to use the Applicant's derivation rule provided in the individual trial Reviewer's Guide to determine the date of censoring. According to this rule, the censoring date was the earliest of the end of observation for the trial and the last recorded date obtained from approximately ten datasets (including labs and exposure datasets). Given the complexity of code needed to derive these dates from several datasets for each of the multiple trials included in the meta-analysis, the reviewer performed calculations by hand for 12 censored patients from a variety of trials. The time to censoring when based on dates obtained by the reviewer using the derivation rule was consistent with the Applicant's integrated time to event analysis dataset for the majority of patients examined. As a result of this and given that the reviewer was able to reproduce the CV event dates from the submitted CEC dataset, the reviewer is reassured that the Applicant's time to event or censoring as reported in the integrated time to event analysis dataset is accurate.

Interim data from a dedicated ongoing CV outcomes trial, trial 1245.25, was included in the CV meta-analysis. Concerns were raised during this review about the firewall procedures that were implemented by the Applicant to ensure the integrity of the data from this trial. In response to information request (IR) sent on October 8, 2013, the Applicant noted that several measures, such as confidentiality agreements and restricted access to electronic systems, were taken to guarantee the confidentiality of the interim results until trial completion. These measures were described in the "Interim Analyses Logistics Plan" (finalized August 23rd, 2012); a draft of this final plan was agreed upon with the FDA (reference written responses to meeting request dated April 9, 2012). The plan also includes a list of personnel who were obliged to sign the confidentiality agreement.

<u>Reviewer Comment</u>: The firewall procedures described in the final plan were similar to that previously agreed upon with the FDA and appear to be generally acceptable. However, based

on information provided by the Applicant in response to the IR sent on October 8, 2013, there is a concern regarding the fact that the data was accessed by two representatives from Eli Lilly, the Applicant's Alliance partner. The Applicant states in the IR response that these two representatives reviewed the interim results from trial 1245.25 at the BI site "on a standalone computer without the option to create an electronics copy or printout of the results". It should be noted that an inability to print does not imply that results are not shared, especially top-line results of the primary analysis. Therefore, there is uncertainty based on the IR response that these two Eli Lilly representatives signed any confidentiality agreement to ensure the preservation of the trial interim results.

3.2 Evaluation of Safety

3.2.1 Design of Trials Included in Meta-Analysis

All trials in the meta-analysis included adults (at least 20 years for patients in Japan and at least 18 years otherwise) who had been diagnosed with T2DM, had insufficient glycemic control, and BMI $\leq 45 kg/m^2$. All trials studied the 10 mg and 25 mg doses of empagliflozin. All trials were double-blind and in most trials, patients were randomized in a 1:1:1 (empagliflozin 10 mg: empagliflozin 25 mg: comparator) ratio. Specific design summaries for each of the trials included in the meta-analysis (completed and ongoing) that were obtained from the respective trial protocols and study reports are provided in the subsections that follow.

Patients were not eligible to enroll in any trial if any of the following criteria applied:

1. Impaired hepatic function, defined as serum levels of alanine transaminase, aspartate transaminase, or alkaline phosphatase above 3 X the upper limit of normal at screening or during the run-in phase.

2. Treatment with anti-obesity drugs within 3 months prior to informed consent.

3. Treatment with systemic steroids at the time of informed consent or change in dosage of thyroid hormones within 6 weeks prior to informed consent.

4. Drug or alcohol dependent or intolerant of any ingredient of the trial medications or the background therapy.

Female patients could not participate in any trial if they were pregnant or breast feeding or if they did not use adequate contraceptive methods.

3.2.1.1 Designs of Completed Trials Included in Meta-analysis

1245.19: A phase III, randomized, double-blind, placebo-controlled parallel group efficacy and safety trial of empagliflozin (10 mg or 25 mg) administered orally once daily over 24 weeks in patients with T2DM with insufficient glycemic control despite a background therapy of

pioglitazone alone or in combination with metformin. Patients were eligible to enroll if they were at least18 years with T2DM with insufficient glycemic control (7.0 % ≤ HbA1c ≤ 10.0%). The primary endpoint of the trial is the HbA1c change from baseline after 24 weeks. All eligible patients underwent a 2-week open-label placebo run-in period before randomization. Patients who successfully completed this period and who still met the trial specified inclusion/exclusion criteria were randomized to either 1 of the 2 doses of empagliflozin or placebo in addition to pioglitazone alone or in combination with metformin for a 24-week treatment period. Of the 762 patients screened, 530 patients started the placebo run-in period. Overall, 499 patients were randomized to double-blind treatment as follows: 166 patients to placebo, 165 patients to10 mg empagliflozin, and 168 patients to 25 mg empagliflozin (approximately 70 patients per background therapy for each randomized group). Following randomization, there were four study visits, one every 6 weeks, during the 24-week double blind treatment period. Patients who completed the planned 24-week treatment period in this trial were eligible to continue their randomized treatment by enrolling in an extension phase (1245.31). Patients who did not enter the extension were to be followed up for 1 week. The trial was conducted between October 12, 2010 and April 11, 2012.

1245.20: A phase III randomized, double-blind, active (100 mg sitagliptin) and placebo-controlled, parallel group, efficacy and safety study of empagliflozin (10 mg or 25 mg) administered orally over 24 weeks, in drug naïve patients with T2DM and insufficient glycemic control despite diet and exercise. There was no background therapy administered in this trial. A triple-dummy design was followed, that is, each patient was randomized to 1 active treatment and 2 placebos matching the alternative active treatments or 3 placebos matching each of the possible active treatments. The primary endpoint was the change from baseline in HbA1c after 24 weeks of treatment. Eligible patients were drug-naïve with T2DM, insufficient glycemic control (7.0 % ≤ HbA1c ≤ 10.0%, Germany: 7.0 % ≤ HbA1c ≤ 9.0%) and age at least 18 years (Japan: ≥20 years, India: ≥18 years and ≤65 years). Patients with an HbA1c exceeding 10% and fulfilling all remaining inclusion criteria were eligible for inclusion in the empagliflozin 25 mg open-label arm. All eligible patients with HbA1c ≤ 10% underwent a 2-week open-label placebo run-in period before randomization; there was no placebo run-in for eligible patients with HbA1c exceeding 10%. Of the 1616 patients screened, 986 patients were entered and all were treated: 899 patients with double-blind trial medication (after the placebo run-in period) and 87 patients with open-label medication. The 899 patients were randomized to 10 mg empagliflozin (224 patients), 25 mg empagliflozin (224 patients), placebo (228 patients), or 100 mg sitagliptin (223 patients). Following randomization, there were four study visits, one every 6 weeks, during the 24-week double blind treatment period. Patients who completed the planned 24-week treatment period in this trial were eligible to continue their randomized treatment by enrolling in an extension trial (1245.31). Patients who did not enter the extension trial were followed up for 1 week. The trial was conducted from August 12, 2010 to March 19, 2012.

1245.23: A phase III randomized, multi-national, double-blind, placebo-controlled, parallel group, efficacy and safety study of empagliflozin (10 mg or 25 mg) administered orally, once daily over 24 weeks in patients with T2DM with insufficient glycemic control despite treatment with metformin alone or metformin in combination with a sulfonylurea. Eligible patients were at

least 18 years old with 7.0 % ≤ HbA1c ≤ 10.0% for the double blind randomized treatments and HbA1c > 10.0% for the open-label treatment arm. The primary endpoint was the change from baseline in HbA1c after 24 weeks. This trial comprised 2 separate sub-studies based on the background medication the patients were taking (metformin only or metformin plus sulfonylurea). Patients suitable after screening and with 7.0 % ≤ HbA1c ≤ 10.0% underwent a 2 week open-label placebo run-in period before randomization. Patients who successfully completed this period and who still met the inclusion/exclusion criteria were randomized to either 1 of the 2 doses of empagliflozin or placebo in addition to a background therapy of metformin or a combination of metformin with sulfonylurea for a 24-week treatment period. Following randomization, there were four study visits, one every 6 weeks, during the 24-week double blind treatment period. Patients with HbA1c values >10% at screening were not eligible for randomization but were offered participation in the open-label treatment group (in either of the 2 sub-studies) in addition to a background therapy of metformin or a combination of metformin with sulfonylurea. There was no placebo run-in period for these patients. Of the 2256 screened patients, 172 were assigned to open-label treatment (69 patients in the metformin sub-study, also known as sub-study A, and 103 patients in the metformin plus sulphonylurea sub-study, also known as sub-study B). From the remaining screened patients, 1454 started placebo run-in period, of which 1307 patients were randomized and distributed equally to the 3 treatment groups: placebo (207 patients in sub-study A and 225 in sub-study B), 10 mg empagliflozin (217 patients in sub-study A and 226 patients in sub-study B), and 25 mg empagliflozin (214 patients in sub-study A and 218 patients in sub-study B). Patients who completed the planned 24-week treatment period in this trial were eligible to continue their randomized treatment by enrolling in an extension (1245.31). Patients who did not enter the extension were followed up for 1 week. The trial was conducted between July 29, 2010 and February 3, 2012.

1245.33: A phase IIb, randomized, multi-national, double-blind, placebo-controlled, parallel group, safety and efficacy study of empagliflozin (10 mg or 25 mg) administered orally, once daily in T2DM patients receiving treatment with basal insulin (glargine, detemir, or NPH insulin only) with or without concomitant metformin and/or sulfonylurea therapy and insufficient glycemic control. The treatment period was 78 weeks after an open-label placebo run-in period of 2 weeks. Eligible patients were at least 18 years old with 7.0 % ≤ HbA1c ≤ 10.0% for the double blind randomized treatments and HbA1c > 11.0% for the open label treatment arm. The primary endpoint was the change from baseline in HbA1c after 18 weeks; secondary endpoints were evaluated after 78 weeks. Eligible patients underwent a 2 week open-label placebo run-in period before randomization. Following randomization, there were three study visits, one every 6 weeks during the first 18 weeks of the double blind treatment period, thereafter visits were scheduled every 12 weeks until week 78. Of the patients screened, 532 patients started the placebo run-in period. Overall, 494 patients were randomized to double-blind treatment with either empagliflozin 10 mg (169 patients), empagliflozin 25 mg (155 patients), or placebo (170 patients). There was a four-week follow-up after the double-blind treatment period. The trial was conducted between November 11, 2009 and May 9, 2012.

1245.36: A phase III, randomized, double-blind, placebo-controlled, parallel group efficacy and safety study of empagliflozin (10 mg and 25 mg) administered orally once daily as add-on

therapy to any pre-existing antidiabetic therapy, excluding SGLT-2 inhibitors, in patients with T2DM with renal impairment. The treatment period was 52 weeks following a two-week open-label placebo run-in period. Randomization was stratified by HbA1c (<8.5% or ≥8.5%), background therapy (insulin/combination; sulfonylurea/glinide as monotherapy; other) and level of renal impairment (mild; moderate; severe) based on estimated Glomerular Filtration Rate (eGFR). The primary endpoint was change from baseline in HbA1c after 24 weeks; other efficacy endpoints assessed change from baseline after 52 weeks of treatment. A total of 1317 patients were screened, of which 826 patients started the placebo run-in period. Overall, 741 patients were randomized to double-blind treatment: 98 patients to 10 mg empagliflozin, 322 patients to 25 mg empagliflozin, and 321 patients to placebo. Note that the 10 mg dose of empagliflozin was administered to patients with mild renal impairment (60≤eGFR<90) only, thereby accounting for the fewer patients in this treatment arm. There was a three-week follow-up after the double-blind treatment period. The trial was conducted between November 11, 2009 and May 9, 2012.

Note that in the two trials (1245.20 and 1245.23), patients with HbA_{1c} >10.0% at screening who participated in the open-label treatment arm were not included in this meta-analysis.

3.2.1.2 Designs of Ongoing Trials Included in Meta-analysis

1245.25: An ongoing phase III, multicenter, multinational, randomized, parallel group, double-blind cardiovascular safety trial of empagliflozin (10 mg or 25 mg) administered orally once daily compared to placebo as add-on to standard of care in T2DM patients with increased cardiovascular risk. The primary objective of this event driven trial is to rule out an excess CV risk of 1.3 based on the three-point composite major adverse cardiovascular events (MACE) endpoint: CV death, nonfatal stroke, or nonfatal myocardial infarction (MI). The 4-point composite endpoint comprising cardiovascular death, nonfatal stroke, nonfatal MI, or hospitalization for unstable angina (MACE+) is considered a key secondary endpoint. The trial is performed in patients with T2DM and high cardiovascular risk[5] who have insufficient glycemic control despite diet and exercise and are either treatment naïve or receiving any antidiabetic background therapy. Drug-naïve is defined as absence of any antidiabetic therapy for 12 weeks prior to randomization and insufficient glycemic control as 7.0 % ≤ HbA1c ≤ 10.0% at Visit 1 (screening) for patients on background therapy or 7.0% ≤ HbA1c ≤ 9.0% at Visit 1 (screening) for drug-naïve patients. All patients suitable after screening were to undergo a 2 week open-label placebo run-in period before randomization. Patients who successfully complete the run-in period and who still met the inclusion/exclusion criteria were to be randomized to the double-blind treatment period of the trial in which they will receive either 1 of the 2 doses of empagliflozin or placebo in addition to the background therapy (if applicable) they are receiving at the time they sign the informed consent. Randomization was to be stratified by baseline BMI

[5] High CV risk defined as at least one of the following: confirmed history of MI, evidence of multivessel coronary artery disease, evidence of single vessel coronary artery disease, last episode of unstable angina > 2 months prior to informed consent with confirmed evidence of coronary multivessel or single vessel disease, history of ischemic or hemorrhagic stroke, or presence of peripheral artery disease. Refer to inclusion criterion #7 (trial 1245.25 protocol) for detailed definition of high CV risk.

and HbA1c, geographical regions, and renal function. A total of 7000 patients with T2DM who meet the entry criteria are planned for inclusion in this trial. Randomized patients were to return to clinic for regularly scheduled follow up visits 4, 8, 12, 16, 28, 40, and 52 weeks after randomization during the first year of trial participation, and every 14 weeks thereafter for the duration of the trial. As of the planned data cut-off date of June 22, 2012 for this trial, there were 4559 patients randomized: 1521 in the 10 mg empagliflozin arm, 1525 in the 25 mg empagliflozin arm, and 1513 in the placebo arm. The planned treatment duration of the trial will be between 6 and 8 years (approximately 300 to 420 weeks); because this is an event driven trial, specific trial duration was not specified.

***Reviewer's Comment**: Of all the trials included in the meta-analysis to evaluate the 1.8 pre-marketing risk margin, trial 1245.25 is the only trial designed and powered to assess cardiovascular safety of empagliflozin on its own. The trial was planned to have one interim analysis in order to support NDA submission after 80 confirmed primary events had been adjudicated or based on the cut-off date of July 15, 2012, whichever occurred first. The actual cut-off date used was June 22, 2012 at which time there were more than the planned 80 adjudicated events. The Applicant explains that this is due to the lag time between reporting of potential CV events and the completion of the adjudication process, leading to some adjudicated events only becoming apparent at time of cut-off.*

Trial 1245.25 was powered to rule out the post-marketing risk margin of 1.3 and had one planned interim analysis to occur at the time of the assessment of the 1.8 pre-marketing risk margin. To account for this interim look at the data, an alpha adjustment (for the 1.3 risk margin) was pre-specified to be based on the Haybittle-Peto method. According to the protocol, the trial will be stopped after 691 patients have experienced an adjudicated MACE. With this many events, the Applicant estimated that the trial will have 90% power to rule the 1.3 post-marketing risk margin. There were no plans to stop the trial early due to interim results** (b) (4) ***see Appendix II for hierarchical testing strategy.

1245.28: An ongoing phase III randomized, double-blind, active-controlled parallel group efficacy and safety trial of 25 mg empagliflozin compared to glimepiride (1 to 4mg) administered orally during 104 weeks with a 104-week extension period in patients with T2DM and insufficient glycemic control despite metformin treatment. The primary endpoint is change from baseline in HbA1c after 104 weeks of treatment. Patients eligible to enroll had type 2 diabetes mellitus and insufficient glycemic control at screening (7.0 % ≤ HbA1c ≤ 10.0%) despite therapy with immediate release metformin at the maximum tolerated dose (≥1500 mg/day) unchanged for at least the last 12 weeks prior to randomization. There was a two-week open-label placebo run-in period preceding randomization. Patients who completed the planned 104-week treatment period in this trial were eligible to continue their randomized treatment in a 104-week, double-blind, double-dummy, extension treatment period. Patients who discontinued trial medication prematurely before the end of the 104-week treatment period were requested to continue to attend trial visits. Patients were to be followed-up for 4 weeks after discontinuation

from the trial or after the end of the treatment period. An interim analysis was planned when all patients continuing in the trial had completed 52 weeks (interim report included in the NDA submission) and final analysis to be conducted after patients completed 104 weeks. There were 2637 patients screened for the trial. Of these patients, 1678 patients started the placebo run-in period. Overall, there were 1549 patients randomized: 769 patients to 25 mg empagliflozin and 780 patients to glimepiride.

1245.31: An ongoing phase III double-blind, multi-national extension trial to investigate the long-term safety and tolerability and long-term efficacy of empagliflozin (10 or 25 mg once daily) as monotherapy or as add-on therapy in patients with T2DM. The trial was a combined study of 4 trials under 1 study number; the studies differed in the background therapy administered: no background therapy (‘drug naïve’ in preceding trial 1245.20), pioglitazone (alone or in combination with metformin in preceding trial 1245.19), metformin only (preceding trial 1245.23), and metformin plus sulfonylurea (preceding trial 1245.23). Patient eligibility was based on a successful completion of the preceding trial and on signature of the informed consent for the extension trial. Patients continued the treatment to which they had been randomized in the preceding trial; no re-randomization was performed in the extension trial. All patients who had received rescue medication in the preceding trial and were still treated with it at the time of Visit 1 of trial 1245.31 continued to take this rescue medication. In addition to the 24 weeks of treatment in the preceding trial, patients were to be treated for at least 52 weeks in the extension trial. Patients were to remain in the extension trial until the last patient entered has been treated for 52 weeks in the extension trial, resulting in a minimum treatment period of 52 weeks in the extension. The maximum treatment duration in the extension (dependent on recruitment time) was estimated to be 130 weeks. The first patient was enrolled in the trial on February 22, 2011. There were 1856 patients who enrolled in the extension trial: 305 patients from 1245.19, 615 patients from 1245.20 and 936 patients from 1245.23 (463 with metformin only and 473 with metformin plus sulfonylurea).

Reviewer’s Comment: The basic designs of trials were generally similar and all patients were treated for T2DM, which suggests that the trials were adequate to be incorporated into the meta-analysis for assessing CV risk. In accordance with the FDA Guidance recommendations the trials included patients with advanced CV disease, elderly patients, and patients with some degree of renal impairment.

3.2.2 Meta-Analysis Endpoints and Adjudication Methods

3.2.2.1 Meta-Analysis Endpoints

The pre-specified primary endpoint for the CV meta-analysis is a composite endpoint consisting of CV death (including fatal stroke and fatal MI), non-fatal MI, non-fatal stroke, or hospitalization due to unstable angina. This endpoint will be referred to as **MACE+** throughout this statistical review.

A pre-specified secondary endpoint for the CV meta-analysis is the composite endpoint of CV death (including fatal stroke and fatal MI), non-fatal MI, or non-fatal stroke. This secondary endpoint will be referred to as **MACE** in this statistical review.

Additional endpoints analyzed in this statistical review are the individual components of MACE+ and all-cause mortality.

Reviewer's Comment: The Applicant pre-specified CV related secondary and tertiary endpoints that are not described above. This review focuses on the evaluation of MACE+ and MACE because these are the typical endpoints used in the evaluation of CV safety for products intended to treat T2DM.

3.2.2.2 Endpoint Adjudication

The meta-analysis endpoints were based on adjudicated CV events from the empagliflozin phase II and III development program. The adverse events data for all trials included in the meta-analysis were searched for SMQs (defined in the adjudication charter) at least weekly to identify events to be sent for adjudication. All identified events were compiled by clinical reviewers from (b) (4), an independent research organization. The adjudication was performed by a Clinical Event Committee (CEC) governed by a charter. The CEC was made up of five independent expert cardiologists and five independent expert neurologists separated into the following two groups:

- CEC Cardiology (CECC) for the adjudication of fatal and non-fatal events suspect of myocardial ischemia and heart failure

- CEC Neurology (CECN) for the adjudication of events suspect of stroke (fatal and not-fatal strokes, TIA)

To ensure independence, members of the CEC were not permitted to have any affiliation with any site participating in any of the trials included in the meta-analysis. Furthermore, members of the CEC were not to have a direct, financial, or intellectual interest in knowing or influencing the outcome of the trials. To ensure objective and unbiased adjudication, the evaluations of potential CV events were performed without knowledge of the treatment assignment in the clinical trial.

Two CEC members were assigned to review and vote on assigned cases. If the two CEC members did not agree on the number of events and type, a third member reviewed and voted on the case. If the third member did not concur with either of the first two members on the total number and type(s) of event(s), the case was reviewed by the CEC panel in order to reach decision. The detailed definitions for the CV events are provided in the CEC charter and are based on draft recommendations of the Center for Drug Evaluation and Research (CDER) "Standardized Definitions for Cardiovascular Outcomes Trials: Draft Recommendation", issued by Division of Cardiovascular and Renal Products, CDER, FDA. The CEC charter indicated that

if there are limited or missing data, the CEC should adjudicate events based on their clinical expertise and the totality of the evidence.

(b) (4) was responsible for the monthly rotation of CEC members. Additionally, (b) (4) maintained and archived all CEC-related documents including all data files, voting data, and decision data. BI received hard copy files and electronic documents from (b) (4) upon completion of the trials.

Reviewer's Comment: The implementation of CEC conforms to recommendations in the FDA Guidance to Industry for evaluating CV safety with T2DM products.

3.2.3 Statistical Methodologies

The main effect measure discussed throughout this review is the hazard ratio (empagliflozin relative to comparator) for the outcomes defined in Section 3.2.2.1. A hazard ratio of one is indicative of equivalent rates between the two treatment groups, a hazard ratio greater than one is indicative of higher rate in the empagliflozin treatment group compared to comparator, and vice versa for hazard ratio less than one.

3.2.3.1 Analysis Populations

The main statistical analyses followed an intent-to-treat principle based on patients in the treated set (TS) defined as all randomized patients who received at least one dose of study drug. Patients were analyzed according to treatment assigned at randomization, regardless of actual treatment received. These analyses included all CV events that occurred during the planned observational period for each trial (or interim analyses cut-off date for ongoing trials) and includes events that occurred after study drug discontinuation, also known as an on-study analyses. Events with an onset date prior to first study medication were not included in the analyses. For the ongoing trials, the data cut-off dates for inclusion in the meta-analysis were June 22, 2012, July 31, 2012, and May 29, 2012 for trials 1245.25, 1245.28, and 1245.31, respectively.

On-treatment analyses were also conducted by the Applicant as supportive analyses of the primary endpoint. The Applicant defines the on-treatment set as those patients who were treated for at least 30 days and the on-treatment analysis included events that occurred no later than 4 weeks after the last intake of study drug or until the end of the entire trial, whichever occurred first.

Reviewer's Comment: The Applicant's definition of the on-treatment set is an example of a per protocol population as it considers only those patients who were treated for at least 30 days, thereby reducing the number of patients in the treated set. The reviewer considers alternate on-treatment analysis, see Section 3.2.3.3 for details.

3.2.3.2 Type I Error Control

In accordance with the FDA Guidance, the Applicant designed the CV meta-analysis to rule out an unacceptable increase in risk based on a pre-market hazard ratio risk margin of 1.8. The meta-analysis plan also included the options to assess excessive risk based on 1.3 risk margin at the time point that the 1.8 risk margin is ruled out and also to demonstrate improvement of CV risk (i.e. superiority), if 1.3 risk margin can be ruled out. Each of these three hypotheses was tested in a hierarchical manner (separately for MACE+ and MACE endpoints), see Table 4.

Table 4 Applicant's Pre-specified Hypotheses to be Tested in the CV Meta-analysis

	Interim analysis	Final analysis to test non-inferiority based on the 1.8 margin[1]	Supportive meta-analysis at time of completion of study 1245.25
$H_{0,1}$: $\gamma_{b/c} \geq 1.8$ non-inferiority based on 1.8 margin	X	X	
$H_{0,2}$: $\gamma_{b/c} \geq 1.3$ non-inferiority based on 1.3 margin	X[2]	X[2]	X
[redacted]			(b) (4)

[1] Only to be performed if the interim analysis could not establish non-inferiority based on the 1.8 margin.
[2] Only to be performed if non-inferiority based on the 1.8 margin had been shown [redacted] (b) (4)

Source: Applicant's meta-analysis report Table 6.5.1:1 (page 31)

Type I Error Control for 1.8 Risk Margin

According to the meta-analysis plan, there was to be one interim analysis and one final analysis for assessing the 1.8 risk margin after observing 182 MACE+. The overall significance level was 2.5%, one-sided. The alpha adjustment for the 1.8 margin was based on the Hwang, Shih and De Cani spending function (reference: Hwang et. al. (1990). "Group Sequential Designs Using a Family of Type I Error Probability Spending Functions". *Statistics in Medicine* 9, 1439-1445) to account for the interim analysis. With this many MACE+ and assuming true hazard ratio of 1.0 and annual event rate between 1.0% and 2.5%, the Applicant estimated that the final meta-analysis would have 95% power to rule out unacceptable increase in CV risk, see Table 5.

***Reviewer's Comment:* Because the number of events at the interim analysis database lock exceeded the planned 182 events for the final analysis, the meta-analysis currently under review, is considered final. Thus no alpha adjustments are required for testing the 1.8 pre-marketing risk margin, as no interim analysis was conducted.**

Table 5 Applicant's Planned Alpha Adjustment and Power for 1.8 Risk Margin

		Interim analysis	Final analysis to test non-inferiority based on the 1.8 margin
$H_{0,1}$: $\gamma_{b/c} \geq 1.8$ non-inferiority based on 1.8 margin	Significance level	0.0171	0.0156
	α spent	0.0171	0.0250
	# of events	130	182
	Power	85%	95%

Source: Applicant's meta-analysis report Table 6.5.1:2 (page 32)

Type I Error Control for 1.3 Risk Margin and Superiority

The Applicant planned to evaluate the 1.3 post-marketing risk margin using a meta-analysis of all trials (including trial 1245.25) and results from trial 1245.25 alone. In either case, an interim analysis was planned at the time point that the 1.8 risk margin was ruled out, both of which incorporated different alpha spending functions to control the Type I error within each data source (i.e. the meta-analysis or trial 1245.25 alone).

Reviewer's Comment: While the Type I error is controlled separately within each data source in April of 2012 the Agency communicated to the Applicant that using both data sources to evaluate the 1.3 risk margin does not control the Type I error across the two data sets. No changes to the evaluation of the 1.3 risk margin were made to address the multiplicity issue across both data sources. Since the multiplicity across the two data sources has not been addressed, the reviewer recommends that conclusions about the CV safety of empagliflozin with respect to the 1.3 risk margin be based on findings from the dedicated cardiovascular outcome trial alone, trial 1245.25, as this trial is adequately powered to rule out the 1.3 risk margin. In addition, the use of a dedicated cardiovascular outcome trial to evaluate the 1.3 risk margin is consistent with advice provided to other sponsors of products intended to treat Type 2 diabetes. As such, what follows is the planned Type I error control for trial 1245.25.

A trial specific alpha spending function based on the Haybittle-Peto boundary (0.0001 alpha spent when extracting data for the meta-analysis and 0.0249 for the final analysis) was used to control the alpha level at 0.025 one-sided for trial 1245.25. The hierarchical testing strategy for this trial is presented in Appendix II of this review.

3.2.3.3 Statistical Analyses

This section describes the pre-specified statistical analyses performed by the Applicant as well as additional post-hoc analyses conducted by the reviewer.

Pre-specified Analyses Performed by the Applicant

The primary meta-analysis to evaluate the 1.8 risk margin, as agreed upon with the FDA, estimates the hazard ratio of MACE+ in patients randomized to empagliflozin (pooling doses of 10mg and 25mg) once daily to patients randomized to all comparators (pooling active and placebo). The hazard ratio (empagliflozin versus comparator) is estimated using a Cox proportional hazard (PH) model stratified by trial. Data from the ongoing extension trial (1245.31) was pre-specified to be incorporated into the completed main trials (1245.19, 1245.20, and 1245.23). Patients who already experienced an event in the main trials were not to be considered at risk for this event during the extension. Each sub-study of trial 1245.23, that is, with metformin only or metformin plus sulfonylurea as background therapy, was stratified as a separate trial in the Cox model. Patients without an event were censored at the time the patient was last known to be event-free, but no later than the planned observation period for completed

trials or the data cut-off date for ongoing trials. Kaplan Meier (KM) plots are provided for graphical comparison of the survival functions between treatment groups.

Reviewer's Comment: The KM plots for meta-analysis of trials with imbalanced randomization ratios do not adequately account for trial-level differences and therefore, may be subject to Simpson's paradox. These plots are provided in this review for descriptive purposes rather than for testing significant differences between the treatment groups.

The Applicant pre-specified supportive analyses of MACE+ using the stratified Cox PH model described above and patients in the on-treatment set or OS (see Section 3.2.3.1 for definition). For the OS analysis, patients who did not experience the event were censored at the end of the planned observation period if the patients completed the planned treatments or at 4 weeks after last intake of study drug if the patient had not completed the treatment as planned.

The Applicant estimated the hazard ratio for the secondary endpoint of MACE as well as the individual components of MACE+ and all-cause mortality each using a Cox PH model stratified by trial. Analyses of MACE+ based on patient subgroups of age, gender, race, geographical region, BMI, renal function, duration of diabetes, and smoking status are also presented in this review.

Exploratory analyses of MACE+ were performed for each empagliflozin dose (10 mg or 25 mg) compared to all comparators as requested by the Agency in a communication dated June 24, 2011. These pairwise comparisons were performed in separate Cox models and included only those trials which included the treatment arms being compared.

Since the majority of events included in the meta-analysis were obtained from trial 1245.25, the ongoing cardiovascular outcomes trial, trial-level assessments of MACE+ for the 1.8 risk margin within patient subgroups, MACE+ components, and all-cause mortality were assessed. Additionally, findings from the trial-specific interim analysis for assessing the 1.3 risk margin based on MACE (the pre-specified primary endpoint for trial 1245.25) are also presented; see Appendix II for details of the testing strategy for this trial. All trial-specific analyses are based on a Cox PH model with treatment as the only predictor.

Additional Analyses Conducted by Reviewer

Because of the observed imbalance in number of patients discontinuing after completion of the main trials and not participating in the extension trial, the reviewer conducted separate post-hoc analyses for MACE+ and MACE, based on stratified Cox model in which the observation time during the extension trial was not included in the time at risk for these events. In these analyses, for patients who continued in the extension, the observation time ended upon enrollment and events that occurred during extension were excluded; no changes in observation time or censoring were made for patients who did not continue into the extension. Note that these post-hoc reviewer analyses were not performed for patient subgroups, MACE+ components, all-cause mortality or dose-level analyses.

Additionally, the reviewer performed alternate on-treatment analysis of MACE+ in which all patients in the treated set were analyzed and only events occurring while the patients was still on treatment or within 4 weeks of last treatment dose were included. Therefore, in the reviewer's analysis only the ascertainment window for events is altered and not the number of patients analyzed as was done in the Applicant's on-treatment analysis.

3.2.4 Patient Disposition, Demographic and Baseline Characteristics

3.2.4.1 Characteristics for All Trials Included in Meta-analysis

There were a total of 10,052 randomized patients enrolled in the trials included in the meta-analysis; 10,036 patients (6206 empagliflozin and 3830 comparator) received at least one dose of randomized treatment, thereby comprising the safety analysis population, also referred to as the treated set (TS). There were sixteen patients (11 empagliflozin and 5 comparators) who were randomized but did not receive any dose of study medication; therefore, these patients were excluded from the TS. Most patients in the empagliflozin treatment group were randomized to the 25 mg dose (3587 patients or 58%) compared to the 10 mg dose (2619 patients or 42%). In the comparator group, the majority of patients were randomized to the placebo arm (2827 patients or 74%) compared to the active arm (1003 patients or 26%). Note that active control was administered in two trials only: 1245.20 in which patients were randomized to sitagliptin and 1245.28 in which patients were randomized to glimepiride.

Table 6 shows that the overall discontinuation rate (for all reasons combined) was slightly lower in the empagliflozin group (13.3%) compared to the comparator group (16.9%) for all patients in the treated set. The lower overall discontinuation rate for empagliflozin patients was also observed within each trial. The majority of trial discontinuations were due to patients who did not continue into the extension phase of trials 1245.19, 1245.20, 1245.23MO and 1245.23MS. For trials 1245.20, 1245.23MO and 1245.23MS, the percentage of patients who did not continue into the extension phase was lower for empagliflozin patients (range: 18% – 19%) than comparator patients (range: 23% – 25%). For trial 1245.19, the percentage of patients who did not continue into the extension phase was still lower in the empagliflozin arm (29.4%) compared to comparator (32.7%), but these rates were higher than the rates observed in the other three trials; *the reasons for this are unclear.*

Reviewer's Comment: Results of post hoc analyses conducted by the reviewer to assess the impact of the discontinuation rates from the main trials to extension phase is provided in Section 3.2.5 of this review.

Table 6 Trial Discontinuation Rates Overall and By Trial in the Treated Set

Trial ID/Overall	N	Discontinuation Reason							All reasons combined
		Death	Lost to follow-up	Prematurely [a] discontinued in preceding trial	Did not continue [a] to extension trial	Withdrawal by subject	Other reason	Reason missing	
		n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
1245-19									
Empagliflozin	333	0 (0.0)	1 (0.3)	23 (6.9)	98 (29.4)	6 (1.8)	0 (0.0)	0 (0.0)	128 (38.4)
Comparator	165	0 (0.0)	0 (0.0)	18 (10.9)	54 (32.7)	4 (2.5)	0 (0.0)	0 (0.0)	76 (46.1)
1245-20									
Empagliflozin	448	0 (0.0)	1 (0.2)	38 (8.5)	86 (19.2)	16 (3.6)	0 (0.0)	0 (0.0)	141 (31.5)
Comparator	451	0 (0.0)	2 (0.4)	58 (12.9)	102 (22.6)	11 (2.4)	0 (0.0)	0 (0.0)	173 (38.3)
1245-23 MO									
Empagliflozin	430	0 (0.0)	1 (0.2)	25 (5.8)	80 (18.6)	13 (3.0)	0 (0.0)	0 (0.0)	119 (27.7)
Comparator	207	0 (0.0)	1 (0.5)	21 (10.1)	48 (23.2)	4 (1.9)	0 (0.0)	0 (0.0)	74 (35.7)
1245-23 MS									
Empagliflozin	441	1 (0.2)	1 (0.2)	34 (7.7)	79 (17.9)	12 (2.7)	0 (0.0)	0 (0.0)	127 (28.8)
Comparator	225	0 (0.0)	1 (0.4)	24 (10.7)	56 (24.9)	7 (3.1)	0 (0.0)	0 (0.0)	88 (39.1)
1245-25 [b]									
Empagliflozin	3046	40 (1.3)	7 (0.2)	--	--	96 (3.2)	0 (0.0)	0 (0.0)	143 (4.7)
Comparator	1513	26 (1.7)	4 (0.3)	--	--	59 (3.9)	0 (0.0)	0 (0.0)	89 (5.9)
1245-28 [b]									
Empagliflozin	765	3(0.4)	7 (0.9)	--	--	25 (3.3)	0 (0.0)	9 (1.2)	44 (5.8)
Comparator	780	2 (0.3)	9 (1.2)	--	--	38 (4.9)	0 (0.0)	10 (1.3)	59 (7.6)
1245-33									
Empagliflozin	324	0 (0.0)	10 (3.1)	--	--	9 (2.8)	63 (19.4)	0 (0.0)	82 (25.3)
Comparator	170	1 (0.6)	11 (6.5)	--	--	9 (5.3)	31 (18.2)	0 (0.0)	52 (30.6)
1245-36									
Empagliflozin	419	2 (0.5)	6 (1.4)	--	--	36 (8.6)	0 (0.0)	0 (0.0)	44 (10.5)
Comparator	319	3 (0.9)	3 (0.9)	--	--	29 (9.1)	0 (0.0)	0 (0.0)	35 (10.9)
Overall									
Empagliflozin	6206	46 (0.7)	34 (0.5)	120 (1.9)	343 (5.5)	213 (3.4)	63 (1.0)	9 (0.1)	828 (13.3)
Comparator	3830	32 (0.8)	31 (0.8)	121 (3.2)	260 (6.8)	161 (4.2)	31 (0.8)	10 (2.6)	646 (16.9)

MO=metformin only background therapy, MS=metformin in combination with sulfonylurea, N=number of randomized patients, n=number of patients discontinued
[a] Applies only to trials 1245.19, 1245.20, 1245.23MO, and 1245.23MS with ongoing extension phase
[b] Ongoing trials
Source: Created by the reviewer using dataset "adsl.xpt"

As shown in Table 7, the distributions for demographic characteristics were similar between the empagliflozin and comparator treatment groups. The majority of the patients were male (64.1% empagliflozin and 61.7% comparator), white (61.6% empagliflozin and 59.7% comparator) with a mean age of approximately 60 years and mean BMI approximately 30 kg/m^2. Most of the patients were enrolled in sites in Europe and Asia.

Table 7 Distribution of Demographic Characteristics in the Treated Set

Demographic Characteristic	Empagliflozin Arms			Comparator Arms		
	10 mg n=2619	25 mg n=3587	All n=6206	Placebo n=2827	Active n=1003	All n=3830
Sex, n (%)						
Male	1685 (64.3)	2290 (63.8)	3975 (64.1)	1800 (63.7)	562 (56.0)	2362 (61.7)
Female	934 (35.7)	1297 (36.2)	2231 (35.9)	1027 (36.3)	441 (44.0)	1468 (38.3)
Race, n (%)						
White	1605 (61.4)	2216 (61.8)	3821(61.6)	1689 (59.8)	595 (59.3)	2284 (59.7)
Black	107 (4.1)	132 (3.7)	239 (3.9)	132 (4.7)	11 (1.1)	143 (3.7)
Amer. Ind.	17 (0.7)	17 (0.5)	34 (0.6)	16 (0.6)	1 (0.1)	17 (0.4)
Asian	884 (34.0)	1216 (33.9)	2100(33.9)	983 (34.8)	396 (39.5)	1379 (36.0)
Hawaiian	3 (0.1)	3 (0.1)	6 (0.1)	5 (0.2)	0 (0.0)	5 (0.1)
Region, n (%)						
Europe	928 (35.4)	1333(37.2)	2261(36.4)	1012 (35.8)	355 (35.4)	1367 (35.7)
North America	626 (23.9)	760 (21.2)	1386(22.3)	666 (23.6)	152 (15.2)	818 (21.4)
Latin America	248 (9.5)	387 (10.8)	635 (10.2)	247 (8.7)	140 (14.0)	387 (10.1)
Asia	817 (31.2)	1107 (30.9)	1924 (31.0)	902 (31.9)	356 (35.5)	1258 (32.9)
Age, in years						
Mean (SD)	60.4 (9.8)	59.7 (10.2)	60.0 (10.0)	55.6 (10.3)	60.8 (9.9)	59.4 (10.3)
Range	27.0 – 88.0	19.0 – 98.0	23.0 - 88.0	23.0 - 83.0	26.0 - 88.0	19.0 – 98.0
Age, n (%)						
Less than 50	366 (14.0)	581 (27.0)	947 (15.3)	357 (12.6)	277 (27.6)	634 (16.6)
50-65	1343 (51.3)	1821 (50.8)	3164 (51.0)	1476 (52.2)	540 (53.8)	2016(52.6)
65-75	730 (27.9)	967 (27.0)	1697 (27.3)	781 (27.6)	155 (15.5)	936 (24.4)
Over 75	180 (6.9)	218 (6.1)	398 (6.4)	213 (7.5)	31 (3.1)	244 (38.0)
BMI, in kg/m^2						
Mean (SD)	30.1 (5.5)	30.1 (5.5)	30.1 (5.5)	30.1 (5.5)	29.8 (5.3)	30.0 (5.5)
Range	17.3 – 69.1	16.8 – 62.5	16.8 – 69.1	17.9 – 61.0	16.6 - 45.4	16.6 – 61.0
BMI, n (%)						
<25	465 (17.8)	633 (17.7)	1098 (17.8)	507 (17.9)	174 (17.4)	681 (17.8)
25-30	941 (35.9)	1282 (35.8)	2223 (35.8)	1011 (35.8)	392 (39.1)	1403 (36.6)
30-35	730 (27.9)	982 (27.4)	1712 (27.6)	777 (27.5)	267 (26.6)	1044 (27.3)
≥35	482 (18.4)	688 (19.2)	1170 (18.9)	531 (18.8)	170 (17.9)	701 (18.3)

Source: Created by the reviewer using dataset "adslcv xpt". Actual BMI values obtained from dataset "adttecv.xpt".

Reviewer's Comment: There were 24 patients (13 empagliflozin and 11 comparator) whose BMI exceeded 45 kg/m^2, the threshold used for trial eligibility. Including these patients in the analyses is not expected to change the overall conclusions.

Table 8 shows similar distributions for CV risk factors of smoking status, duration of diabetes and renal function between the empagliflozin and comparator treatment groups. Renal function was measured using the Modification of Diet in Renal Disease scale: eGFR<30 (severe), 30≤eGFR<60 (moderate), 60≤eGFR<90 (mild), and eGFR≥90 (normal). The majority of patients randomized in the trials had duration of diabetes exceeding 5 years (69.4 % empagliflozin and 64.2% comparator), had never smoked (51.4% empagliflozin and 53.9% comparator), and had mild renal impairment (52.1% empagliflozin and 51.0% comparator).

Table 8 Distribution of Cardiovascular Risk Factors in the Treated Set

Cardiovascular Risk Factor	Empagliflozin Arms			Comparator Arms		
	10 mg N=2619	25 mg N=3587	All N=6206	Placebo N=2827	Active N=1003	All N=3830
Diabetes Duration, in yrs. n (%)						
Less than 1	179 (6.8)	255 (7.1)	434 (7.0)	151 (5.3)	186 (18.5)	337 (8.8)
1 – 5	552 (21.1)	915 (25.5)	1467 (23.6)	613 (21.7)	422 (42.1)	1035 (27.0)
More than 5	1888 (72.1)	2417 (67.4)	4305 (69.4)	2063 (73.0)	395 (39.4)	2458 (64.2)
Smoking Status n (%)						
Never	1308 (49.4)	1884 (52.5)	3192 (51.4)	1458 (51.6)	606 (60.4)	2064 (53.9)
Ex-Smoker	966 (36.9)	1241 (34.6)	2207 (35.6)	1026 (36.3)	241 (24.0)	1267 (33.1)
Current Smoker	345 (13.2)	462 (12.9)	807 (13.0)	343 (12.1)	156 (15.6)	499 (13.0)
Renal Function (eGFR), n (%)						
Normal	713 (27.3)	1032 (28.8)	1745 (28.2)	715 (25.3)	417 (41.6)	1132 (29.6)
Mild	1404 (53.7)	1827 (51.0)	3231 (52.1)	1393 (49.3)	560 (55.8)	1953 (51.0)
Moderate	491 (18.8)	672 (18.8)	1163 (18.8)	665 (23.5)	26 (2.6)	691 (18.1)
Severe	7 (0.3)	53 (1.5)	60 (1.0)	52 (1.8)	0 (0.0)	52 (1.4)

Source: Created by the reviewer using dataset “adslcv xpt” and “adttecv xpt” for diabetes duration and renal function

3.2.4.2 Characteristics for CV Outcomes Trial 1245.25

At the time of this statistical review, there were 4559 patients (3046 empagliflozin and 1513 placebo) randomized prior to the data cut-off of June 22, 2012 and included in the treated set. Recall that no active controls were administered in this trial as the trial included standard of care. The trial discontinuation rates and reasons for discontinuation for 1245.25 are shown in Table 6 of this review. At the time of this review, the overall trial discontinuation rates were low (4.7% empagliflozin and 5.9% placebo). There are no relevant issues noted with the discontinuation reasons for this trial.

Table 9 and Table 10 show similar distributions of demographic characteristics and cardiovascular risk factors for empagliflozin and placebo patients randomized in trial 1245.25.

Table 9 Distribution of Demographic Characteristics in Trial 1245.25

Demographic Characteristic	Empagliflozin Arms 10 mg n=1521	25 mg n=1525	All n=3046	Placebo n=1513
Sex, n (%)				
Male	1069 (70.3)	1111 (72.9)	2180 (71.6)	1105 (73.0)
Female	452 (29.7)	414 (27.1)	866 (28.4)	408 (27.0)
Race, n (%)				
White	1070 (70.4)	1083 (71.0)	2153 (70.8)	1048 (69.4)
Black	78 (5.1)	76 (5.0)	154 (5.1)	89 (5.9)
American Indian	9 (0.6)	12 (0.8)	21 (0.7)	9 (0.6)
Asian	360 (23.7)	351 (23.0)	711 (23.4)	363 (24.0)
Hawaiian	1 (0.1)	0 (0.0)	1 (0.0)	2 (0.1)
Region, n (%)				
Europe	654 (43.0)	655 (43.0)	1309 (43.0)	648 (42.8)
North America	325 (21.4)	326 (21.4)	651 (21.4)	323 (21.3)
Latin America	223 (14.7)	228 (15.0)	451 (14.8)	224 (14.8)
Asia	319 (21.0)	316 (20.7)	635 (20.9)	318 (21.0)
Age, in years				
Mean (SD)	62.9 (8.7)	62.9 (8.6)	62.9 (8.7)	63.1 (8.9)
Range	31 – 88	30 – 88	30 – 88	33 – 88
Age, n (%)				
Less than 50	106 (7.0)	104 (6.8)	210 (6.9)	88 (5.8)
50-65	759 (49.9)	753 (49.3)	1512 (49.6)	768 (50.8)
65-75	516 (33.9)	544 (35.7)	1060 (34.8)	513 (33.9)
Over 75	140 (9.2)	124 (8.1)	264 (8.7)	144 (9.5)
BMI, in kg/m^2				
Mean (SD)	30.5 (5.3)	30.5 (5.4)	30.5 (5.4)	30.5 (5.3)
Range	17.3 – 69.1	17.1 – 62.5	18.2 – 69.1	17.1 – 61.1
BMI, n (%)				
<25	229 (15.1)	209 (13.7)	438 (14.4)	198 (13.1)
25-30	523 (34.4)	554 (36.3)	1077 (35.4)	549 (12.0)
30-35	481 (31.6)	455 (29.8)	936 (30.7)	466 (30.8)
≥35	287 (18.9)	305 (20.0)	592 (19.4)	299 (20.0)

Source: Created by the reviewer using datasets “adsl.xpt” and “adttecv.xpt”

Table 10 Distribution of Cardiovascular Risk Factors in Trial 1245.25

Cardiovascular Risk Factors	Empagliflozin Arms			Placebo n=1513
	10 mg n=1521	25 mg n=1525	All n=3046	
Diabetes Duration, in yrs. n (%)				
Less than 1	40 (2.6)	37 (2.4)	77 (2.5)	33 (2.2)
1 – 5	230 (15.1)	252 (16.5)	482 (15.8)	242 (16.0)
More than 5	1251 (82.3)	1236 (81.1)	2487 (81.7)	1238 (81.8)
Smoking Status n (%)				
Never	628 (41.3)	619 (40.6)	1247 (40.9)	626 (41.4)
Ex-Smoker	696 (45.8)	728 (47.7)	1424 (46.8)	713 (47.1)
Current Smoker	197 (13.0)	178 (11.7)	375 (12.3)	174 (11.5)
Renal Function (eGFR), n (%)				
Normal	306 (20.1)	320 (21.0)	626 (20.6)	310 (20.5)
Mild	807 (53.1)	797 (52.3)	1604 (52.7)	794 (52.5)
Moderate	399 (26.2)	395 (25.9)	794 (26.1)	402 (26.6)
Severe	5 (0.3)	10 (0.7)	15 (0.5)	5 (0.3)

Source: Created by the reviewer using datasets “adsl.xpt” and “adttecv.xpt”

3.2.5 Results and Conclusions

3.2.5.1 Descriptive Statistics for MACE+

(b) (4)

14 Page(s) have been Withheld in Full as b4 (CCI/TS) immeditately following this page

5 SUMMARY AND CONCLUSIONS

5.1 Collective Evidence and Statistical Issues

The cardiovascular meta-analysis was conducted in accordance with the meta-analysis plan that was finalized July 2012 and agreed upon with the FDA. The agreed upon population of interest for the meta-analysis comprised all randomized patients who received at least one dose of study medications. Data from 8 Phase II and III trials, including interim data from an ongoing dedicated cardiovascular outcomes trial (trial 1245.25), were included in the meta-analysis. The primary comparison was based upon the pooled empagliflozin treatment (10 mg and 25 mg doses) and the pooled comparator (placebo and active). The agreed upon primary safety endpoint of this meta-analysis was MACE+, a composite endpoint comprising cardiovascular death, non-fatal myocardial infarction, non-fatal stoke, or hospitalization for unstable angina. The key secondary endpoint was MACE, a composite endpoint comprising cardiovascular death, non-fatal myocardial infarction, or non-fatal stoke. All events included in the meta-analysis were based on positively adjudicated events by a Clinical Event Committee using standardized definitions. The primary objective of the meta-analysis was to demonstrate that the hazard ratio (empagliflozin to comparators) was smaller than the pre-market risk of 1.8 as stipulated in the FDA Diabetes Guidance for assessing CV safety. The pre-specified primary statistical analysis used a Cox proportional hazards model, stratified by trial.

In addition to the evaluation of the pre-marketing risk margin of 1.8, the Applicant pre-specified an interim meta-analysis to test the post-marketing risk margin of 1.3 for MACE+ and MACE at the time point that the 1.8 risk margin was ruled out. Additionally, the Applicant pre-specified to use the interim data from trial 1245.25 alone to assess the 1.3 risk margin. While the Type I error is controlled separately within each data source (i.e. the meta-analysis and trial 1245.25 alone), in April 2012[6] the Agency communicated that using both data sources to evaluate the 1.3 risk margin does not control the Type I error across the two data sets. No changes to the evaluation of the 1.3 risk margin were made to address the multiplicity issue across both data source. Because this multiplicity issue has not been addressed, we recommend that conclusions about the CV safety of empagliflozin with respect to the 1.3 risk margin be based on findings from the dedicated cardiovascular outcome trial alone, trial 1245.25, as this is consistent with advice provided to other sponsors of products intended to treat Type 2 diabetes.

There were a total of (b) (4)%) empagliflozin patients compared to (b) (4)%) comparator patients with positively adjudicated MACE+. Most of these events occurred in the trial 1245.25; specifically (b) (4)%) in empagliflozin patients and (b) (4)%) in placebo patients. The estimated hazard ratio for MACE+ across all trials included in the meta-analysis was (b) (4) with 95% CI **(** (b) (4) **)**. The upper bound of the 95% CI for MACE+ was less than the 1.8 pre-market risk margin. Post-hoc analyses of MACE+ performed by the reviewer yielded consistent results. Analyses were conducted to compare the risk of MACE+ within each empagliflozin dose relative to all comparators. For the 10 mg dose, the HR estimate

[6] Refer to preliminary meeting responses dated April 9, 2012.

of MACE+ was [redacted] with 95% CI [redacted]). For the 25 mg dose, which is the proposed marketing dose, the HR estimate was [redacted] with 95% CI [redacted]).

[redacted] Analyses of MACE and the individual components of MACE+ were also conducted; see summary of findings below in Table 24. Note that the analyses of individual components are meant for descriptive purposes only, [redacted]

Table 24 Summary of Meta-analysis Findings for MACE and MACE+ Individual Components

Outcome	Number of Patients with Events		HR (95% CI)
	Empagliflozin[1] (N=6206)	Comparator[2] (N=3830)	
MACE+	[redacted]		
MACE			
CV Death			
Non-fatal MI			
Non-fatal Stroke			
Hospitalization for UA			

[1]Pooled 10mg and 25 mg empagliflozin doses
[2]Pooled active and placebo comparators
CV=cardiovascular, MI=myocardial infarction, UA=unstable angina
Source: Created by the reviewer using dataset "adttecv.xpt"

The primary endpoint in trial 1245.25 was MACE as defined previously. Based on the pre-specified alpha spending function, testing at the interim analysis for trial 1245.25 is based on a two-sided α=0.002 level (i.e. a 99.98% confidence level). At the time of this statistical review, there was [redacted]%) MACE observed in empagliflozin patients compared to [redacted]%) MACE observed in placebo patients in trial 1245.25. Note that there were no active comparators in this trial as patients received standard of care. The estimated HR for MACE was [redacted] with corresponding 99.98% CI [redacted]). [redacted] Note that the trial is still ongoing with the final evaluation planned to occur at the completion of the trial with 691 planned events.

5.2 Conclusions and Recommendations

This is a statistical safety review of a CV meta-analysis report submitted by Boehringer Ingelheim, the Applicant for this NDA, to assess the CV safety of empagliflozin tablets relative to all comparators. The meta-analysis included Phase II and III trials in the empagliflozin development program, including an ongoing dedicated CV outcomes trial, trial 1245.25. The estimated HR for MACE+ (the primary endpoint) was (b) (4) with corresponding 95% CI (b) (4)). The upper bound of this confidence interval was less than 1.8 and therefore ruled out the unacceptable risk margin of 1.8 set forth in the FDA Guidance to establish CV safety of new antidiabetic products.

The Applicant pre-specified an interim meta-analysis to test the 1.3 risk margin at the time point that the 1.8 risk margin was ruled out as well as separate interim analysis of the ongoing cardiovascular outcome trial to rule out the 1.3 risk margin. The Applicant was advised on April 9, 2012 that using two data sources for the evaluation of the 1.3 risk margin did not control the Type I error rate and to revise the plan their plan for evaluating the 1.3 risk margin. No revisions were made to address the multiplicity issue across two data sources. As trial 1245.25 was adequately powered to evaluate the 1.3 risk margin on its own and to be consistent with other products intended for the treatment of T2DM, we recommend the evaluation of the 1.3 risk margin be based on trial 1245.25 alone. The planned two-sided alpha-level for the interim analysis of this trial was 0.002 to evaluate the 1.3 risk margin which was in agreement with the Agency. The interim analysis of MACE for trial 1245.25 resulted in an estimated HR of (b) (4) with corresponding (b) (4) % CI of (b) (4)). (b) (4)

1 Page has been Withheld in Full as b4 (CCI/tS) immediately following this page

APPENDIX II HIERARCHIAL TESTING STRATEGY FOR TRIAL 1245.25

The steps below outline the hierarchical testing strategy planned for assessing the 1.3 risk margin in the ongoing trial 1245.25, as described in the protocol. At each step, the hypothesis test is based on an overall one-sided $\alpha = 0.025$. The trial was designed to have one interim analysis at the time of data extraction for the CV meta-analysis and one final analysis after the pre-specified number of primary events are observed. The overall alpha is controlled at this level using the Haybittle-Peto method that assigns $\alpha = 0.0001$ one-sided at the interim analysis and $\alpha = 0.0249$ one-sided at the final analysis.

Step 1

Test the null hypothesis that HR$\geq$ 1.3 for the primary endpoint (MACE). The alternate hypothesis is HR < 1.3. If the upper bound of the two-sided α-adjusted CI is less than 1.3, then non-inferiority can be concluded for a margin of 1.3 for empagliflozin and proceed to step 2. Otherwise, the procedure will stop.

Step 2

Test the null hypothesis that HR$\geq$ 1.3 for the key secondary endpoint (MACE+). The alternate hypothesis is HR < 1.3. If the upper bound of the two-sided α-adjusted CI is less than 1.3, then non-inferiority can be concluded for a margin of 1.3 for empagliflozin and proceed to step 3. Otherwise, the procedure will stop.

(b) (4)

1 Page has been Withheld in Full as b4 (CCI/TS) immediately following this page

SIGNATURE DISTRIBTUION LIST

Primary Statistical Reviewer: Janelle K. Charles
Date: November 1, 2013

Concurring Reviewers: Mat Soukup, Team Leader DBVII
Mark Levenson, Deputy Division Director DBVII

cc:
Jean-Marc Guettier, Acting Division Director, DMEP
Karen Mahoney, Medial Team Leader, DMEP
William Chong, Medical Officer, DMEP
Patricia Madara, Project Manager, DMEP
Mark Levenson, Deputy Director, DBVII
Mat Soukup, Team Lead, DBVII
Janelle Charles, Statistical Reviewer, DBVII
Dongmei Liu, Statistical Reviewer, DBII
Mark Rothman, Statistical Team Leader, DBII
Lillian Patrician, Office of Biostatistics

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

/s/

JANELLE K CHARLES
11/01/2013

MATTHEW J SOUKUP
11/01/2013
Concur

MARK S LEVENSON
11/01/2013

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

CLINICAL STUDIES

NDA/Serial Number:	204629/0000
Drug Name:	(b) (4) (empagliflozin tablet)
Indication(s):	As an adjunct to diet and exercise to improve glycemic control in adults with type 2 diabetes mellitus
Applicant:	Boehringer Ingelheim
Date(s):	Stamp date: March 5, 2013 PDUFA goal date: March 5, 2014 Primary review due date: November 5, 2013
Review Priority:	Standard
Biometrics Division:	Division of Biometrics II
Statistical Reviewer:	Dongmei Liu, Ph.D.
Concurring Reviewers:	Mark Rothmann, Ph.D., Statistical Team Leader Thomas Permutt, Ph.D., Division Director
Medical Division:	Division of Metabolism and Endocrinology Products
Clinical Team:	William Chong, M.D., Medical Officer Karen M Mahoney, M.D., Diabetes Team Leader Jean-Marc Guettier, M.D., Acting Medical Division Director
Project Manager:	Patricia Madara

Keywords: NDA review, clinical studies

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1 EXECUTIVE SUMMARY

Boehringer Ingelheim proposes to register empagliflozin, a novel, orally administered, sodium-dependent glucose co-transporter-2 (SGLT-2) inhibitor, for treatment of type 2 diabetes mellitus (T2DM), in conjunction with diet and exercise (b) (4) Based on evaluation of change in glycosylated hemoglobin (HbA1c) from baseline to Week 24, the applicant claims insulin empagliflozin is effective in improving glycemic control in adult patients with T2DM.

1.1 Conclusions and Recommendations

The primary endpoint of change in HbA1c was met in each of eight randomized, controlled, phase 3 efficacy trials that were included in this submission. Based on my review, empagliflozin is effective in improving glycemic control in adult patients with T2DM. This NDA is approvable from statistical point of view. Summary of the primary endpoint in the six trials included in this review is given in Figure 1 in Section 1.3.

1.2 Brief Overview of Clinical Studies

This review focuses on efficacy data from 6 trials. Four of them (1245.20, 1245.23$_{(met)}$, 1245.23$_{(met+SU)}$, and 1245.19) are randomized, double-blind, placebo-controlled, parallel-group, phase 3 trials of empagliflozin (10 mg or 25 mg once daily) with a treatment duration of 24 weeks. The 4 trials were identical in their main design features and trial procedures, but differed in the required antidiabetic background medication. These four trials are considered pivotal trials to evaluate the efficacy and safety of empagliflozin in the target patient population.

Trial 1245.20 investigated empagliflozin as monotherapy and thus comprised patients without antidiabetic background medication. This trial also included an active control group sitagliptin. Patients in trial 1245.23$_{(met)}$ received metformin background therapy, and patients in 1245.23$_{(met+SU)}$ received metformin plus a sulphonylurea as background medication. Trial 1245.19 included patients taking background medication of pioglitazone with or without metformin.

All 4 pivotal trials have an ongoing 52 weeks double-blind, controlled extension to collect supportive long-term data. During the extension, patients in these 4 trials continued on the randomized trial treatment and the background medication they had taken in the initial trials. The 4 extensions are conducted under a single trial number (1245.31) with one trial protocol. This review focuses on data from the initial 24 weeks.

Other than the 4 pivotal trials, two supportive trials, 1245.33 (phase 2b) and 1245.28 (phase 3), are also reviewed for efficacy. Trial 1245.33 was conducted in patients with basal insulin as background therapy. This trial is double-blind and placebo controlled. It included both strength of empagliflozin 10 mg and empagliflozin 25 mg. In the first 18 weeks, the dose of the basal insulin was fixed, but could be adjusted by the investigator after week 18. Trial 1245.28 is double-blinded and active controlled, compares 25 mg empagliflozin once daily with glimepiride

(maximal tolerated dose between 1 and 4 mg) in patients with a metformin background therapy. Trial 1245.28 is ongoing. The result from interim analysis of this trial, with cutoff date at week 52, is included in this review.

Other than the 6 trials included in this review, there were two other supportive phase 3 trials submitted to this NDA. One is Trial 1245.36, the other is Trial 1245.48. Trial 1245.36 is a randomized, double-blind, placebo-controlled, parallel group, efficacy and safety study of empagliflozin (10 mg and 25 mg) as add on to pre-existing antidiabetic therapy over 52 weeks in patients with T2DM and renal impairment and insufficient glycaemic control. Trial 1245.48 is a randomized, double-blind, placebo-controlled, parallel group, efficacy and safety study of empagliflozin (10 mg, 25 mg) over 12 weeks in hypertensive patients with T2DM. Trial 1245.36 is reviewed by the clinical pharmacology team, mainly focus on dose response in patients with renal impairment. Trial 1245.48 is reviewed by the clinical team, mainly focus on safety issues in hypertensive patients.

1.3 Statistical Issues and Findings

The primary efficacy endpoint in the 4 pivotal trials is change from baseline in HbA1c after 24 weeks of treatment. In each pivotal trial, both doses of empagliflozin provided statistically significant and clinically meaningful reductions in HbA1c compared with placebo after 24 weeks of treatment. The differences between empagliflozin and placebo were largest (-0.83%) in the monotherapy trial 1245.20. Empagliflozin as an add-on therapy to metformin (Trial 1245.23$_{(met)}$), as add-on to metformin and a sulphonylurea (Trial 1245.23$_{(met+SU)}$), and as add-on to pioglitazone with or without metformin (Trial 1245.19) led to similar reductions of 0.5% or more in HbA1c compared with placebo. In each trial, except for 1245.23$_{(met+SU)}$, treatment with empagliflozin 25 mg provided numerically greater placebo-adjusted HbA1c reductions than treatment with empagliflozin 10 mg.

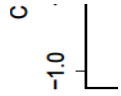
The key secondary efficacy endpoints in the pivotal trials varied and were ranked in different orders. Because the design and study population were fairly similar in the pivotal trials and there is no obvious reason why the secondary endpoints should differ across trials and be prioritized in different orders, secondary efficacy endpoints are collectively evaluated in this review across trials. The collective evidence shows consistent reduction in fasting plasma glucose (FPG), body weight and systolic blood pressure (SBP) by empagliflozin compared to placebo. Reduction in diastolic blood pressure (DBP) is not all significant in the 4 pivotal trials.

In Trial 1245.33 (patients with basal insulin as background therapy), the primary efficacy endpoint is change from baseline in HbA1c at 18 weeks. Both doses of empagliflozin provided statistically significant and clinically meaningful reductions in HbA1c compared with placebo after 18 weeks of treatment. There are two key secondary efficacy endpoints: change from baseline in basal insulin dose at 78 weeks and change from baseline in HbA1c at 78 weeks. The basal insulin doses during the first 18 weeks were similar for placebo and empagliflozin, both are around 47 IU. There were statistically significant differences in basal insulin dose at Week 78 in both empagliflozin groups compared to the placebo group and the basal insulin dose was significantly lower in the empagliflozin groups. Comparing to placebo, the basal insulin dose at Week 78 was reduced by 3.5IU (95% CI: 0.07-6.8) in empagliflozin 10mg group and by 4.1IU

(95% CI: 1.1-7.0) in empagliflozin 25mg group. For change from baseline in HbA1c at Week 78, both doses of empagliflozin provided statistically significant and clinically meaningful reductions in HbA1c compared with placebo. For both doses of empagliflozin, the reduction in HbA1c at Week 78 is smaller (around 0.7%) than the reduction at Week 18 (less than 0.7%).

In Trial 1245.28 (with glimepiride as active control), the primary efficacy endpoint for the interim analysis is change from baseline in HbA1c at 52 weeks. This trial used a 97.5% confidence interval and a p-value related to a non-inferiority margin of 0.3%. As mentioned in the Diabetes Guidance, when HbA1c is the primary efficacy endpoint, the division accepts a non-inferiority margin of 0.3% or 0.4% for active-controlled trials. When oral antidiabetic drug is the active control, the non-inferiority margin is usually set at 0.3%. The adjusted mean difference between empagliflozin and glimepiride in the change in HbA1c from baseline at Week 52 was -0.07% with a 97.5% CI of (-0.16, 0.02). With the upper limit of the 97.5% CI less than the pre-specified non-inferiority margin 0.3%, the null hypothesis for primary analysis was rejected; empagliflozin was non-inferior to glimepiride treatment. Empagliflozin showed a numerically greater HbA1c reduction compared with glimepiride at 52 weeks: -0.73% (SE 0.03) for the empagliflozin group compared with -0.66% (SE 0.03) for the glimepiride group. The key secondary endpoints include change from baseline in body weight, SBP, and DBP after 52 weeks of treatment and confirmed hypoglycemia episodes at Week 52. Empagliflozin 25mg was superior to glimepiride for all key secondary endpoints. At Week 52, the reductions in body weight, SBP, and DBP, and the lower incidence of confirmed hypoglycemic adverse events in the empagliflozin group compared with glimepiride were statistically significant.

Figure 1 Summary of the primary efficacy endpoint in all six trials included in the review (Adjusted mean difference of change in HbA1c between empagliflozin and comparator with 97.5% CI).



2 INTRODUCTION

2.1 Overview

2.1.1 Class and Indication

Empagliflozin is a novel, orally administered, sodium-dependent glucose co-transporter-2 (SGLT-2) inhibitor developed by Boehringer Ingelheim. Empagliflozin is proposed to be registered for treatment of type 2 diabetes mellitus (T2DM) in conjunction with diet and exercise, (b) (4)

Empagliflozin tablets are to be administered once daily, with or without food. The recommended therapeutic dose is 25 mg. However, in this submission, strengths of both 10 mg and 25 mg are evaluated.

2.1.2 History of drug development

The clinical development of empagliflozin started in January 2007. On 10 April 2008, Boehringer Ingelheim submitted the Investigational New Drug Application (IND 102145) for the development of empagliflozin (BI 10773) for the treatment for type 2 diabetes. The End-of-Phase II Meeting was held on 4 May 2010. Based on the results from two phase 2b trials 1245.9 and 1245.10, 10 mg and 25 mg empagliflozin once daily were chosen for the phase 3 trials. FDA provided written response to comment on the protocols of the phase 3 studies. The pre-NDA meeting took place on 27 November 2012 and the NDA was submitted on March 5, 2013.

The clinical program presented in this application comprises 30 phase 1 trials, 5 dose-finding phase 2 trials, and 13 phase 2b/3 trials. Overall, 13767 subjects were treated in these clinical trials. A total of 8506 patients with T2DM and 549 subjects without T2DM (516 healthy volunteers and 33 patients with renal or hepatic impairment) were treated with empagliflozin. Of the 8506 patients with T2DM, 3311 patients received empagliflozin 10 mg, 4563 patients received empagliflozin 25 mg and 632 received other dosages of empagliflozin.

2.1.3 Specific studies reviewed

This review focuses on 6 trials. Summary of the trial design are given in Table 1.

Four of them (1245.20, 1245.23$_{(met)}$, 1245.23$_{(met+SU)}$, and 1245.19) are randomized, double-blind, placebo-controlled, parallel-group, phase 3 trials of empagliflozin (10 mg or 25 mg once daily) with a treatment duration of 24 weeks. The 4 trials were identical in their main design features and trial procedures, but differed in the required antidiabetic background medication. These four trials are considered pivotal trials to evaluate the efficacy and safety of empagliflozin in the target patient population. Trial 1245.20 investigated empagliflozin as monotherapy and thus comprised patients without antidiabetic background medication. This trial also included an active control group sitagliptin. Patients in trial 1245.23$_{(met)}$ received metformin background therapy, and patients in 1245.23$_{(met+SU)}$ received metformin plus a sulphonylurea as background medication. Trials 1245.23$_{(met)}$ and 1245.23$_{(met+SU)}$ were conducted under a single trial number (1245.23) with one trial protocol, but are considered 2 independent trials for the purpose of the

analyses. Trial 1245.19 included patients taking background medication of pioglitazone with or without metformin. All 4 trials have an ongoing 52 weeks double-blind, controlled extension to collect supportive long-term data. During the extension, patients in these 4 trials continued on the randomized trial treatment and the background medication they had taken in the initial trials. The 4 extensions are conducted under a single trial number (1245.31) with one trial protocol. This review focuses on data from the initial 24 weeks.

Recommended by the clinical review team, efficacy data from 2 other trials are included in this review. One is trial 1245.33 (phase 2b), the other is trial 1245.28 (phase 3). Trial 1245.33 was conducted in patients with basal insulin as background therapy. This trial was originally designated as a phase 2b trial. Since it had confirmatory testing introduced via a protocol amendment, it is considered to be equivalent to a phase 3 trial for the assessment of the efficacy and safety of empagliflozin. This trial is double-blind and placebo controlled. It included both strength of empagliflozin 10 mg and empagliflozin 25 mg. In the first 18 weeks, the dose of the basal insulin was fixed, but could be adjusted by the investigator after week 18. Trial 1245.28 is double-blind and active controlled, compares 25 mg empagliflozin once daily with glimepiride (maximal tolerated dose between 1 and 4 mg) in patients with a metformin background therapy. Trial 1245.28 is ongoing. The result from interim analysis of this trial, with cutoff date at week 52, is included in this review.

In all 6 trials, patients were randomized equally to the available treatment arms. To ensure a homogeneous distribution of patients with more severe hyperglycemia across the randomized treatment groups, randomization was stratified by the HbA1c value at screening (<8.5% versus ≥8.5%). In Trials 1245.20, 1245.23$_{(met)}$, 1245.23$_{(met+SU)}$, and 1245.28, randomization was also stratified by geographic region and baseline renal impairment status. In Trial 1245.19, randomization was also stratified by background therapy and baseline renal impairment status.

Table 1 Summary of trial design.

Trial No.	Background therapy	Number of patients	Trial duration (weeks)	Treatment arms
1245.23$_{(met)}$	metformin	706	24 (initial) + 52 (extension)	placebo empagliflozin 10 mg empagliflozin 25 mg
1245.23$_{(met+SU)}$	metformin + sulphonylurea	767	24 (initial) + 52 (extension)	placebo empagliflozin 10 mg empagliflozin 25 mg
1245.19	pioglitazone ± metformin	498	24 (initial) + 52 (extension)	placebo empagliflozin 10 mg empagliflozin 25 mg
1245.20	monotherapy	986	24 (initial) + 52 (extension)	placebo empagliflozin 10 mg empagliflozin 25 mg sitagliptin 100 mg
1245.33	basal insulin ± metformin ±	494	18 (fixed dose insulin) +	placebo empagliflozin 10 mg

	sulphonylurea		60 (insulin dose adjustable)	empagliflozin 25 mg
1245.28	metformin	1549	104 (completed) + 104 (ongoing)	glimepiride 1~4 mg empagliflozin 25 mg

2.2 Data Sources

The data and final study report were submitted electronically. The submission was archived under the network path location <\\CDSESUB1\EVSPROD\NDA204629\204629.enx>. The information needed for this review was contained in Module 1 FDA Regional Information (cover letter, meeting correspondence, and labeling), Module 2.5 Clinical Overview, Module 2.7 Clinical Summary, and Module 5 Clinical Study Report. This review focuses on documents submitted to serial number 0000.

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

This submission is in electronic common technical document (eCTD) format with xml backbone. All required documents that are necessary for statistical review are submitted. No additional information request was made for statistical review.

Study datasets are provided as SAS XPORT transport files version 5 in CDISC standard. Both tabulation and analysis datasets are provided. Tabulation datasets include the source data without any derivations or enrichments, whereas analysis datasets also include derived and enriched data (such as formatted variables, populations, derived endpoints, data imputation information, etc.). The tabulation and analysis datasets are joinable by the unique record identifier (USUBJID).

The datasets are in good organization. Variable names are slightly varied across trials, but with clear description in the Define.pdf file. The reported analysis results are in good quality. My analysis on the primary and secondary efficacy endpoints gives approximately the same results as those reported in the clinical study report (CSR).

3.2 Evaluation of Efficacy

3.2.1 Study Design and Endpoints

Summary of the trial designs for the 6 trials included in this review are given in Table 1.

In all 6 trials, the primary efficacy endpoint is the change from baseline in HbA1c at a specific time point (as specified in Table 2). Key secondary efficacy endpoints varied across trials. The key secondary efficacy endpoints, ranked by the hierarchical testing order, are given in Table 2.

Table 2 Summary of key secondary efficacy endpoints.

Trial No.	1245.20	1245.23$_{(met)}$	1245.23$_{(met+SU)}$	1245.19	1245.33	1245.28
			Time to access the efficacy endpoints			
	24 weeks	24 weeks	24 weeks	24 weeks	18 weeks	52 weeks
			Key secondary efficacy endpoints in hierarchical testing order			
1	Change in body weight	Change in body weight	Change in body weight	Change in FPG	Change in basal insulin dose at week 78	Change in body weight
2	Change in SBP	Change in MDG	Change in MDG	Change in body weight	Change in HbA1c at week 78	Confirmed hypoglycemia
3	Change in DBP					Change in SBP
4						Change in DBP

- FPG: fasting plasma glucose
- MDG: mean daily glucose
- SBP: systolic blood pressure
- DBP: diastolic blood pressure

3.2.2 Patient Disposition, Demographic and Baseline Characteristics

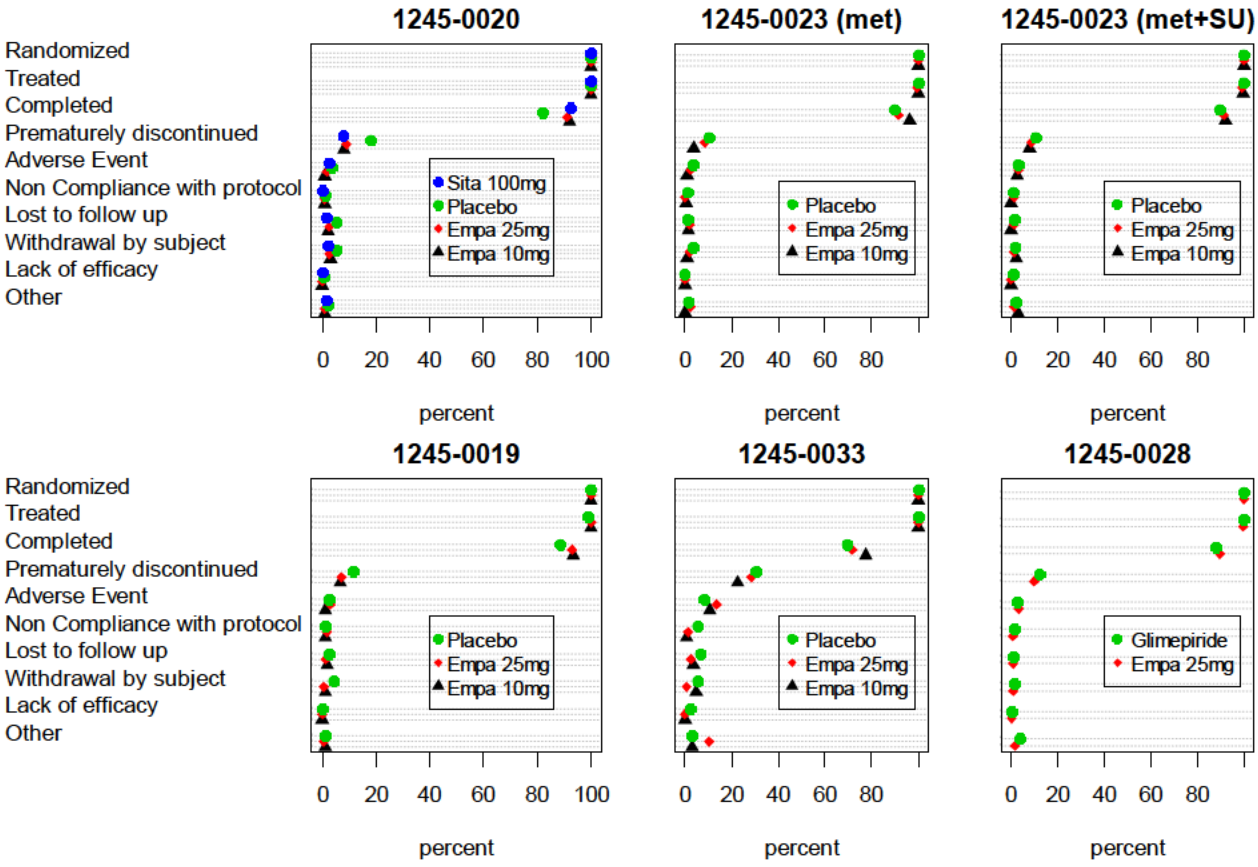
A total of 4634 patients were screened into the pivotal placebo-controlled trials. In each of the pivotal trials, the main reason for not entering patients was that HbA_{1c} values were not within the range specified as inclusion criteria. A total of 2705 patients were randomized to receive placebo (826 patients), empagliflozin 10 mg (832 patients), empagliflozin 25 mg (824 patients), or sitagliptin 100 mg (223 patients). A further 259 screened patients were assigned to open-label treatment with empagliflozin 25 mg. Of the 2477 patients in the pivotal trials who were randomized and treated with empagliflozin or placebo, around 90% completed the 24-week treatment duration as planned. The overall percentage of discontinuations from study medication was lower for patients in the empagliflozin groups than for patients in the placebo group. The most frequent reason for premature discontinuation was adverse events (4% overall).

In Trial 1245.33, 494 patients were randomized to double-blind treatment with either empagliflozin 10 mg (169 patients), empagliflozin 25 mg (155 patients), or placebo (170 patients). At Week 78, 360 patients (73%) had completed the treatment period: 131 patients (78%) in the empagliflozin 10 mg group, 111 patients (72%) in the empagliflozin 25 mg group, and 118 patients (69%) in the placebo group. A total of 134 patients (27%) prematurely discontinued trial medication during the 78-week treatment period including a lower proportion of patients in the empagliflozin 10 mg group (23%) compared with the empagliflozin 25 mg group (28%) and the placebo group (31%). The most frequent reason for premature discontinuation was adverse events (11% overall).

In Trial 1245.28, 1549 patients were randomized to double-blind treatment with empagliflozin (769 patients) or glimepiride (780 patients). At database lock, 1373 (89%) of the randomized and treated patients had not prematurely discontinued from the trial. The proportion of treated patients who prematurely discontinued study medication was higher in the glimepiride group (12%) than in the empagliflozin group (10%). The most frequent reason for premature discontinuation was adverse events (3% overall).

Further details of the disposition of randomized patients in all 6 trials included in this review are given in Figure 2.

Figure 2 Summary of patient dispositions.



The main inclusion and exclusion criteria were harmonized across the trials. Men and women diagnosed with T2DM and with insufficient glycaemic control (HbA1c ≥7.0%; upper limit 10% in most trials/sites), with a BMI of ≤45 kg/m^2 and at least 18 years of age (for patients in Japan the minimum age was 20 years) were to be included in the trials. Each of the trials had specific additional eligibility requirements, e.g. with regard to the permitted antidiabetic background medication.

For the pivotal trials, the key demographic characteristics were generally balanced across the randomized treatment groups. Just over half (55%) of the randomized patients were male.

Overall, 56% of patients were Asian, 41% were White, and 2% were Black or African American. Other races (American Indian / Alaska Native and Hawaiian / Pacific Islander) accounted for less than 1% of all patients.

In line with the data for race, just over half the patients (54%) were from Asia, approximately equal proportions were from North America (23%) and Europe (21%), and 3% were from Latin America. The mean age was 56 years, and the mean BMI was 29 kg/m_2. Based on eGFR, most patients had normal renal function (41%) or mild renal impairment (51%); 7% of patients had moderate renal impairment. Just over half (56%) of patients had a history of hypertension.

The percentages of patients who had been diagnosed with diabetes for more than 5 years reflected the eligibility criteria for the individual trials. The proportion of patients who had been diagnosed with diabetes for more than 5 years was lowest in trial 1245.20 (22%), which recruited treatment-naïve patients, and highest in trial 1245.23*(met+SU)* (78%), which recruited patients taking both metformin and a sulphonylurea as background therapy.

Trial 1245.20 included a randomized treatment arm with sitagliptin 100 mg. The demographic characteristics of the sitagliptin 100 mg group were similar to those of the patients randomized to placebo and empagliflozin.

3.2.3 Statistical Methodologies

Analysis Methods

The primary analysis of the HbA1c change from baseline was performed using an analysis of covariate model (ANCOVA) with baseline HbA1c as covariate, treatment and stratification factors for randomization as fixed effects. For the analysis of the key secondary endpoints, an ANCOVA was performed for the change from baseline in each of the key secondary endpoints. The model was similar to that used for the primary analysis, except that the respective secondary endpoint at baseline was added as a linear covariate. Descriptive statistics were calculated for hypoglycemic events and the time to onset of the first confirmed hypoglycemic event was summarized by Kaplan-Meier curve.

Hypothesis Testing

For the five placebo-controlled trials, two strength of empagliflozin were studied. In each trial, two null and alternative primary hypotheses were tested:

- H_{o1}: No difference in change of HbA1c between empagliflozin 10 mg and placebo
- H_{a1}: A difference in change of HbA1c between empagliflozin 10 mg and placebo
- H_{o2}: No difference in change of HbA1c between empagliflozin 25 mg and placebo
- H_{a2}: A difference in change of HbA1c between empagliflozin 25 mg and placebo

Both hypothesis tests were considered of equal importance and each was tested in parallel at 2.5% level (two-sided test). Depending on the significance of the 2 aforementioned hypothesis tests, 2 parallel hierarchical procedures on the key secondary efficacy endpoints are set up so to preserve the family-wise error rate at 2.5% for each dose. Confirmatory tests for primary and key

secondary endpoints were based on 2-sided tests at a 2.5% level (and corresponding 97.5% confidence intervals), preserving the overall 5% level of the trial.

The active-controlled trial 1245.28 only evaluated the effect of empagliflozin 25 mg. This trial used a 97.5% confidence interval and a p-value related to a non-inferiority margin of 0.3%. As mentioned in the Diabetes Guidance, when HbA1c is the primary efficacy endpoint, the division accepts a non-inferiority margin of 0.3% or 0.4% for active-controlled trials. When oral antidiabetic drug is the active control, the non-inferiority margin is usually set at 0.3%. The null hypothesis of inferiority of empagliflozin was to be rejected if the confidence interval was entirely below the non-inferiority margin 0.3%. Superiority of HbA1c lowering was not tested in a confirmatory way after 52 weeks of treatment in the interim analysis. If non-inferiority for HbA1c was established, superiority on the key secondary efficacy endpoints will be tested in a hierarchical order.

Analysis Population

All efficacy analyses in this review are based on the full analysis set (FAS) unless otherwise indicated. The FAS included all randomized patients who received at least one dose of study medication and had a baseline HbA1c measurement. The influence of important protocol violations and premature discontinuations was evaluated by analysis on the per-protocol set (PPS) and completer set. PPS comprised all patients in the FAS who did not have any important protocol violations leading to exclusion from the PPS. Completer set excluded patients who prematurely discontinued treatment.

Missing Data and Sensitivity Analysis

For the analysis of the primary and secondary efficacy endpoints, a last observation carried forward (LOCF) approach was used to replace missing data. Missing values within a course of measurements on treatment were interpolated based on the last observed value before the missing visit and the first observed value after the missing visit. Baseline values were carried forward if there was no post-randomization value available. Values measured after a patient had taken rescue medication were excluded and imputed using the LOCF method. As mentioned in the 2010 report on missing data by the National Academy of Sciences (NAS), *The Prevention and Treatment of Missing Data in Clinical Trials*, LOCF method is not recommended for missing data imputation. In this NDA submission, sponsor conducted sensitivity analysis based on several other imputation methods which gave similar results to the primary analysis.

Sensitivity analysis on the primary efficacy endpoint includes: 1) observed case (OC) approach, in which all available data were analyzed as observed, missing data was not imputed, and all values observed after a patient started rescue medication were excluded from the analysis; 2) original results (OR) analysis, which was performed for the use of rescue therapy and for the time to start of rescue therapy where missing data were not imputed, but values obtained after start of rescue therapy were used as measured; This approach is also indicated as IR, meaning 'including values on rescue medication'; 3) mixed model repeated measures (MMRM) approach, which using a restricted maximum likelihood (REML)-based MMRM approach with the model including the original covariate in the ANCOVA model and two additional covariates, 'visit' and

'visit by treatment interaction', as fixed effects. MMRM methods provide unbiased treatment estimates under the missing at random assumption. It estimates the treatment effects assuming the withdrawn patients have the same statistical behavior as those who continued.

3.2.4 Results and Conclusions

Pivotal Trials

Summary of the primary efficacy endpoint is given in Table 3. In each pivotal trial, both doses of empagliflozin provided statistically significant and clinically meaningful reductions in HbA1c compared with placebo after 24 weeks of treatment. The differences between empagliflozin and placebo were largest (-0.83%) in the monotherapy trial 1245.20. Empagliflozin as an add-on therapy to metformin (Trial 1245.23$_{(met)}$), as add-on to metformin and a sulphonylurea (Trial 1245.23$_{(met+SU)}$), and as add-on to pioglitazone with or without metformin (Trial 1245.19) led to similar reductions of 0.5% or more in HbA1c compared with placebo. In each trial, except for 1245.23$_{(met+SU)}$, treatment with empagliflozin 25 mg provided numerically greater placebo-adjusted HbA1c reductions than treatment with empagliflozin 10 mg.

Table 3 Summary of the primary efficacy endpoints in the 4 pivotal trials.

			Change from	Difference from Placebo		
Trial /Treatment Group	Number of subjects	Baseline HbA1c Mean (SE)	Baseline at Week 24 Mean (SE)	Adjusted Mean Difference	97.5% CI	P value
1245.20 (monotherapy)						
Placebo	228	7.91 (0.05)	0.06 (0.05)	---	---	---
Empagliflozin 10 mg	224	7.87 (0.06)	-0.66 (0.05)	-0.72	(-0.89, -0.56)	<0.0001
Empagliflozin 25 mg	224	7.86 (-.06)	-0.77 (0.06)	-0.83	(-0.99, -0.68)	<0.0001
Sitagliptin	223	7.85 (0.05)	-0.65 (0.05)	-0.70	(-0.86, -0.54)	<0.0001
1245.23$_{(met)}$ (metformin background)						
Placebo	207	7.90 (0.06)	-0.13 (0.05)	---	---	---
Empagliflozin 10 mg	217	7.94 (0.05)	-0.72 (0.05)	-0.57	(-0.72, -0.42)	<0.0001
Empagliflozin 25 mg	213	7.86 (0.06)	-0.75 (0.06)	-0.64	(-0.79, -0.48)	<0.0001
1245.23$_{(met+SU)}$ (metformin + sulphonylurea background)						
Placebo	225	8.15 (0.06)	-0.18 (0.05)	---	---	---
Empagliflozin 10 mg	225	8.07 (0.05)	-0.80 (0.05)	-0.64	(-0.79, -0.49)	<0.0001
Empagliflozin 25 mg	216	8.10 (0.06)	-0.77 (0.05)	-0.60	(-0.76, -0.44)	<0.0001
1245.19 (pioglitazone ± metformin background)						
Placebo	165	8.16 (0.07)	-0.14 (0.08)	---	---	---
Empagliflozin 10 mg	165	8.07 (0.07)	-0.57 (0.07)	-0.48	(-0.70, -0.26)	<0.0001
Empagliflozin 25 mg	168	8.06 (0.06)	-0.70 (0.07)	-0.63	(-0.85, -0.41)	<0.0001

The alternative missing data imputation methods includes: 1) observed case (OC) approach; 2) including values on rescue medication (IR) approach; and 3) mixed model repeated measures (MMRM) approach. The different imputation methods are applied to analysis population FAS, completer set, and PPS. The result of sensitivity analysis on the primary efficacy endpoint for

Trial 1245.20 is given in Figure 3 as an example. The results by LOCF and the sensitivity analysis are close.

Figure 3 Summary of sensitivity analysis result on change in HbA1c for Trial 1245.20.

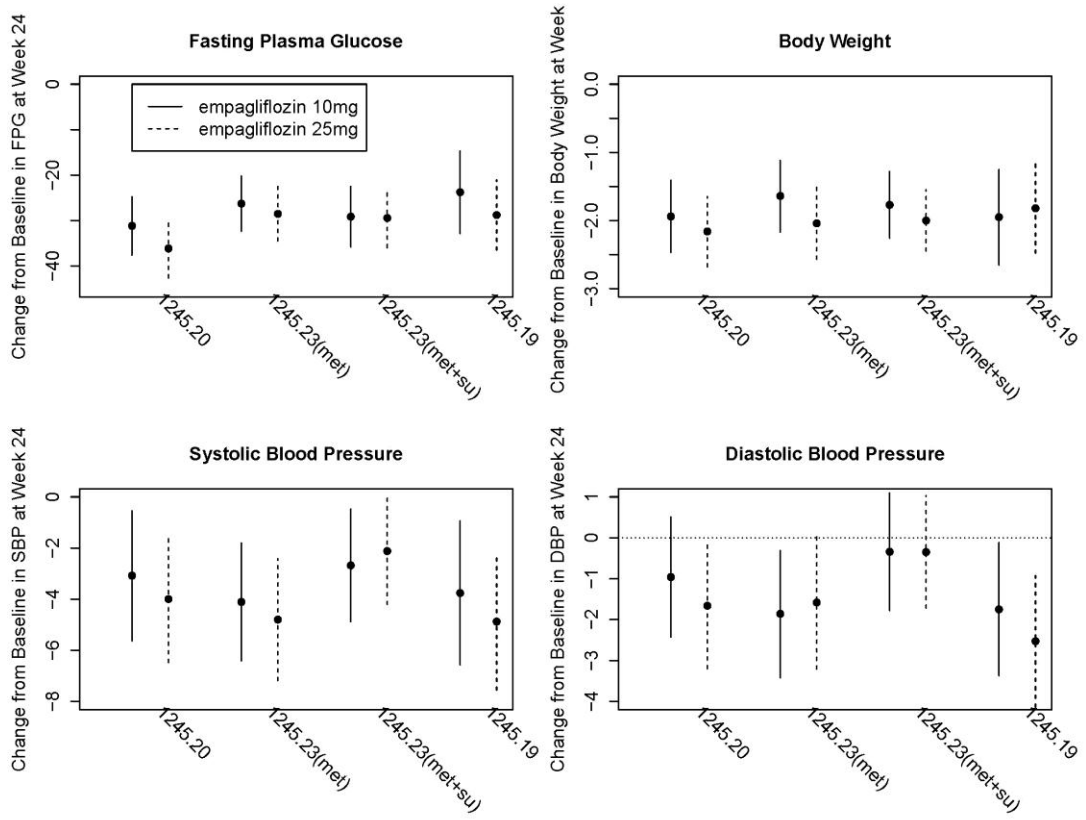
Summary of the secondary efficacy endpoints in the 4 pivotal trials is given in **Error! Reference source not found.**.

Change from baseline in body weight after 24 weeks of treatment is a key secondary efficacy endpoint in all 4 pivotal trials. In each trial, both doses of empagliflozin were superior to placebo, with similar reductions of up to 2 kg obtained. In each trial, except for Trial 1245.19, the higher empagliflozin dose (25mg) led to numerically larger weight reductions than the lower dose (10mg).

Change from baseline in blood pressure after 24 weeks of treatment is a key secondary efficacy endpoint in monotherapy Trial 1245.20. In other pivotal trials, it is an exploratory secondary endpoint. In Trial 1245.20, change in SBP was superior to placebo in both doses; but for change in DBP, reductions were not statistically significant. The treatment differences to placebo for empagliflozin as add-on to a metformin background (Trial 1245.23$_{(met)}$) or a metformin + sulphonylurea background (Trial 1245.23$_{(met+SU)}$) or a pioglitazone ± metformin background (Trial 1245.19) were in a similar range as for empagliflozin monotherapy (Trial 1245.20). The reduction for SBP is around 2~4mmHg; the reduction for DBP is around 1~2mmHg. For Trials 1245.20 and 1245.19, the higher empagliflozin dose provided numerically larger SBP and DBP reductions than the lower dose.

Change from baseline in FPG is a key secondary efficacy endpoint in Trial 1245.19 (pioglitazone ± metformin as background therapy). It is an exploratory secondary endpoint in other pivotal trials. In each trial, both doses of empagliflozin provided significant and clinically meaningful reductions in FPG compared with placebo. The reduction was around 30 mg/dL.

Figure 4 Summary of the secondary efficacy endpoints in the 4 pivotal trials (Adjusted mean difference of change in FPG, body weight, and blood pressure between empagliflozin and placebo with 97.5% CI).



Change from baseline in mean daily glucose (MDG) after 24 weeks of treatment is a key secondary efficacy endpoint only in Trials 1245.23$_{(met)}$ and 1245.23$_{(met+SU)}$. After 24 weeks of treatment, both doses provided statistically significant and clinically meaningful reductions in MDG in both trials. The reduction in MDG is around 10mg/dL. Analysis result is given in Table 4.

Table 4 Summary of mean daily glucose (MDG) in Trials 1245.23$_{(met)}$ and 1245.23$_{(met+SU)}$.

Trial /Treatment Group	Number of subjects	Baseline MDG Mean (SE)	Change from Baseline at Week 24 Mean (SE)	Difference from Placebo: Adjusted Mean Difference	Difference from Placebo: 97.5% CI	Difference from Placebo: P value
1245.23$_{(met)}$ (metformin background)						
Placebo	154	170 (3.1)	-1 (2.3)	---	---	---
Empagliflozin 10 mg	173	168 (2.5)	-11 (1.7)	-8	(-14, -2)	0.0055
Empagliflozin 25 mg	172	168 (2.6)	-15 (2.2)	-13	(-19, -7)	<0.0001
1245.23$_{(met+SU)}$ (metformin + sulphonylurea background)						
Placebo	179	170 (2.3)	1 (1.5)	---	---	---
Empagliflozin 10 mg	174	170 (2.2)	-11 (1.8)	-11	(-16, -5)	<0.0001
Empagliflozin 25 mg	153	173 (3.1)	-16 (2.2)	-14	(-19, -8)	<0.0001

Trial 1245.33

Summary of efficacy result for Trial 1245.33 is given in Table 5.

The primary efficacy endpoint is change from baseline in HbA1c at 18 weeks. Both doses of empagliflozin provided statistically significant and clinically meaningful reductions in HbA1c compared with placebo after 18 weeks of treatment.

There are two key secondary efficacy endpoints: change from baseline in basal insulin dose at 78 weeks and change from baseline in HbA1c at 78 weeks. The basal insulin doses during the first 18 weeks were similar for placebo and the two empagliflozin groups, all were around 47 IU. There were statistically significant differences in basal insulin dose at Week 78 in both empagliflozin groups compared to the placebo group and the basal insulin dose was significantly lower in the empagliflozin groups. Comparing to placebo, the basal insulin dose at Week 78 was reduced by 3.5IU (95%CI: 0.07-6.8) in empagliflozin 10mg group and by 4.1IU (95% CI: 1.1-7.0) in empagliflozin 25mg group. For change from baseline in HbA1c at Week 78, both doses of empagliflozin provided statistically significant and clinically meaningful reductions in HbA1c compared with placebo. For both doses of empagliflozin, the reduction in HbA1c at Week 78 is smaller (around 0.7%) than the reduction at Week 18 (less than 0.7%).

Table 5 Summary of efficacy results for Trial 1245.33.

Endpoint /Treatment Group	Number of subjects	Baseline Mean (SE)	Change from Baseline Mean (SE)	Difference from Placebo: Adjusted Mean Difference	Difference from Placebo: 97.5% CI	Difference from Placebo: P value
18 Weeks (no insulin adjustment)						
HbA1c (%)						
Placebo	170	8.18 (0.06)	0.06 (0.06)	---	---	---
Empagliflozin 10 mg	169	8.27 (0.06)	-0.56 (0.07)	-0.62	(-0.82, -0.42)	<0.0001
Empagliflozin 25 mg	155	8.27 (0.07)	-0.69 (0.06)	-0.74	(-0.93, -0.56)	<0.0001

78 Weeks (adjustable insulin dose after 18 weeks)						
HbA1c (%)						
Placebo	170	8.18 (0.06)	0.08 (0.07)	---	---	---
Empagliflozin 10 mg	169	8.27 (0.06)	-0.34 (0.07)	-0.40	(-0.62, -0.18)	<0.0001
Empagliflozin 25 mg	155	8.27 (0.07)	-0.61 (0.07)	-0.67	(-0.88, -0.46)	<0.0001
Basal Insulin Dose (IU)						
Placebo	170	46.9 (2.2)	33.1 (1.2)	---	---	---
Empagliflozin 10 mg	169	46.7 (2.3)	-0.3 (1.3)	-3.5	(-6.8, -0.07)	0.0237
Empagliflozin 25 mg	155	46.8 (2.3)	-0.7 (0.8)	-4.1	(-7.0, -1.1)	0.0024

Trial 1245.28

Summary of efficacy results for Trial 1245.28 is given in Table 6, Table 7, and Figure 5.

The primary efficacy endpoint for the interim analysis is change from baseline in HbA1c at 52 weeks. The adjusted mean difference between treatment groups in the change in HbA1c from baseline at Week 52 was -0.07% with a 97.5% CI of (-0.16, 0.02). With the upper limit of the 97.5% CI less than the pre-specified non-inferiority margin 0.3%, the null hypothesis for primary analysis was rejected; empagliflozin was non-inferior to glimepiride treatment. Empagliflozin showed a numerically greater HbA1c reduction compared with glimepiride at 52 weeks: -0.73% (SE 0.03) for the empagliflozin group compared with -0.66% (SE 0.03) for the glimepiride group. Superiority of HbA1c lowering was not tested in a confirmatory way in the interim analysis.

Table 6 Summary of efficacy results for Trial 1245.28.

			Change from	Difference from Placebo		
Endpoint /Treatment Group	Number of subjects	Baseline MDG Mean (SE)	Baseline at Week 24 Mean (SE)	Adjusted Mean Difference	97.5% CI	P value for superiority
HbA1c (%)						
Glimepiride	780	7.92 (0.03)	-0.66 (0.03)	---	---	---
Empagliflozin 25 mg	765	7.92 (0.03)	-0.73 (0.03)	-0.07	(-0.16, 0.02)	0.0743
Body Weight						
Glimepiride	780	83 (0.7)	1.6 (0.1)	---	---	---
Empagliflozin 25 mg	765	83 (0.7)	-3.2 (0.1)	-4.8	(-5.2, -4.5)	<0.0001
SBP						
Glimepiride	780	133 (0.6)	0.9 (0.5)	---	---	---
Empagliflozin 25 mg	765	133 (0.6)	-4.8 (0.5)	-5.7	(-7.2, -4.3)	<0.0001
DBP						
Glimepiride	780	79 (0.3)	0.1 (0.3)	---	---	---
Empagliflozin 25 mg	765	80 (0.4)	-2.4 (0.3)	-2.5	(-3.3, -1.7)	<0.0001

The key secondary endpoints include change from baseline in body weight, SBP, and DBP after 52 weeks of treatment and confirmed hypoglycemia episodes at Week 52. Empagliflozin 25mg was superior to glimepiride for all key secondary endpoints. At Week 52, the reductions in body

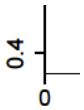
weight, SBP, and DBP, and the lower incidence of confirmed hypoglycemic adverse events in the empagliflozin group compared with glimepiride were statistically significant.

Table 7 Occurrence of confirmed hypoglycemic episodes at Week 52 in Trial 1245.28.

	Empagliflozin 25mg	Glimepiride
Number of subjects	765	780
Number of subjects with confirmed hypoglycemia episodes at Week 52 (%)	12 (2%)	159 (20%)
Comparison vs. glimepiride*		
Adjusted odds ratio	0.06	
97.5% CI	(0.03, 0.11)	
P value	<0.0001	

- * by Cochran-Mantel-Haenszel test adjusting for HbA1c at baseline

Figure 5 Kaplan-Meier curve for the time to onset of the first confirmed hypoglycemia episodes in Trial 1245.28.



3.3 Evaluation of Safety

Meta-analyses on cardiovascular events were reviewed by Dr. Janelle Charles from Division of Biometrics VII. Other safety events were reviewed by Dr. William Chong from Medical Division of Metabolism and Endocrinology Products. The reader is referred to those reviews for the evaluation of safety.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

Patients in the FAS of the four pivotal trials are pooled for subgroup analysis. Subgroup analyses on HbA1c were conducted by the ANCOVA model, similar to the one used for the primary efficacy analysis, with the additional covariate on the subgroups being analyzed and interaction between treatment effect and subgroups, and stratified by trial. Comparison of the primary efficacy endpoint in subpopulations is summarized in Figure 6 to Figure 16. Subgroups with number of patients less than 20 are excluded from the report.

The factors considered for subgroup analyses include:

1. Intrinsic factors:
   - Age
   - Sex
   - Race
   - Ethnicity
   - Region
   - BMI

2. Disease-related factors:
   - Baseline HbA1c
   - Diabetes duration
   - Renal function
   - Homeostasis model assessment index to assess insulin resistance (HOMA-IR)
   - Homeostasis model assessment index to assess insulin secretion (HOMA-IS)

In general, the subgroup analysis results are consistent with the results of overall efficacy of empagliflozin compared with placebo is consistent across the subgroup categories for race (Figure 8**), geographic region (**

Figure **10**), baseline BMI (Figure **11**), and time since diagnosis of diabetes (Figure **13**). Several subgroup populations show marginal or significant treatment-by-subgroup interaction effect (with P value by F test on the subgroup by treatment interaction ≤ 0.10). When appropriate (i.e. the subgroups are considered as ordinal, such as age groups or baseline renal impairment status), further analysis on trend along subgroups is done by chi-square test.

When looking at subgroup analysis results, we should keep in mind that subgroup analyses in this submission are considered exploratory due to the problem of inadequate power, multiplicity, interpretability, etc. Any potential heterogeneity or trend among subgroups discussed below should be considered as hypothesis that need further testing.

Age (shown in Figure 6)
For comparison between empagliflozin 25mg and placebo, the subgroup analysis on gender shows a greater treatment effect in younger patients than in older patients. The adjusted mean

difference from placebo in change of HbA1c from baseline at week 24 is -0.86 with SE of 0.1 in patients less than 50 years old. In age group 50 to 65 years old, the treatment effect is -0.64 with SE of 0.05. In age group 65 to 75 years old, the treatment effect is -0.56 with SE of 0.1. In patients with age greater than 75 years, the treatment effect is -0.54 with SE of 0.28. The p value for the treatment-by-age interaction term is 0.02. Further test on trend shows a p value of 0.15.

However, for the comparison between empagliflozin 10mg and placebo, heterogeneity of treatment effect across age groups is not as strong as that in comparison between empagliflozin 25mg and placebo (P=0.13). The treatment effects by empagliflozin 10mg in the three younger age groups are similar to that in empagliflozin 25 mg, but no significant reduction in HbA1c is shown in patients greater than 75 years old. Keep in mind that, there were only 45 patients in this subgroup, thus the estimation of treatment effect in this subgroup should be interpreted with caution.

Gender (shown in Figure 7)
For comparison between empagliflozin 25mg and placebo, the subgroup analysis on gender shows a greater treatment effect in males (mean= -0.76, SE=0.05) than in females (mean=-0.58, SE=0.06). The p value for the treatment-by-gender interaction term is 0.03. However, for the comparison between empagliflozin 10mg and placebo, no heterogeneity of treatment effect is detected in males vs. females (P=0.18).

Ethnicity (shown in Figure 9)
The subgroup analysis on ethnicity shows a greater treatment effect in patients not of Hispanic/Latino ethnicity compared with patients of Hispanic/Latino ethnicity in both doses. For empagliflozin 10mg, the treatment effect in patients of Hispanic/Latino ethnicity is -0.28 with SE of 0.15; the treatment effect in patients of not Hispanic/Latino ethnicity is -0.64 with SE of 0.04. For empagliflozin 25mg, the treatment effect in patients of Hispanic/Latino ethnicity is -0.41 with SE of 0.16; the treatment effect in patients of not Hispanic/Latino ethnicity is -0.70 with SE of 0.04. The p values for the treatment-by-ethnicity interaction term are 0.01 for empagliflozin 10mg and 0.05 for empagliflozin 25mg. There is a notable imbalance in the numbers of patients in both ethnicities, the number of patients who are not Hispanic/Latino ethnicity is 10 times more than the number of patients who are Hispanic/Latino ethnicity. Therefore, results of the subgroup analyses by ethnicity should be interpreted with caution.

Baseline HbA1c (shown in Figure **12**)
Treatment with empagliflozin provides greater HbA1c reductions from baseline in patients with a higher baseline HbA1c. For empagliflozin 10mg, the treatment effect in patients with baseline HbA1c <8.5% is -0.49 with SE of 0.04; the treatment effect in patients with baseline HbA1c ≥8.5% is -0.84 with SE of 0.08. For empagliflozin 10mg, the treatment effect in patients with baseline HbA1c <8.5% is -0.59 with SE of 0.04; the treatment effect in patients with baseline HbA1c ≥8.5% is -0.85 with SE of 0.08. The p-values for the treatment-by-baseline HbA1c interaction term are <0.0001 for empagliflozin 10mg and 0.004 for empagliflozin 25mg.

Baseline renal impairment status (shown in Figure 14)

The subgroup analysis by renal impairment category shows a reduction in the efficacy of empagliflozin 25mg with increasing degree of renal impairment. For patients with normal renal function at baseline, the treatment effect is -0.84 with SE of 0.06; for patients with mild renal impairment at baseline, the treatment effect is -0.58 with SE of 0.05; for patients with moderate renal impairment at baseline, the treatment effect is -0.32 with SE of 0.16. The p-value for the treatment-by-renal impairment interaction term is 0.001. Further test on trend shows a p value of 0.01. Empagliflozin is a SGLT-2 inhibitor. Considering SGLT-2 is mainly expressed in the renal proximal tubules and accounts for approximately 90% of renal glucose reabsorption, therefore kidney serves as the target of empagliflozin in improving glycemic control, it may be reasonable to see renal impairment influence the efficacy of the treatment.

However, for empagliflozin 10mg, heterogeneity of treatment effect across subgroups with different levels of baseline renal impairment is not as strong as that in empagliflozin 25mg (P=0.60). There is still a reduction in the efficacy with increasing degree of renal impairment, but the differences among groups are much smaller. For patients with normal renal function at baseline, the treatment effect is -0.66 with SE of 0.06; for patients with mild renal impairment at baseline, the treatment effect is -0.57 with SE of 0.05; for patients with moderate renal impairment at baseline, the treatment effect is -0.52 with SE of 0.16.

Baseline HOMA-IR (shown in Figure 15)
For comparison between empagliflozin 25 mg and placebo, the treatment difference is greatest for the patients in the highest baseline HOMA-IR category (>8.5 mU/L × mmol/L). However, there is no obvious trend across the other 3 HOMA-IR categories. For patients with baseline HOMA-IR ≤4 mU/L × mmol/L, the treatment effect is -0.66 with SE of 0.06; for patients with baseline HOMA-IR >4 to ≤5.5 mU/L × mmol/L, the treatment effect is -0.45 with SE of 0.12; for patients with baseline HOMA-IR >5.5 to ≤8.5 mU/L × mmol/L, the treatment effect is -0.69 with SE of 0.1; for patients with baseline HOMA-IR >8.5 mU/L × mmol/L, the treatment effect is -1.18 with SE of 0.12. The p-value for the treatment-by-baseline HOMA-IR interaction term is <0.0001. Further test on overall trend along the 4 subgroups by baseline HOMA-IR level shows a p value <0.0001.

For empagliflozin 10mg, heterogeneity of treatment effect across subgroups with different levels of baseline HOMA-IR is not as strong as that in empagliflozin 25mg (P=0.58). There is still an overall increase in the efficacy with increasing baseline HOMA-IR, but the differences among groups are much smaller. For patients with baseline HOMA-IR ≤4 mU/L × mmol/L, the treatment effect is -0.62 with SE of 0.05; for patients with baseline HOMA-IR >4 to ≤5.5 mU/L × mmol/L, the treatment effect is -0.60 with SE of 0.11; for patients with baseline HOMA-IR >5.5 to ≤8.5 mU/L × mmol/L, the treatment effect is -0.69 with SE of 0.1; for patients with baseline HOMA-IR >8.5 mU/L × mmol/L, the treatment effect is -0.81 with SE of 0.13. The p-value for the treatment-by-baseline HOMA-IR interaction term is 0.58. Further test on overall trend along the 4 subgroups by baseline HOMA-IR level shows a p value 0.32.

It should be noted that the 4 categories of baseline HOMA-IR are defined by the sponsor with arbitrary cutoffs with no clear rational. HOMA-IR measures are not used in routine clinical practice, but are primarily used in research. The subgroup analysis on HOMA-IR is considered exploratory and it is not clear what the clinical relevance of this result is.

Baseline HOMA-IS (shown in Figure 16)
For comparison between empagliflozin 10 mg and placebo, the treatment difference is greatest for the patients in the lowest baseline HOMA-IS category (<25 mU/mmol) and shows a reduction in efficacy along with increasing baseline HOMA-IS level. For patients with baseline HOMA-IS ≤25 mU/mmol, the treatment effect is -0.86 with SE of 0.09; for patients with baseline HOMA-IS >25 to ≤40 mU/mmol, the treatment effect is -0.66 with SE of 0.09; for patients with baseline HOMA-IS >40 to ≤70 mU/mmol, the treatment effect is -0.61 with SE of 0.08; for patients with baseline HOMA-IS >70 mU/mmol, the treatment effect is -0.53 with SE of 0.08. The p-value for the treatment-by-baseline HOMA-IS interaction term is 0.04. Further test on overall trend along the 4 subgroups by baseline HOMA-IS level shows a p value 0.01.

For comparison between empagliflozin 10 mg and placebo, the treatment difference is greatest for the patients in the lowest baseline HOMA-IS category (<25 mU/mmol), but shows an increment in efficacy along with increasing baseline HOMA-IS level in the rest 3 categories. For patients with baseline HOMA-IS ≤25 mU/mmol, the treatment effect is -0.83 with SE of 0.09; for patients with baseline HOMA-IS >25 to ≤40 mU/mmol, the treatment effect is -0.62 with SE of 0.09; for patients with baseline HOMA-IS >40 to ≤70 mU/mmol, the treatment effect is -0.64 with SE of 0.07; for patients with baseline HOMA-IS >70 mU/mmol, the treatment effect is -0.69 with SE of 0.08. The p-value for the treatment-by-baseline HOMA-IS interaction term is 0.27. Further test on overall trend along the 4 subgroups by baseline HOMA-IS level shows a p value 0.33.

Subgroup analysis on baseline HOMA-IS has the same issues with HOMA-IR, the categories are defined by the sponsor with arbitrary cutoffs with no clear rational. HOMA-IS measures are not used in routine clinical practice, but are primarily used in research. The subgroup analysis on HOMA-IS is considered exploratory and it is not clear what the clinical relevance of this result is.

Figure 6 Subgroup analysis on Change in HbA1c by age group.

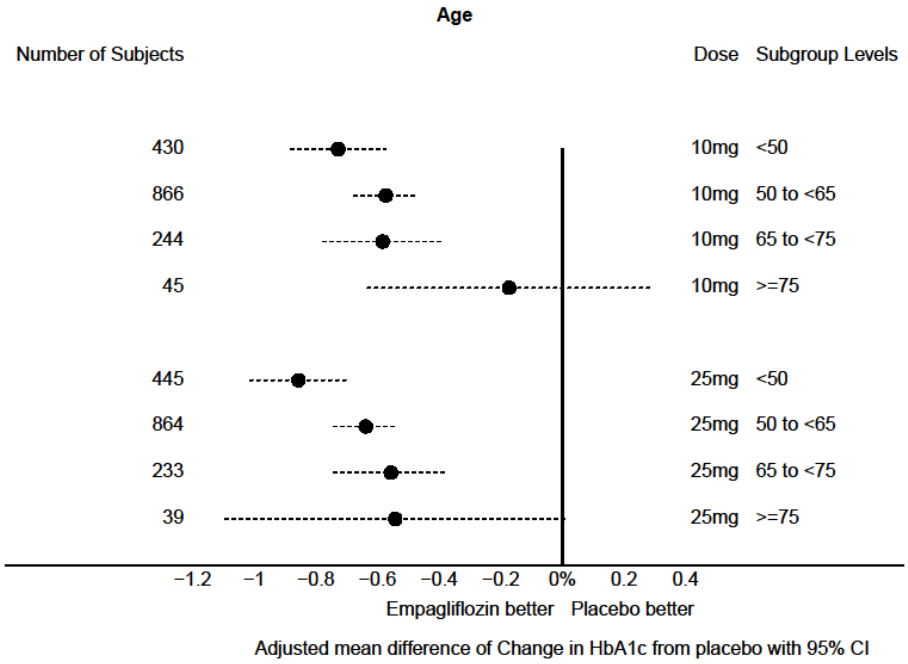


Figure 7 Subgroup analysis on Change in HbA1c by gender.

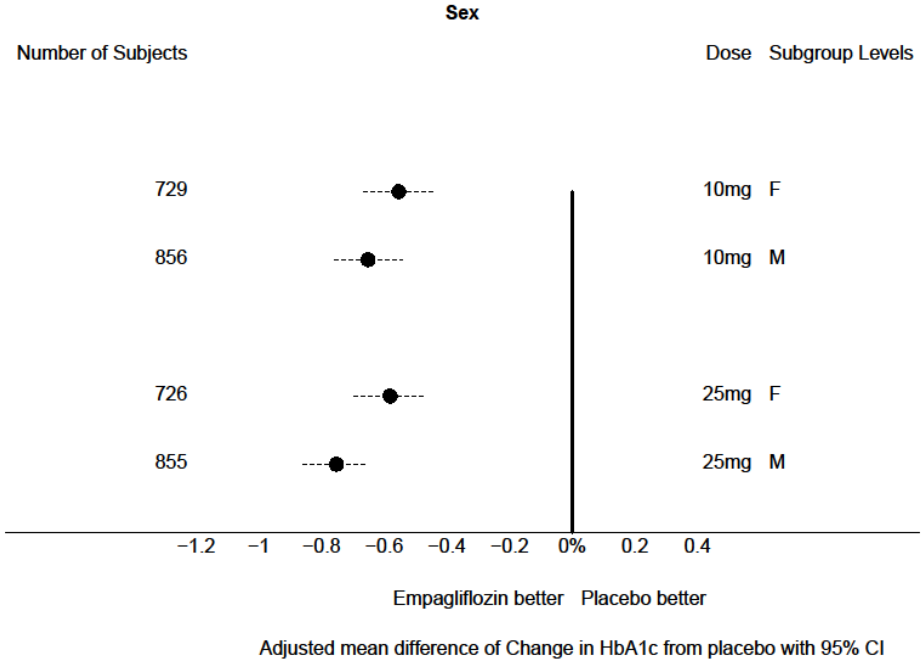


Figure 8 Subgroup analysis on Change in HbA1c by race.

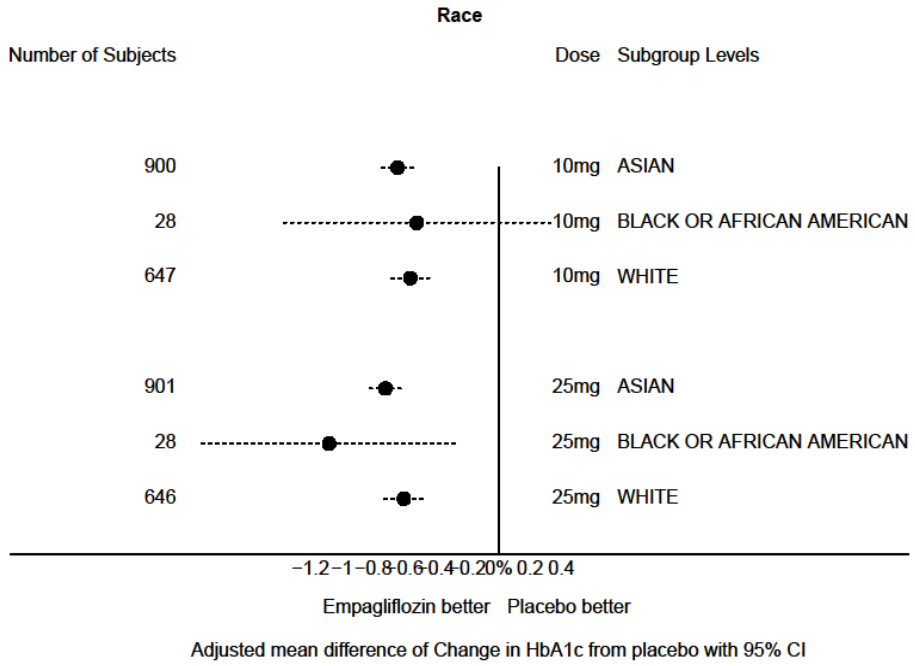


Figure 9 Subgroup analysis on Change in HbA1c by ethnicity.

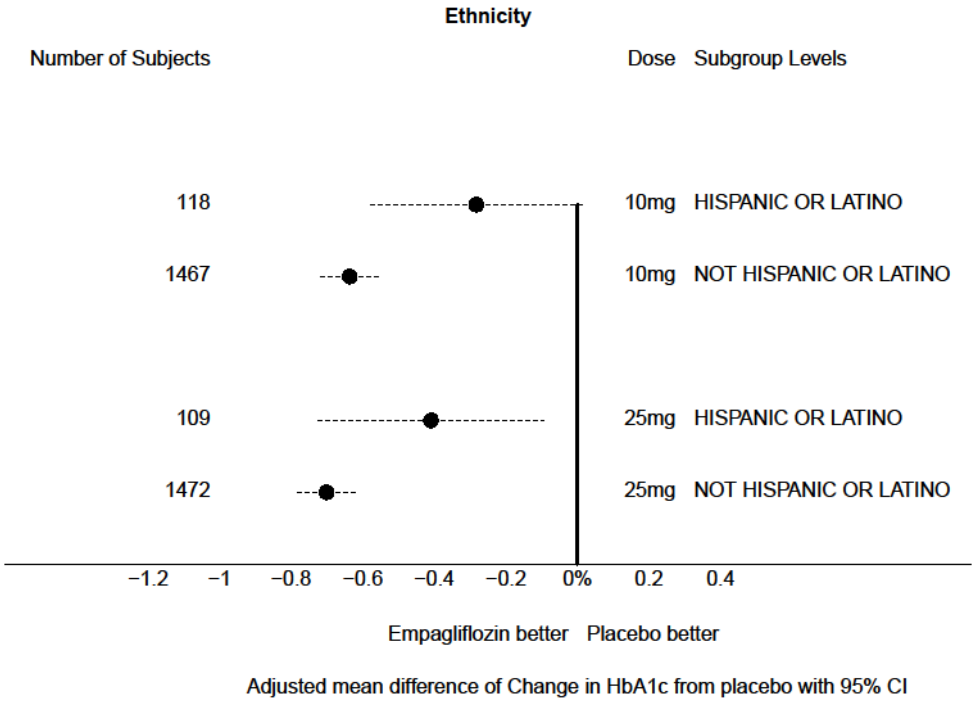


Figure 10 Subgroup analysis on Change in HbA1c by geographic region.

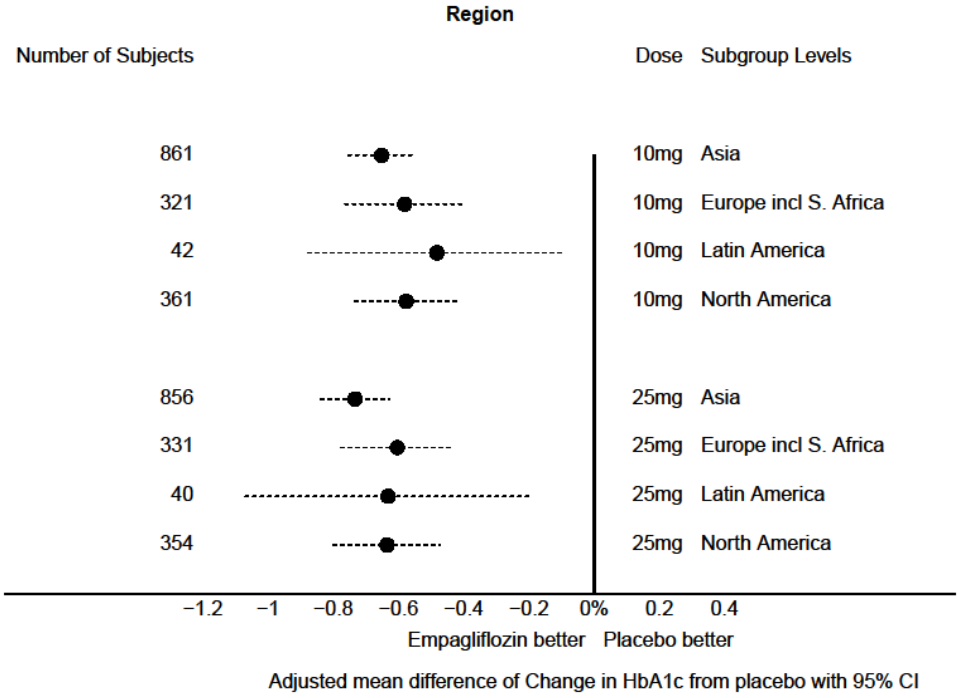


Figure 11 Subgroup analysis on Change in HbA1c by BMI.

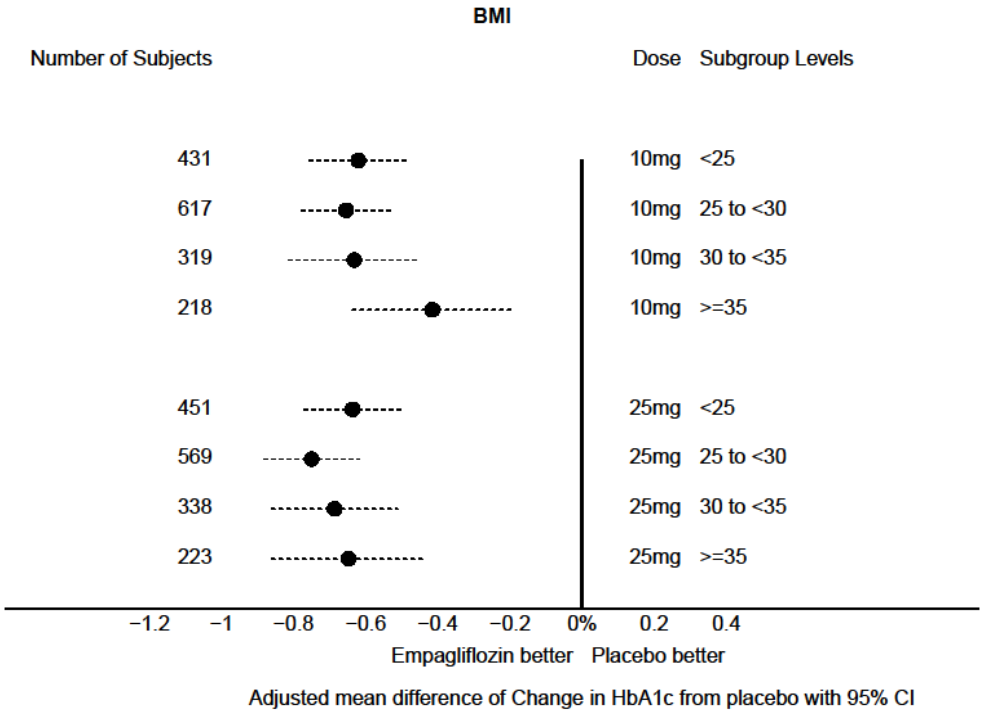


Figure 12 Subgroup analysis on Change in HbA1c by baseline HbA1c group.

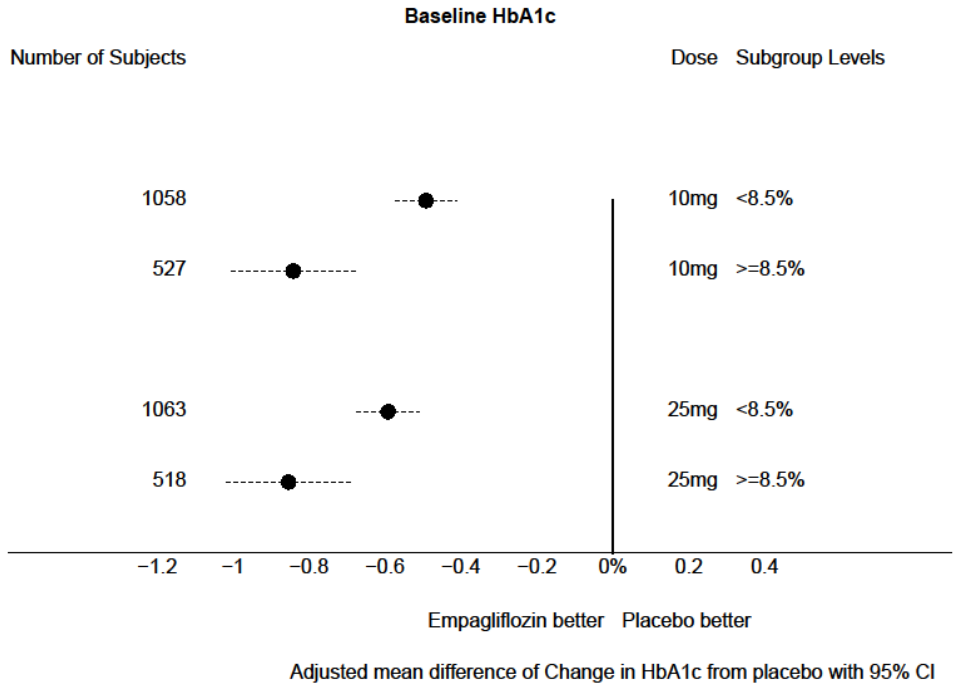


Figure 13 Subgroup analysis on Change in HbA1c by duration of diabetes.

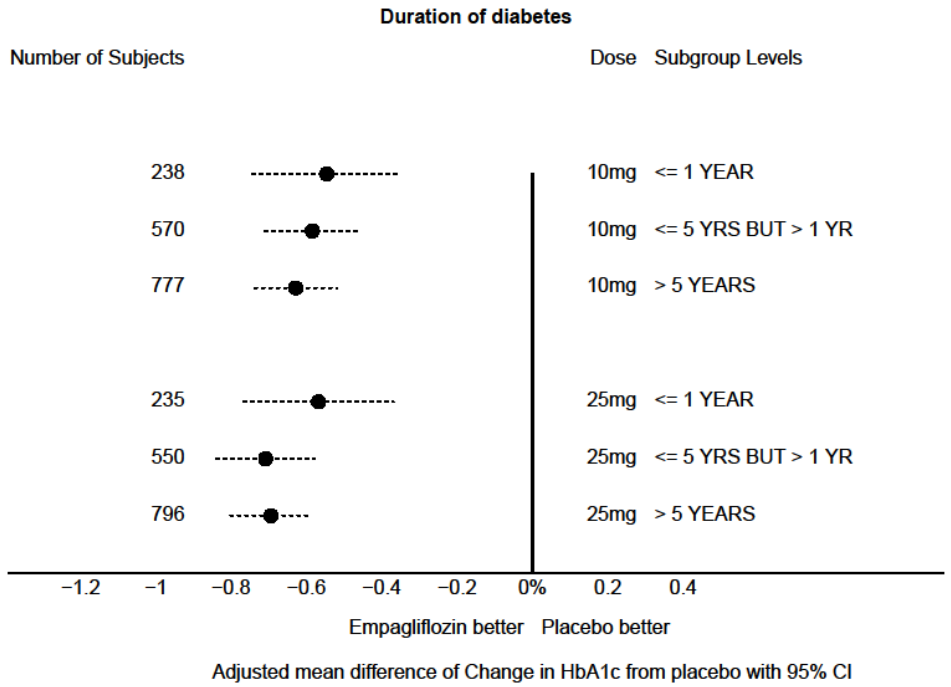


Figure 14 Subgroup analysis on Change in HbA1c by baseline renal impairment status.

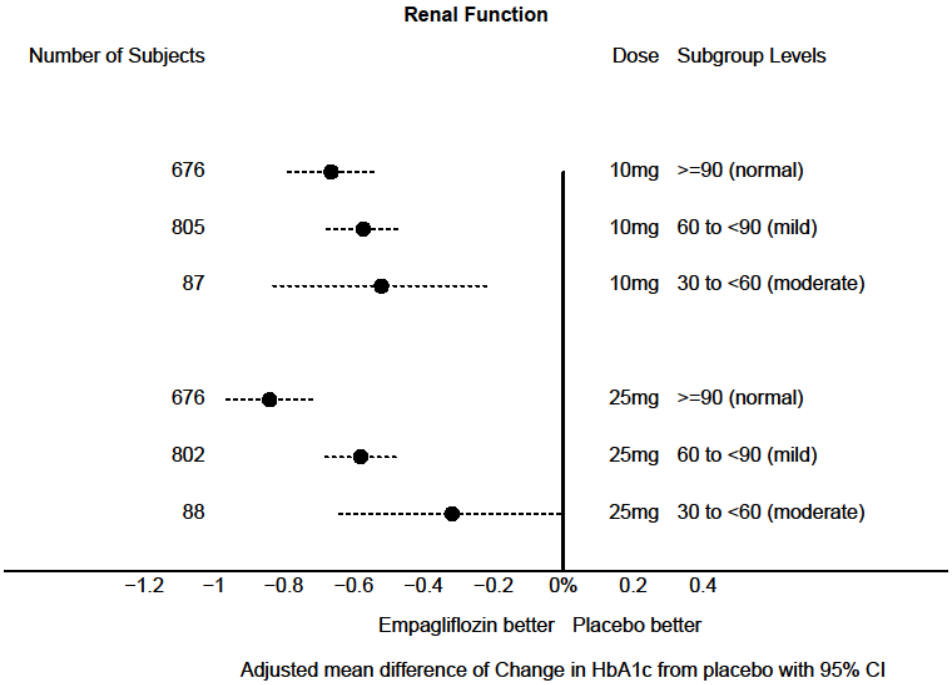
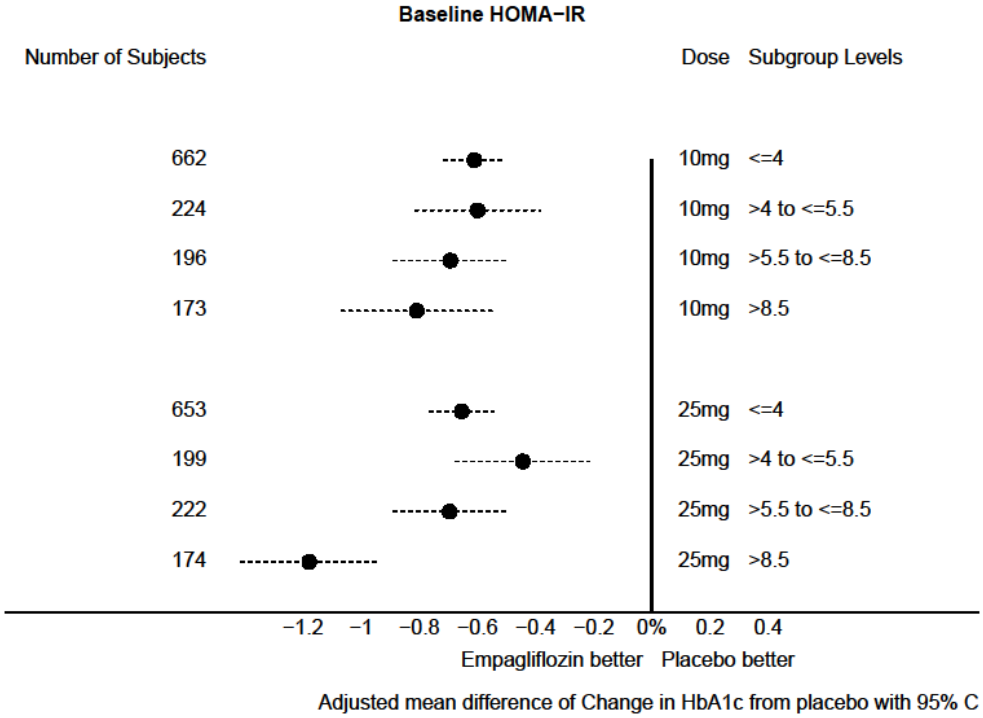


Figure 15 Subgroup analysis on Change in HbA1c by baseline HOMA-IR.



5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

The main statistical issue in this submission is analysis on key secondary efficacy endpoints. The secondary endpoints differ across trials and were prioritized in different orders. This submission did not provide explanation on how the secondary endpoints were selected and prioritized. Because the design and study population were fairly similar in the pivotal trials and there is no obvious reason why the secondary endpoints should differ across trials and be prioritized in different orders, it is more reasonable to evaluate the results of secondary efficacy endpoints collectively across trials than to evaluate them separately in different order in each single trial. The collective evidence across trials shows consistent reduction in FPG, body weight and SBP by empagliflozin compared to placebo. It is better to include these secondary efficacy endpoints consistently in the label, rather than by how they were prespecified and ordered in the protocol.

The other statistical issue is several subpopulations showed marginal or significant treatment-by-subgroup interaction effect. However, considering the problem of inadequate power, multiplicity, interpretability, etc., the subgroup analysis in this submission should be considered as exploratory. Any potential heterogeneity or trend among subgroups discussed in this review should be considered as hypothesis that need further testing.

5.2 Conclusions and Recommendations

The review on efficacy supports the claim of using empagliflozin for improving glycemic control in adult patients with T2DM. This NDA is approvable from statistical point of view.

5.3 Labeling Recommendations

(b) (4)

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

/s/

DONGMEI LIU
10/30/2013

MARK D ROTHMANN
10/30/2013
I concur

THOMAS J PERMUTT
10/30/2013
concur

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 STUDIES

IND/NDA Number:	NDA 204-629
Drug Name:	BI 10773
Applicant:	Sponsor: Boehringer Ingelheim Pharmaceuticals, Inc.900 Ridgebury Road Ridgefield, Connecticut 06877 USA Test Facility: (b) (4)
Documents Reviewed:	Electronic data submitted on April 15, 2013, Also include the sponsor's reports submitted.
Review Priority:	Standard
Biometrics Division:	Division of Biometrics -6
Statistical Reviewer:	Min Min, Ph.D.
Concurring Reviewer:	Karl Lin, Ph.D.
Medical Division:	Division of Metabolism and Endocrinology Products
Reviewing Pharmacologist:	Mukesh Summan Ph.D.
Project Manager:	
Keywords:	Carcinogenicity, Dose response

Table of Contents

1. Background

In this submission, the sponsor included reports of two animal carcinogenicity studies, one in rats and one in mice. These studies were intended to further assess the carcinogenic potential of BI 10773 in rats and mice, when administered daily via oral gavage to mice and rats for at least 104 weeks.

Results of this review have been discussed with the reviewing pharmacologist Dr. Summan who suggested doing analysis for rat and female mouse studies.

2. 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 two control groups. Two hundred and fifty Crl:WI(Han) rats of each sex were randomly allocated to treated and control groups. There were 50 animals per sex in each of groups. Treated animals each received dose preparations 100, 300, or 700 mg/kg.

Male and female Crl:WI(Han) rats were assigned to groups, and doses were administered as indicated in the following table. Rats were dosed via oral gavage.

Group	No. of Animals		Dose Level (mg/kg/day)	Dose Concentration[a] (mg/mL)
	Male	Female		
Carcinogenicity Animals				
1 (Vehicle Control)[b]	50	50	0	0
2 (Vehicle Control)[b]	50	50	0	0
3 (Low)	50	50	100	10
4 (Mid)	50	50	300	30
5 (High)	50	50	700	70
Toxicokinetic Animals[c]				
6 (Vehicle Control)[b]	6	6	0	0
7 (Low)	12	12	100	10
8 (Mid)	12	12	300	30
9 (High)	12	12	700	70
Sentinel Animals[d]				
10 (Sentinel Animals)	10	10	e	e

a Rats were dosed at a volume of 10 mL/kg.
b Groups 1, 2, and 6 received vehicle control article [0.5% (w/v) hydroxyethylcellulose in reverse osmosis water] only.
c Three extra rats/sex in Groups 6 through 9 were assigned to the study and used, as needed, as replacement rats (dependent on survival). All remaining rats were sacrificed and discarded with surviving toxicokinetic rats after the final blood collections.
d Five rats/sex served as replacement rats, as needed, to compensate for possible mortality.
e Sentinel rats were not dosed.

2.1. Sponsor's analyses

2.1.1. Survival analysis

Evaluations of trend and heterogeneity of survival data were performed using the Cox-Tarone binary regression on life tables and Gehan-Breslow nonparametric methods using the National Cancer Institute (NCI) Life Table Package. The Cox-Tarone method is more sensitive to late deaths, and the Gehan-Breslow method is more sensitive to early deaths due to a test article. As a result, they are both important tools to evaluate observable incidence data. Week 105 of the dosing phase was treated as the end of the study in the NCI package for males and females. Those rats sacrificed at the scheduled interval and rats sacrificed for other reasons or accidental deaths, if any, were censored in the analyses. Continuity-corrected, one-sided tail probabilities for trend and group comparisons were evaluated at the 0.05 significance level.

Sponsor's findings: The Kaplan-Meier product-limit survival curves from the sponsor's report are presented in Figure 1 and Figure 2 for males and females, respectively. Sponsor's analysis showed that administration of BI 10773 had no effect on overall male or female survival rates. The survival rates to scheduled sacrifice (Week 105 of the dosing phase) were 64, 70, 78, and 64% in males and 57, 64, 66, and 62% in females given 0 (Groups 1 and 2 combined), 100, 300, or 700 mg/kg/day, respectively. The number of BI 10773-treated animals that survived through 104 weeks of dose administration was sufficient to allow for evaluation of carcinogenicity.
In conclusion, administration of BI 10773 daily via oral gavage to male and female Crl:WI(Han) rats at a dose level of 100, 300, or 700 mg/kg/day for at least 104 weeks had no effect on survival.

Figure 1: Kaplan-Meier plot of Survival in Male Rats

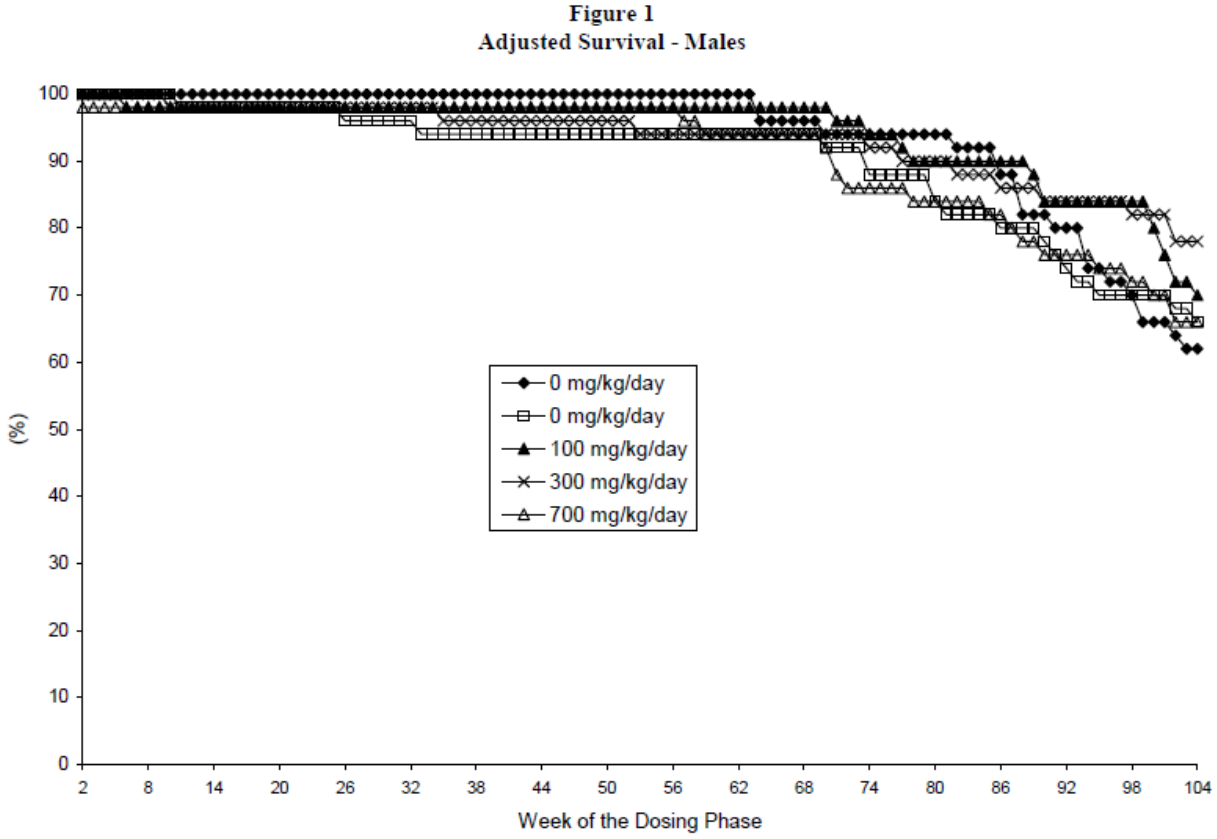
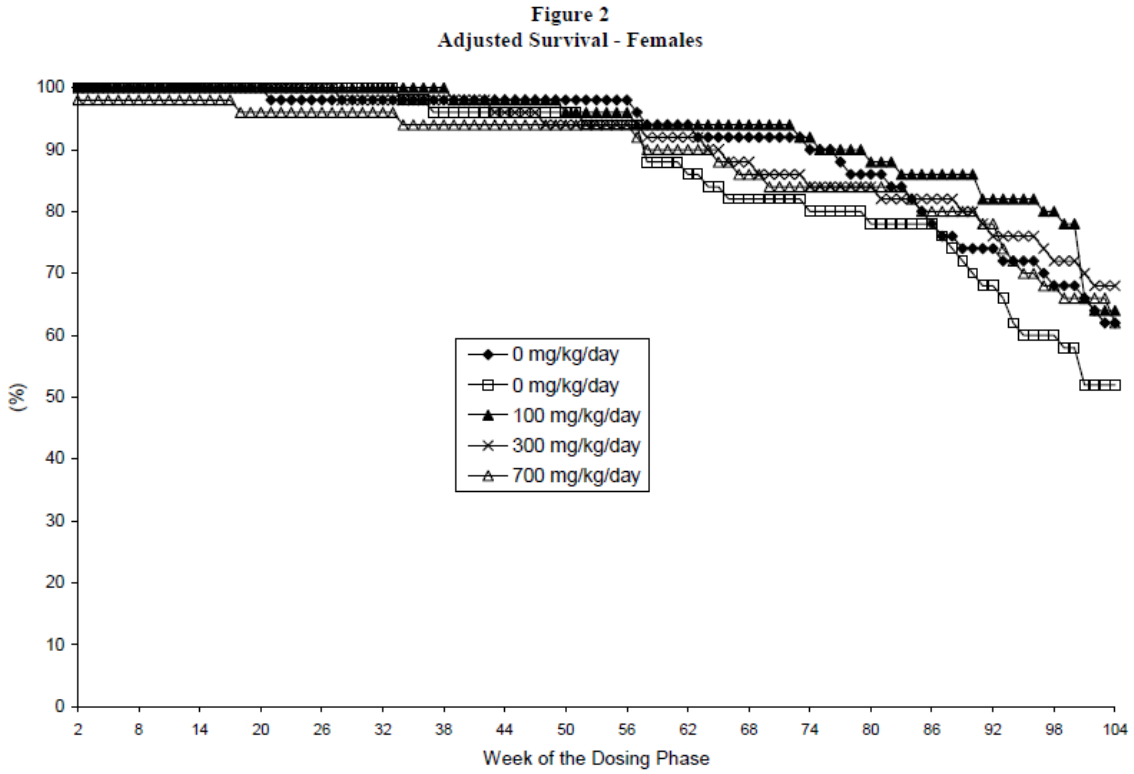


Figure 2: Kaplan-Meier plot of Survival in Female Rats



2.1.2. Tumor data analysis

Neoplastic lesions were chosen for statistical analyses if the incidence in at least one of the treated groups was increased by at least two occurrences over either control group.
The incidental tumors (i.e., tumors not assigned as the cause of death of the rats by the study pathologist) were analyzed by fixed International Agency for Cancer Research (IARC) interval-based prevalence tests (Peto et al., 1980). These intervals were Weeks 0 to 52, 53 to 78, 79 to 105, and 105 (terminal sacrifice) for both sexes. Fatal and palpable (superficial) tumors were analyzed by the Cox-Tarone binary regression method using the death time or the first palpation time (as applicable) as a surrogate for the tumor onset time. In the case of any particular tumor type where the study pathologist assigned the tumor in question as the cause of death of a subset of the rats and the rest of the rats were assumed to be dead of other competing risks, IARC cause of death analysis (Peto et al., 1980) was performed. Tumor types in which the cause of death was undetermined were treated as incidental for statistical purposes.

Sponsor's findings: In the males, several instances were noted in which the second control group showed significantly decreased incidences compared with the first control group. These included benign adenoma of the pituitary ($p = 0.0222$), benign follicular cell adenoma of the thyroid ($p = 0.0090$), and benign and malignant combined follicular cell lesions of the thyroid ($p = 0.0137$). These significant differences are indicative of background variations in these lesions in this strain of rodents. A statistically significant increase in the malignant lymphosarcoma incidence rate of the body, whole/cavity ($p = 0.0014$) was noted in rats given 300 mg/kg/day versus the combined control groups.

Also, males given 300 mg/kg/day showed statistically significant increase over the second control group only in benign follicular cell adenoma (p = 0.0023) and benign and malignant follicular cell lesions combined (p = 0.0049) of the thyroid. In none of these cases was there any significant trend, and the incidences in rats given 700 mg/kg/day were comparable to those of controls. Finally, two instances of statistically significant increases were observed in males given 300 or 700 mg/kg/day, consisting of benign interstitial cell tumor of the testis (p = 0.0051 for males given 300 mg/kg/day and p = 0.0048 for males given 700 mg/kg/day) and benign hemangioma of body, whole/cavity (p = 0.0476 for males given 300 mg/kg/day and p = 0.0005 for males given 700 mg/kg/day). In the first case, the positive trend was not significant for common tumors (p = 0.0168), but in second case, the positive trend was significant for common tumors (p = 0.0002). When the hemangioma incidences were examined only from the mesenteric lymph node, the positive trend was significant (p = 0.0001) for common tumors; so was the control versus 700 mg/kg/day dose level increased over control (p = 0.0005). The incidence rate at the 300 mg/kg/day group was not statistically significant (p = 0.1017) in this case compared to control.

A single instance (benign granular cell tumor of the brain) was noted in females in which the second control group showed a statistically significant increase over the first control group (p = 0.0448). Finally, benign endometrial stromal polyp showed a statistically significant increase in females given 100 mg/kg/day versus the combined controls (p = 0.0449). The trend was not significant in this case (p = 0.4398) and the other two treated groups had incidences comparable to the two control groups.

The results of the statistical analyses did not indicate any BI 10773-related effect in survival in either sex. No BI 10773-related increase in any neoplastic lesion was observed in the females. In males, the hemangioma (both from the mesenteric lymph nodes and body, whole/cavity) and interstitial cell tumor of the testis showed some statistically significant findings.

Text Table 3
Incidence of Selected Neoplastic Findings

	Sex	Males					Females				
	Group	1	2	3	4	5	1	2	3	4	5
	Dose Level (mg/kg/day)	0	0	100	300	700	0	0	100	300	700
Mesenteric Lymph Node											
	Number Examined	50	50	50	50	50	50	48	50	50	50
Hemangioma		3	0	1	4	9	0	1	0	0	0
Body Whole/Cavity											
	Number Examined	50	50	50	50	50	50	50	50	50	50
Hemangioma		3	0	2	5	9	0	1	0	0	0
Hemangiosarcoma		0	0	0	0	0	1	1	0	0	0
Testis											
	Number Examined	50	50	50	50	50	NA	NA	NA	NA	NA
Interstitial Cell Tumor		2	0	4	7	6	NA	NA	NA	NA	NA

NA = Not applicable.

2.2. Reviewer's analyses

To verify sponsor's analyses and to perform the 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. Three sets of survival and tumor analysis were done in the reviewer's analysis, set one including vehicle control 1 with the three treated groups, set two including vehicle control 2 and the three treated groups and set 3 including the combined vehicle control 1 & 2 with three treated groups.

2.2.1. Survival analysis

The survival distributions of animals in all five treatment groups (three treated groups and two vehicle control groups) were estimated by the Kaplan-Meier product limit method. The dose response relationship and homogeneity of survival distributions were tested using the Cox test (Cox, 1972). The inter-current mortality data are given in Tables 1A and 1B in the appendix for five treatment groups in males and females, respectively. The Kaplan-Meier curves for survival rate are given in Figures 1A and 1B in the appendix for five treatment groups in males and females, respectively. Results for the tests for dose response relationship and homogeneity of survivals, are given in Tables 2A and 2B in the appendix for three sets of data in males and females, respectively.

Reviewer's findings: The test results showed no statistically significant dose-response mortality and statistically significant difference in mortality in both females and males when compared separately with vehicle control 1, vehicle control 2 and the combined vehicle control groups. Also the test results showed no statistically significant difference in mortality in both females and males when compared between vehicle control 1 and 2. There were some differences between reviewer's and sponsor's survival rates and the differences may be caused by the different dates of starting the terminal killing.

2.2.2. Tumor data analysis

The tumor data were analyzed for dose response relationships and pair-wise comparisons of each of the two vehicle control groups and combined vehicle control groups with each of the treated groups were performed using the Poly-k method described in the paper of Bailer and Portier (1988) and Bieler and Williams (1993). One critical point for Poly-k test is the choice of the appropriate value of k. 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 of the tested tumor types are listed in Tables 3A (1), 3A (2), 3A (3) 3B (1), 3B (2) and 3B (3) in the appendix for three sets of data in males and females, respectively.

According to pharmacologist request, we have the following tumor combinations only in rat study:

- Hemangioma and hemagiosarcoma for the whole body (all sites) and also for the mesenteric lymph node.
- Skin and subcutis basal cell adenoma and carcinoma.
- Skin and subcutis squamous cell papilloma, carcinoma and keratoacanthoma.
- Lung bronchio-alveolar adenoma and carcinoma.
- Thyroid follicular cell b-adenoma and carcinoma
- Thyroid c-cell b-adenoma and carcinoma.
- Body, whole lymphosarcoma

In male mice study:

- Carcinoma and adenoma in Kidney

Multiple testing adjustment: Adjustment for the multiple dose response relationship testing was done using the criteria developed by Lin and Rahman (1998). The criteria recommend the use of a significance level α=0.025 for rare tumors and α=0.005 for common tumors for a submission with two species, and a significance level α=0.05 for rare tumors and α=0.01 for common tumors for a submission with only one species study in order to keep the false-positive rate at the nominal level of approximately 10%. A rare tumor is defined as one in which the spontaneous tumor rate is less than 1%. The adjustment for multiple pair-wise comparisons was done using the criteria developed by Haseman (1983) that recommends the use of significance level α=0.05 for rare

tumors and α=0.01 for common tumors, in order to keep the false-positive rate at the nominal level of approximately 10%.
It should be noted that the recommended test levels by Lin and Rahman for the adjustment of multiple testing were originally based on the result of a simulation and an empirical study using the Peto method for dose response relationship analysis. However, some later simulation results by Rahman and Lin (2008) indicate that the criteria apply equally well to the analysis using the poly-3 test.

Reviewer's findings: Following tumor types showed p-values less than or equal to 0.05 either tests for dose response relationship and/or pair-wise comparisons between control (each of two vehicle controls and the combined vehicle controls) and each of individual treated groups for three sets of analyses.

Tumor Types with P-Values ≤ 0.05 for Dose Response Relationship or Pair-wise Comparisons (each of vehicle controls or combined controls, low, medium and high dose groups)

			Vehicle Cont1 N=50	100 mg Low N=50	300 mg Med N=50	700 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Male										
	Body, Whole/Cav	B-Hemangioma	3	2	5	9	0.008	0.834	0.370	0.053
			[36]	[34]	[34]	[35]	.	.	.	.
		M-Histiocytic Sarcom	1	0	0	3	0.048	.	.	0.291
			[37]	[34]	[34]	[35]	.	.	.	.
		M-Lymphosarcoma	0	0	5	0	0.428	.	0.033	.
			[36]	[34]	[36]	[35]	.	.	.	.

			Vehicle Cont2 N=50	100 mg Low N=50	300 mg Med N=50	700 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Male										
	Body, Whole/Cav	B-Hemangioma	0	2	5	9	0.001	0.271	0.033	0.001
			[27]	[34]	[34]	[35]	.	.	.	.
		M-Histiocytic Sarcoma	0	0	0	3	0.018	.	.	0.125
			[27]	[34]	[34]	[35]	.	.	.	.
		M-Lymphosarcoma	0	0	5	0	0.491	.	0.037	.
			[27]	[34]	[36]	[35]	.	.	.	.
	Testis	B-Interstitial Cell tumor	0	4	7	6	0.057	0.070	0.008	0.014
			[27]	[34]	[34]	[35]	.	.	.	.
	Thyroid									
		B-Adenoma, Follicular	2	7	13	4	0.582	0.102	0.003	0.350
			[27]	[34]	[34]	[35]	.	.	.	.
		FOLLICULAR_CELL_ADEN NOMA+CARCINOMA	3	8	13	6	0.478	0.121	0.007	0.241
			[27]	[34]	[34]	[35]	.	.	.	.

			Combined	100 mg	300 mg	700 mg				
			Conts	Low	Med	High	P_Value	P_Value	P_Value	P_Value
	Organ Name	Tumor Name	N=100	N=50	N=50	N=50	Dos Resp	C vs. L	C vs. M	C vs. H
Male										
	Body, Whole/Cav	B-Hemangioma	3	2	5	9	0.001	0.567	0.089	0.002
			[63]	[34]	[34]	[35]	.	.	.	.
		M-Histiocytic Sarcoma	1	0	0	3	0.030	.	.	0.105
			[63]	[34]	[34]	[35]	.	.	.	.
		M-Lymphosarcoma	0	0	5	0	0.294	.	0.005	.
			[63]	[34]	[36]	[35]	.	.	.	.
	Testis	B-Interstitial Cell tumor	2	4	7	6	0.022	0.111	0.008	0.017
			[63]	[34]	[34]	[35]	.	.	.	.
	Thyroid									
		B-Adenoma, Follicula	11	7	13	4	0.681	0.435	0.021	0.801
			[63]	[34]	[34]	[35]	.	.	.	.
		FOLLICULAR_CELL_ADEN MA+CARCINOMA	13	8	13	6	0.583	0.447	0.048	0.654
			[63]	[34]	[34]	[35]	.	.	.	.
Female										
	Cervix									
		B-Polyp, Endometrial	0	3	0	1	0.416	0.044	.	0.336
			[67]	[39]	[35]	[33]	.	.	.	.

For vehicle control 1 vs. three treated groups:
Based on the criteria of adjustment for multiple testing of trends proposed by Lin and Rahman, the incidence of none of the above or any other tested tumor types in either sex was considered to have a statistically significant positive dose response relationship. Also based on the criteria by Haseman, the increased tumor incidences of Lymphosarcoma of body whole/cavity in medium dose group in male rats was considered to be statistically significant when compared to the vehicle control 1 because the p-value was less than 0.05.

For vehicle control 2 vs. three treated groups:
Based on the criteria of adjustment for multiple testing of trends proposed by Lin and Rahman, the incidence of none of the above or any other tested tumor types in females was considered to have a statistically significant positive dose response relationship. However the dose response relationship in the incidence of Hemangioma and Histiocytic sarcoma from body whole/cavity in male rats were considered to be statistically significant since the p-values were less than 0.025. Also based on the criteria by Haseman, the increased tumor incidences of Hemangioma of body whole/cavity and interstitial cell tumor of testis in high dose group and Hemangioma and Lymphosarcoma of body whole/cavity, interstitial cell tumor of testis, follicular adenoma and combined follicular adenoma and carcinoma of thyroid in medium dose group in male rats were considered to be statistically significant when compared to the vehicle control 2.

For combined vehicle controls vs. three treated groups:
Based on the criteria of adjustment for multiple testing of trends proposed by Lin and Rahman, the incidence of none of the above or any other tested tumor types in females was considered to have a statistically significant positive dose response relationship. However the dose response relationship in the incidence of

Hemangioma from body whole/cavity in male rats was considered to be statistically significant since the p-values were less than 0.005. Also based on the criteria by Haseman, the increased tumor incidence of Hemangioma of body whole/cavity in high dose group and Lymphosarcoma of body whole/cavity and interstitial cell tumor of testis in medium dose group in males and endometrial stromal polyp of cervix in low dose group in females were considered to be statistically significant when compared to the combined vehicle controls.
In females, vehicle control 2 showed a statistically significant increase over vehicle control 1 (p = 0.049) since the p-value is less than 0.05.

3. 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 two control groups. Two hundred and fifty Crl:CD1(ICR) mice of each sex were randomly allocated to each dose group of 50 animals. The dose levels for treated groups were 100, 300, and 1000 mg/kg/day for males and females for 24 months. In this review these dose groups would be referred to as the low, medium, and high dose group, respectively.

Male and female mice were assigned to study groups as follows.

Group	No. of Animals Male	No. of Animals Female	Dose Level (mg/kg/day)	Dose Concentration[a] (mg/mL)
Carcinogenicity Animals				
1 (Vehicle Control)[b]	50	50	0	0
2 (Vehicle Control)[b]	50	50	0	0
3 (Low)	50	50	100	10
4 (Mid)	50	50	300	30
5 (High)	50	50	1000	100
Toxicokinetic Animals[c]				
6 (Vehicle Control)[b]	12	12	0	0
7 (Low)	57	57	100	10
8 (Mid)	57	57	300	30
9 (High)	57	57	1000	100
10 (Sentinel Animals)[d]	25	25	e	e

a Animals were dosed at a volume of 10 mL/kg.
b Groups 1, 2, and 6 received vehicle control article only [0.5% (w/v) hydroxyethylcellulose in reverse osmosis water].
c Three extra animals/sex in Groups 6 through 9 were assigned to the study and used as replacement animals (dependent on survival). If not used for blood collection, all remaining animals were sacrificed and discarded after final blood collection.
d For selection of sentinel animals, the last four animals/sex (identified by temporary animal numbers) were used for the predose phase blood collection. The remaining 21 animals/sex were selected randomly. Five animals/sex served as replacement animals to compensate for possible mortality.
e Sentinel animals were not dosed.

3.1. Sponsor's analyses

3.1.1. Survival analysis

Survival data from the mouse study were analyzed by the sponsor using the same statistical methodologies that were used to analyze the survival data from the rat study. All statistical analysis was performed for males

and females separately.

Sponsor's findings: Kaplan-Meier product limit survival curves are presented in Figure 1 (males) and Figure 2 (females). The males in the two vehicle control groups were not statistically different from each other ($p > 0.1979$ by both the Cox-Tarone and Gehan-Breslow tests). Combining the two vehicle control groups and testing against the treated groups (Groups 3 through 5) gives a significant positive trend in mortality with $p > 0.0035$ by both the Cox-Tarone and Gehan-Breslow tests. This positive trend is due to the increased mortality in the group given 1000 mg/kg/day ($p \geq 0.0064$ by both the Cox-Tarone and Gehan-Breslow tests). For females, the second vehicle control group (Group 2) showed a significantly increased mortality ($p > 0.004$ by both the Cox-Tarone and Gehan-Breslow tests). The positive trend for the treated female groups versus the first vehicle control group was not significant ($p > 0.0690$ by both the Cox-Tarone and Gehan-Breslow tests). However, the group given 100 mg/kg/day in this case showed a significant reduction in mortality ($p \geq 0.0178$ by both the Cox-Tarone and Gehan-Breslow tests). A negative trend was observed for the treated females versus the second vehicle control group which was not significant. All three treated groups showed nonsignificant decreases in mortality rates versus the second vehicle control group.
Significantly increased mortality in males given 1000 mg/kg/day caused a significant positive trend as well. The females did not show any consistent pattern in mortality.

Figure 3: Kaplan-Meier plot of Survival in Male Mice

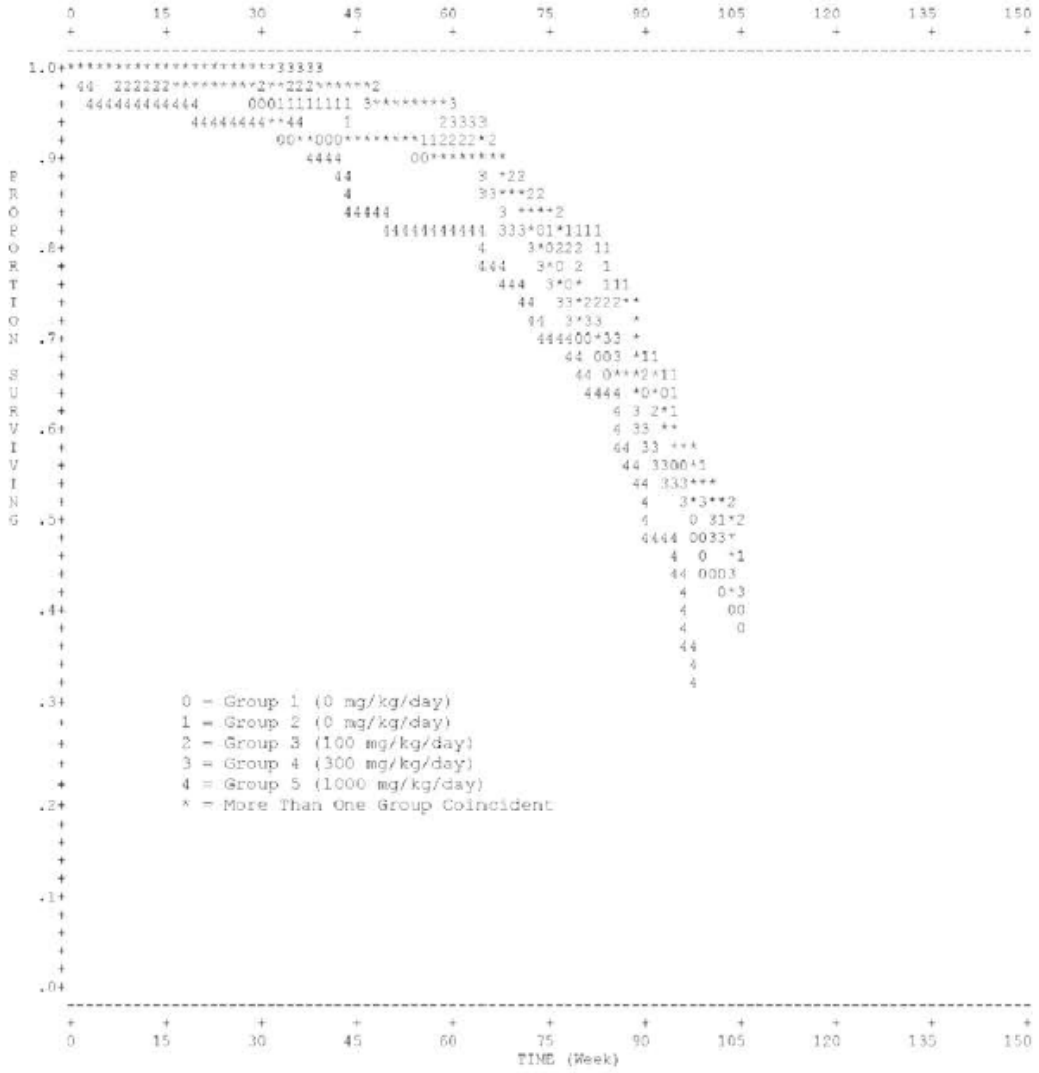
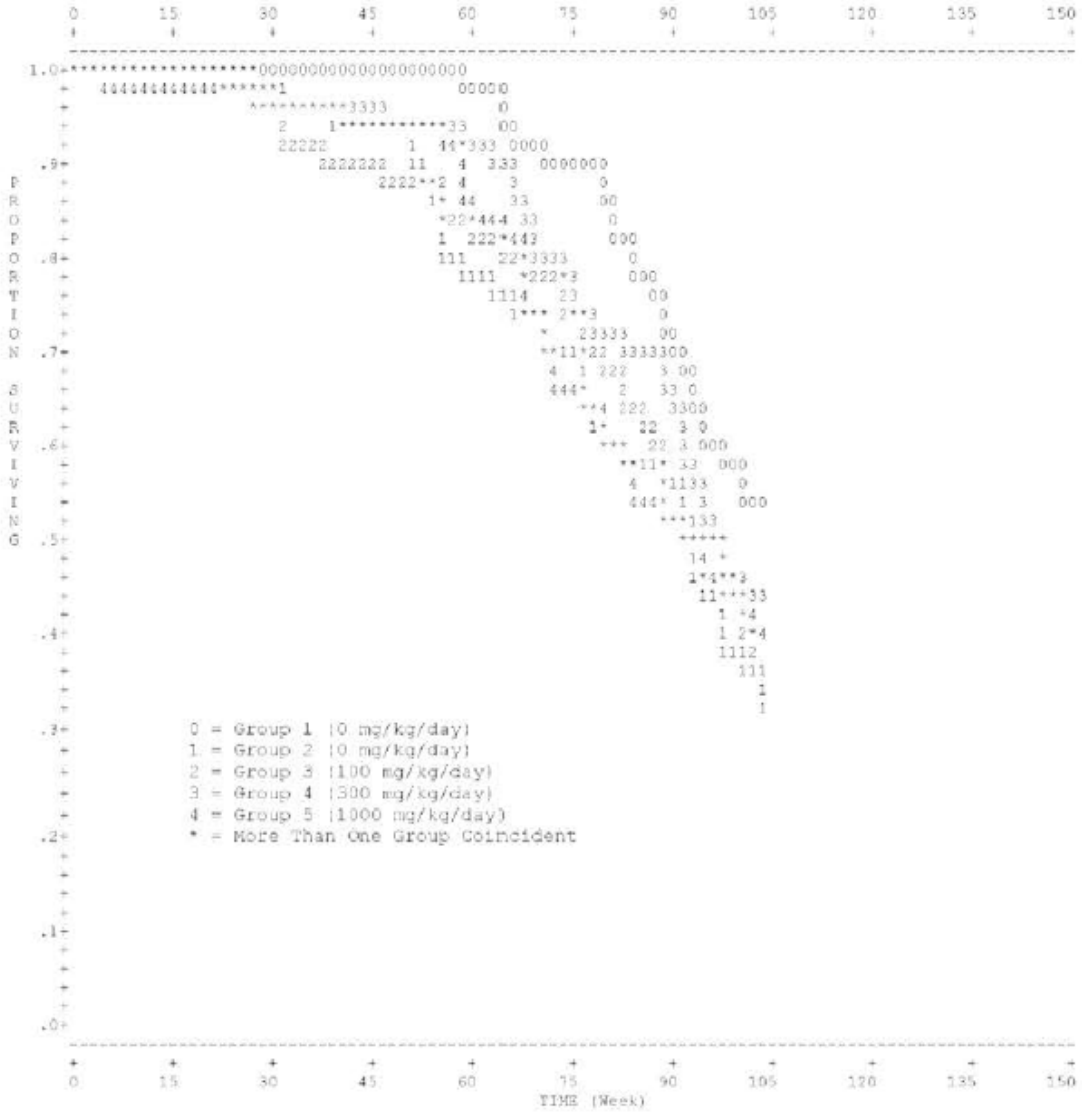


Figure 4: Kaplan-Meier plot of Survival in Female Mice



3.1.2. Tumor data analysis

Tumor data from mouse study were also analyzed by the sponsor using the same statistical methodologies that were used to analyze the tumor data from the rat study.

Sponsor's findings: the incidence of kidney tubular cell adenoma exhibited a significant positive trend (p = 0.0060) and a significant increase in the group given 1000 mg/kg/day (p = 0.0184) for rare tumors. On the other hand, the combined tubular cell adenoma and carcinoma incidences showed a significant positive trend (p = 0.0000) with a significant increase at 1000 mg/kg/day (p = 0.0001) for rare tumors. No other significantly increased neoplastic effect was noted in any other case in the males.

In the females, significant neoplastic findings of increased incidences of body, whole/cavity histiocytic sarcoma (p = 0.0037), bronchiolar-alveolar carcinoma of the lungs (p = 0.0182), and body, whole/cavity lymphosarcoma (p = 0.0194) were noted in the second control group compared to the first control group. Histiocytic sarcoma was significantly increased in the group given 100 mg/kg/day (p = 0.0064) versus the first vehicle control group. Lymphosarcoma was significantly increased in the group given 300 mg/kg/day (p = 0.0361) versus the first vehicle control, but not so versus the second control group.

The only significant neoplastic effect was in the kidney tubular cell tumors where the adenoma and combined adenoma and carcinoma incidences showed significant positive trend arising from a significant increase in males given 1000 mg/kg/day. In the case of the kidney tubular cell carcinoma alone, there was no significant positive trend but a significant increase at the 1000 mg/kg/day group. No other significant neoplastic effect was observed in the males. The females did not exhibit any consistent significant effect in any of the neoplastic findings in this study.

3.2. Reviewer's analyses

This reviewer independently performed survival and tumor data analyses for mouse study. For the mouse data analyses this reviewer used similar methodologies that she used to analyze the data from the rat study. Data used in this reviewer's analyses were provided by the sponsor electronically. Three sets of analysis comparing each of the vehicle controls and combined vehicle controls separately with the treated groups were done in the reviewer's analysis.

3.2.1. Survival analysis

The intercurrent mortality data are given in Tables 4A and 4B in the appendix for five treatment groups in males and females, respectively. The Kaplan-Meier curves for death rate are given in Figures 2A and 2B in the males and females, respectively. Results for the tests for dose response relationship and homogeneity of survivals, are given in Tables 5A and 5B in the appendix for three sets of data in males and females, respectively.

Reviewer's findings: The test showed a statistically significant dose-response in survival across either vehicle control 1, or vehicle control 2, or the combined vehicle controls and treated groups in males, respectively, and pair-wise differences between low dose group and vehicle control 1, between low dose group and vehicle control 2, between low dose group and the combined vehicle controls in survivals in females and between high dose group and vehicle control 1, between high dose group and vehicle control 2 and between high dose group and the combined vehicle controls in survivals in males. There were some differences between reviewer's and sponsor's survival rates and the differences may be caused by the different dates of starting the terminal killing.

3.2.2. Tumor data analysis

The tumor rates and the p-values of the tumor types tested for dose response relationship and pair-wise comparisons of control and treated groups are given in Table 6A (1), 6A (2), 6A (3), 6B (1), 6B (2) and 6B (3) in the appendix for three sets of data in males and females, respectively. As suggested by the reviewing pharmacologist Dr. Summan.

Reviewer's findings: Following tumor types showed p-values less than or equal to 0.05 either tests for dose response relationship or pair-wise comparisons between vehicle control 1 or vehicle control 2 or the combined vehicle controls and each of individual treated groups, respectively.

Tumor Types with P-Values ≤ 0.05 for Dose Response Relationship or Pair-wise Comparisons (Vehicle control 1, 2 & combined vehicle controls, low, medium and high dose groups)

					1000 m				
			Vehicle	100 mg	300 mg	g			
			Cont1	Low	Med	High	P_Value	P_Value	P_Value	P_Value
	Organ Name	Tumor Name	N=50	N=50	N=50	N=50	Dos Resp	C vs. L	C vs. M	C vs. H
Male										
	Kidney	B-Adenoma, Tubular	0	0	1	3	0.007	.	0.493	0.105
			[32]	[24]	[23]	[19]	.	.	.	.
		M-Carcinoma, Tubular	0	0	0	2	0.036	.	.	0.226
			[32]	[24]	[23]	[19]	.	.	.	.
		CARCINOMA+ADENOMA	0	0	1	5	< .0001	.	0.493	0.021
			[32]	[24]	[23]	[19]	.	.	.	.
Female										
	Body, Whole/Cav									
		M-Histiocytic Sarcom	1	6	3	6	0.035	0.023	0.277	0.023
			[25]	[22]	[23]	[18]	.	.	.	.
	Uterus									
		M-Carcinoma, Endomet	0	0	0	2	0.042	.	.	0.184
			[25]	[20]	[23]	[18]	.	.	.	.

						1000 m				
			Vehicle	100 mg	300 mg	g				
			Cont2	Low	Med	High	P_Value	P_Value	P_Value	P_Value
	Organ Name	Tumor Name	N=50	N=50	N=50	N=50	Dos Resp	C vs. L	C vs. M	C vs. H
Male										
	Kidney	B-Adenoma, Tubular	0	0	1	3	0.008	.	0.478	0.096
			[31]	[24]	[23]	[19]	.	.	.	.
		M-Carcinoma, Tubular	0	0	0	2	0.037	.	.	0.214
			[31]	[24]	[23]	[19]	.	.	.	.
		CARCINOMA+ADENOMA	0	0	1	5	<.0001	.	0.478	0.019
			[31]	[24]	[23]	[19]	.	.	.	.
Female										
	Uterus									
		M-Carcinoma, Endomet	0	0	0	2	0.044	.	.	0.237
			[23]	[20]	[23]	[18]	.	.	.	.

						1000 m				
			Combined	100 mg	300 mg	g				
			Conts	Low	Med	High	P_Value	P_Value	P_Value	P_Value
	Organ Name	Tumor Name	N=100	N=50	N=50	N=50	Dos Resp	C vs. L	C vs. M	C vs. H
Male										
	Kidney	B-Adenoma, Tubular	0	0	1	3	0.002	.	0.320	0.028
			[63]	[24]	[23]	[19]	.	.	.	.
		M-Carcinoma, Tubular	0	0	0	2	0.021	.	.	0.094
			[63]	[24]	[23]	[19]	.	.	.	.
		CARCINOMA+ADENOMA	0	0	1	5	<.0001	.	0.320	0.002
			[63]	[24]	[23]	[19]	.	.	.	.

Female										
	Uterus									
		M-Carcinoma, Endomet	0	0	0	2	0.026	.	.	0.087
			[48]	[20]	[23]	[18]	.	.	.	.

Based on the criteria of adjustment for multiple testing of trends by Lin and Rahman, the dose response relationship in the incidence of tubular adenoma and combined tubular adenoma and carcinoma of kidney for vehicle control 1, vehicle control 2 and combined vehicle controls with three treated groups, and of tubular carcinoma of kidney for the combined vehicle controls with three treated groups in male mice were considered to be statistically significant since the p-values were less than 0.025. Also based on the criteria of Haseman, the increased tumor incidence of tubular adenoma of kidney in high dose group in male mice was considered to be statistically significant when compared to the combined vehicle controls because the p-value is less than 0.05. In addition, the increased tumor incidence of combined tubular adenoma and carcinoma of kidney in high dose group in male mice was considered to be statistically significant when compared to the vehicle control 1, vehicle control 2 and combined vehicle controls because the p-value is less than 0.05.

4. Evaluation of validity of the designs of the female rat and mouse studies

As having been noted, the tumor data analyses from female rat study including vehicle control 1 or vehicle control 2 with three treated groups and female mouse study including vehicle control 1, or vehicle control 2 or combined vehicle controls showed no statistically significant dose-response relationship in any tested single tumor type. Before drawing any conclusion regarding the carcinogenic or non-carcinogenic potential of the drug in rats and mice, it is important to look into the following two issues, as have been pointed out in the paper by Haseman (1984).

(i) Were enough animals exposed, for a sustained amount of time, to the risk of late developing tumors?
(ii) Were dose levels high enough to pose a reasonable tumor challenge to the animals?
There is no consensus among experts regarding the number of animals and length of time at risk, although most carcinogenicity studies are designed to run for two years with fifty animals per treatment group. The following are some rules of thumb regarding these two issues as suggested by experts in this field:

Haseman (1985) has done an investigation on the first issue. He gathered data from 21 studies using Fischer 344 rats and B6C3Fl mice conducted at the National Toxicology Program (NTP). It was found that, on the average, approximately 50% of the animals in the high dose group survived the two-year study period. Also, in a personal communication with Dr. Karl Lin of Division of Biometrics-6, Haseman suggested that, as a rule of thumb, a 50% survival of 50 initial animals or 20 to 30 animals still alive in the high dose group, between weeks 80-90, would be consider as a sufficient number and adequate exposure. In addition Chu, Cueto and Ward (1981), suggested that" to be considered adequate, an experiment that has not shown a chemical to be carcinogenic should have groups of animals with greater than 50% survival at one-year."

It appears, from these three sources that the proportions of survival at 52 weeks, 80-90 weeks, and two years are of interest in determining the adequacy of exposure and number of animals at risk.

Regarding the question of adequate dose levels, it is generally accepted that the high dose should be close to the maximum tolerated dose (MTD). In the paper of Chu, Cueto and Ward (1981), the following criteria are mentioned for dose adequacy. A high dose is considered as close to MTD if any of the criteria is met.

(i) "A dose is considered adequate if there is a detectable loss in weight gain of up to 10% in a dosed group relative to the controls."

(ii) "The administered dose is also considered an MTD if dosed animals exhibit clinical signs or severe histopathologic toxic effects attributed to the chemical."

(iii) "In addition, doses are considered adequate if the dosed animals show a slight increased mortality compared to the controls."
We will now investigate the validity of the BI 10773 female rat and mouse studies, in the light of the above guidelines.

4.1. Female Rat Study

The following is the summary of survival data of rats in the high dose groups:

Percentage of survival in the high dose group at the end of Weeks 52, 78, and 91

	Percentage of survival		
	End of 52 weeks	End of 78 weeks	End of 91 weeks
Female	92%	70%	56%

Based on the survival criterion Haseman proposed, it could be concluded that enough female rats were exposed to the high dose for a sufficient amount of time.

The following table shows the percent difference in mean body weight gain from the concurrent combined control, defined as

$$\text{Percent difference} = \frac{(\text{Final BW} - \text{Baseline BW})_{\text{Treated}} - (\text{Final BW} - \text{Baseline BW})_{\text{Control}}}{(\text{Final BW} - \text{Baseline BW})_{\text{Control}}} \times 100$$

Percent Difference in Mean body Weight Gain From Vehicle control 1

Female		
100mg	300 mg	700 mg
-16.4	-25.2	-30

Percent Difference in Mean body Weight Gain From Vehicle control 2

Female		
100mg	300 mg	700 mg
-19.6	-28.1	-32.7

Therefore, relative to the vehicle control 1 or vehicle control 2, there had been more than 15% decrement in body weight gain in all three treated groups in females.

The mortality rates at the end of the experiment were as follows:

Mortality Rates at the End of the Experiment

	Cont. 1	**100 mg**	**300mg**	**700 mg**
Female	38%	36%	32%	38%

	Cont. 2	**100mg**	**300mg**	**700 mg**
Female	48%	38%	32%	38%

This shows that the morality rate of in the high dose group in females is 10% lower than the vehicle control 2 but the same as vehicle control 1. Thus, from the body weight gain and mortality data it can be concluded that for females the used high dose level might not have reached or exceeded the MTD in neither of vehicle control 2 and vehicle control 1 data set. For a final determination of the adequacy of the doses used, other clinical signs and histopathological toxic effects must be considered.

4.2. Female Mouse Study

The following is the summary of survival data of rats in the high dose groups:

Percentage of survival in the high dose group at the end of Weeks 52, 78, and 91

	Percentage of survival		
	End of 52 weeks	End of 78 weeks	End of 91 weeks
Female	80%	54%	42%

Based on the survival criterion Haseman proposed, it could be concluded that enough female mice were exposed to the high dose for a sufficient amount of time.

The following table shows the percent difference in mean body weight gain from the concurrent combined control, defined as

$$\text{Percent difference} = \frac{(\text{Final BW} - \text{Baseline BW})_{\text{Treated}} - (\text{Final BW} - \text{Baseline BW})_{\text{Control}}}{(\text{Final BW} - \text{Baseline BW})_{\text{Control}}} \times 100$$

Percent Difference in Mean body Weight Gain From Vehicle Control 1

Female		
100mg	300 mg	1000 mg
-3.9	-10.4	-15.8

Percent Difference in Mean body Weight Gain From Vehicle Control 2

Female		
100mg	300 mg	1000 mg
13.5	5.8	-0.65

Percent Difference in Mean body Weight Gain
From combined vehicle controls

Female		
100mg	300 mg	1000 mg
4.14	-2.96	-8.88

Therefore, relative to the control, there had been more than 10% decrement in body weight gain in high dose groups in females only in the vehicle control 1 data set not for the vehicle control 2 data set or the combined vehicle controls data set.

The mortality rates at the end of the experiment were as follows:

Mortality Rates at the End of the Experiment

	Cont.1	**100 mg**	**300 mg**	**1000mg**
Female	50%	68%	64%	56%

Mortality Rates at the End of the Experiment

	Cont.2	**100 mg**	**300 mg**	**1000mg**
Female	66%	68%	64%	56%

Mortality Rates at the End of the Experiment

Combined Cont.1 &2		**100 mg**	**300 mg**	**1000mg**
Female	58%	68%	64%	56%

This shows that the morality rate of in the high dose group in females is 6% higher than the vehicle control 1 but lower than the vehicle control 2 and the combined vehicle controls. Thus, from the body weight gain and mortality data it can be concluded that for males the used high dose level might have reached or exceeded the MTD only for vehicle control 1 data set but not for vehicle control 2 data set or the combined vehicle control data set. For a final determination of the adequacy of the doses used, other clinical signs and histopathological toxic effects must be considered.

5. Summary

In this submission the sponsor included reports of two animal carcinogenicity studies, one in rats and one in mice. These studies were intended to further assess the carcinogenic potential of BI 10773 in rats and mice, when administered daily via oral gavage to mice and rats for at least 104 weeks.

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 two control groups. Two hundred and fifty Crl:WI(Han)rats of each sex were randomly allocated to treated and control groups. There were 50 animals per sex in each of groups. Treated animals each received dose preparations 100, 300, or 700 mg/kg.

The test results showed no statistically significant dose-response mortality and statistically significant difference in mortality in both females and males when compared separately with the vehicle control 1 or vehicle control 2 or the combined vehicle controls. Also the test results showed no statistically significant difference in mortality in both females and males when compared between vehicle control 1 and 2.
For vehicle control 1 vs. three treated groups:
Based on the criteria of adjustment for multiple testing of trends proposed by Lin and Rahman, the incidence of none of the above or any other tested tumor types in either sex was considered to have a statistically significant positive dose response relationship. Also based on the criteria by Haseman, the increased tumor incidences of Lymphosarcoma of whole boday cavity in medium dose group in male rats was considered to be statistically significant when compared to the vehicle control 1 because the p-value was less than 0.05.

For vehicle control 2 vs. three treated groups:
Based on the criteria of adjustment for multiple testing of trends proposed by Lin and Rahman, the incidence of none of the above or any other tested tumor types in females was considered to have a statistically significant positive dose response relationship. However the dose response relationship in the incidence of Hemangioma and Histiocytic sarcoma from body whole/cavity in male rats were considered to be statistically significant since the p-values were less than 0.025. Also based on the criteria by Haseman, the increased tumor incidences of Hemangioma of whole boday cavity and interstitial cell tumor of testis in high dose group and Hemangioma and Lymphosarcoma of body whole/cavity, interstitial cell tumor of testis, follicular adenoma and combined follicular adenoma and carcinoma of thyroid in medium dose group in male rats were considered to be statistically significant when compared to the vehicle control 2.

For combined vehicle controls vs. three treated groups:
Based on the criteria of adjustment for multiple testing of trends proposed by Lin and Rahman, the incidence of none of the above or any other tested tumor types in females was considered to have a statistically significant positive dose response relationship. However the dose response relationship in the incidence of Hemangioma from body whole/cavity in male rats was considered to be statistically significant since the p-values were less than 0.005. Also based on the criteria by Haseman, the increased tumor incidence of Hemangioma of body whole/cavity in high dose group and Lymphosarcoma of body whole/cavity and interstitial cell tumor of testis in medium dose group in males and endometrial stromal polyp of cervix in low dose group in females were considered to be statistically significant when compared to the combined vehicle controls.

As having been noted, the tumor data analyses from female rat study including vehicle control 1 or vehicle control 2 with three treated groups showed no statistically significant dose-response relationship in any tested single tumor type. In females, vehicle control 2 showed a statistically significant increase over vehicle control 1 ($p = 0.049$) since the p-value is less than 0.05.

From the body weight gain and mortality data it can be concluded that for females the used high dose level might not have reached or exceeded the MTD in neither of vehicle control 2 and vehicle control 1 data set. For a final determination of the adequacy of the doses used, other clinical signs and histopathological toxic effects must be considered.

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 two control groups. Two hundred and fifty Crl:CD1(ICR) mice of each sex were randomly allocated to each dose group of 50 animals. The dose levels for treated groups were 100, 300, and 1000 mg/kg/day for males and females for 24 months. In this review these dose groups would be referred to as the low, medium, and high dose group, respectively.
The test showed a statistically significant dose-response in survival across either vehicle control 1, or vehicle control 2, or the combined vehicle controls and treated groups in males, respectively, and pair-wise differences between low dose group and vehicle control 1, between low dose group and vehicle control 2, between low dose group and the combined vehicle controls in survivals in females and between high dose group and vehicle control 1, between high dose group and vehicle control 2 and between high dose group and the combined vehicle controls in survivals in males. There were some differences between reviewer's and sponsor's survival rates and the differences may be caused by the different dates of starting the terminal killing. Also the test results showed no statistically significant difference in mortality in both females and males when compared between vehicle control 1 and 2.
Based on the criteria of adjustment for multiple testing of trends by Lin and Rahman, the dose response relationship in the incidence of tubular adenoma and combined tubular adenoma and carcinoma of kidney for vehicle control 1, vehicle control 2 and combined vehicle controls with three treated groups, and of tubular carcinoma of kidney for the combined vehicle controls with three treated groups in male mice were considered to be statistically significant since the p-values were less than 0.025. Also based on the criteria of Haseman, the increased tumor incidence of tubular adenoma of kidney in high dose group in male mice was considered to be statistically significant when compared to the combined vehicle controls because the p-value is less than 0.05. In addition, the increased tumor incidence of combined tubular adenoma and carcinoma of kidney in high dose group in male mice was considered to be statistically significant when compared to the vehicle control 1, vehicle control 2 and combined vehicle controls because the p-value is less than 0.05.

Min Min, Ph.D.
Mathematical Statistician

Concur: Karl Lin, Ph.D.
Team Leader, Biometrics-6

cc:
Archival NDA 20-4629
Dr. Summan
Dr. Tiwari
Dr. Nevius
Lillian Patrician
Dr. Tsong
Dr. Lin
Dr. Min

6. Appendix

Table 1A: Intercurrent Mortality Rate
Male Rats

Week	VEHICLE CONTROL1 NO.OF DEATH	PERCENT	VEHICLE CONTROL2 NO.OF DEATH	PERCENT	LOW NO.OF DEATH	PERCENT	MEDIUM NO.OF DEATH	PERCENT	HIGH NO.OF DEATH	PERCENT
0-52	.	.	3	6.0%	1	2.0%	2	4.0%	1	2.0%
53-78	3	6.0%	3	12.0%	4	10.0%	3	10.0%	7	16.0%
79-92	7	20.0%	7	26.0%	3	16.0%	3	16.0%	4	24.0%
93-104	9	38.0%	4	34.0%	7	30.0%	3	22.0%	5	34.0%
Term. Sac.	31	100.0%	33	100.0%	35	100.0%	39	100.0%	33	100.0%

Table 1B: Intercurrent Mortality Rate
Female Rats

Week	VEHICLE_CONTROL1 NO.OF DEATH	PERCENT	VEHICLE_CONTROL2 NO.OF DEATH	PERCENT	LOW NO.OF DEATH	PERCENT	MEDIUM NO.OF DEATH	PERCENT	HIGH NO.OF DEATH	PERCENT
0-52	1	2.0%	3	6.0%	2	4.0%	3	6.0%	3	6.0%
53-78	6	14.0%	7	20.0%	3	10.0%	5	16.0%	5	16.0%
79-92	6	26.0%	6	32.0%	4	18.0%	4	24.0%	3	22.0%
93-104	6	38.0%	8	48.0%	9	36.0%	4	32.0%	8	38.0%
Term. Sac.	31	100.0%	26	100.0%	32	100.0%	34	100.0%	31	100.0%

Table 2A: Intercurrent Mortality Comparison
Male Rats

Test	P-Value (across four groups)	P-Value (vehicle_control1 vs low)	P-Value (vehicle_control1 vs medium)	P-Value (vehicle_control1 vs high)
Dose Response	0.6066	0.5956	0.3856	0.8869
Homogeneity	0.2925	0.3598	0.1115	0.9403

Test	P-Value (across four groups)	P-Value (vehicle_control2 vs low)	P-Value (vehicle_control2 vs medium)	P-Value (vehicle_control2 vs high)
Dose Response	0.8428	0.7489	0.4860	0.9972
Homogeneity	0.6124	0.5693	0.1865	0.8606

Test	P-Value (across four groups)	P-Value (combined vehicle_controls vs low)	P-Value (combined vehicle_controls vs medium)	P-Value (combined vehicle_controls vs high)
Dose Response	0.5627	0.6327	0.3656	0.9350
Homogeneity	0.3227	0.3990	0.1021	0.9563

Table 2B: Intercurrent Mortality Comparison
Female Rats

Test	P-Value (across four groups)	P-Value (vehicle_control1 vs low)	P-Value (vehicle_control1 vs medium)	P-Value (vehicle_control1 vs high)
Dose Response	0.4587	0.8250	0.7497	0.9890
Homogeneity	0.6306	0.7135	0.7087	0.9820

Test	P-Value (across four groups)	P-Value (vehicle_control2 vs low)	P-Value (vehicle_control2 vs medium)	P-Value (vehicle_control2 vs high)
Dose Response	0.2813	0.3422	0.3203	0.5120
Homogeneity	0.1955	0.1358	0.1550	0.3144

Test	P-Value (across four groups)	P-Value (combined vehicle_controls vs low)	P-Value (combined vehicle_controls vs medium)	P-Value (combined vehicle_controls vs high)
Dose Response	0.2557	0.5058	0.4519	0.7159
Homogeneity	0.2415	0.2942	0.3032	0.5687

Table 3A (1): Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Rats (Vehicle control1, low, medium and high dose goups)

		Vehicle	100 mg	300 mg	700 mg				
		Cont1	Low	Med	High	P_Value	P_Value	P_Value	P_Value
Organ Name	Tumor Name	N=50	N=50	N=50	N=50	Dos Resp	C vs. L	C vs. M	C vs. H
Adipose Tissue		(50)	(50)	(49)	(50)	.	.	.	.
	M-Schwannoma, Malignant	0	0	0	1	0.252	.	.	0.494
		[36]	[34]	[34]	[35]	.	.	.	.
Adrenal, Cortex		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma	2	0	0	0	1.000	1.000	1.000	1.000
		[36]	[34]	[34]	[35]	.	.	.	.
	M-Carcinoma	0	1	0	0	0.741	0.511	.	.
		[36]	[34]	[34]	[35]	.	.	.	.
Adrenal, Medull		(50)	(50)	(50)	(50)	.	.	.	.
	B-Pheochromocytoma	2	3	1	3	0.392	0.522	0.884	0.489
		[36]	[35]	[34]	[35]	.	.	.	.
	M-Complex Malignant Pheochro	1	0	0	0	1.000	1.000	1.000	1.000
		[36]	[34]	[34]	[35]	.	.	.	.
Body, Whole/Cav		(50)	(50)	(50)	(50)	.	.	.	.
	B-Hemangioma	3	2	5	9	0.008	0.834	0.370	0.053
		[36]	[34]	[34]	[35]	.	.	.	.
	M-Hibernoma	0	2	1	0	0.730	0.259	0.511	.
		[36]	[35]	[35]	[35]	.	.	.	.
	M-Histiocytic Sarcoma	1	0	0	3	0.048	1.000	1.000	0.291
		[37]	[34]	[34]	[35]	.	.	.	.
	M-Lymphosarcoma	0	0	5	0	0.428	.	0.033	.
		[36]	[34]	[36]	[35]	.	.	.	.
Brain		(50)	(50)	(50)	(50)	.	.	.	.
	B-Granular Cell Tumor	1	1	0	0	0.934	0.764	1.000	1.000
		[36]	[34]	[34]	[35]	.	.	.	.
	B-Meningioma	0	1	0	0	0.743	0.511	.	.
		[36]	[35]	[34]	[35]	.	.	.	.
Duodenum		(49)	(50)	(50)	(50)	.	.	.	.
	M-Carcinoma	1	0	0	0	1.000	1.000	1.000	1.000
		[36]	[34]	[34]	[35]	.	.	.	.
	M-Leiomyosarcoma	0	1	0	0	0.741	0.511	.	.
		[36]	[34]	[34]	[35]	.	.	.	.
Epididymis		(50)	(50)	(50)	(50)	.	.	.	.
	B-Benign Mesothelioma	0	1	0	0	0.741	0.511	.	.
		[36]	[34]	[34]	[35]	.	.	.	.
	M-Schwannoma, Malignant	1	0	0	0	1.000	1.000	1.000	1.000
		[36]	[34]	[34]	[35]	.	.	.	.

Table 3A (1) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Rats (Vehicle control1, low, medium and high dose goups)

Organ Name	Tumor Name	Vehicle Cont1 N=50	100 mg Low N=50	300 mg Med N=50	700 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Gl, Preputial		(50)	(50)	(50)	(50)	.	.	.	.
	M-Schwannoma, Malignant	0	1	0	0	0.741	0.511	.	.
		[36]	[34]	[34]	[35]	.	.	.	.
Gl, Zymbal's		(50)	(50)	(50)	(50)	.	.	.	.
	M-Carcinoma	0	0	1	0	0.500	.	0.511	.
		[36]	[34]	[35]	[35]	.	.	.	.
Jejunum		(48)	(50)	(50)	(49)	.	.	.	.
	M-Carcinoma	1	0	0	0	1.000	1.000	1.000	1.000
		[36]	[34]	[34]	[35]	.	.	.	.
Kidney		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Tubule Cell	0	0	0	1	0.252	.	.	0.494
		[36]	[34]	[34]	[35]	.	.	.	.
	B-Lipoma	1	0	0	0	1.000	1.000	1.000	1.000
		[36]	[34]	[34]	[35]	.	.	.	.
Liver		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Hepatocellular	0	0	1	0	0.496	.	0.506	.
		[36]	[34]	[34]	[35]	.	.	.	.
Mammary, Male		(49)	(50)	(50)	(49)	.	.	.	.
Mesentery		(50)	(50)	(50)	(50)	.	.	.	.
	M-Schwannoma, Malignant	0	0	0	1	0.252	.	.	0.494
		[36]	[34]	[34]	[35]	.	.	.	.
Muscle, Bi Fem		(50)	(50)	(50)	(50)	.	.	.	.
	B-Granular Cell Tumor	1	0	0	0	1.000	1.000	1.000	1.000
		[36]	[34]	[34]	[35]	.	.	.	.
Pancreas		(50)	(50)	(50)	(49)	.	.	.	.
	B-Adenoma, Islet Cell	0	1	0	1	0.310	0.511	.	0.494
		[36]	[34]	[34]	[35]	.	.	.	.
Parathyroid		(46)	(48)	(49)	(48)	.	.	.	.
	B-Adenoma	0	0	1	0	0.496	.	0.506	.
		[36]	[34]	[34]	[35]	.	.	.	.

Table 3A (1) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Rats (Vehicle control1, low, medium and high dose goups)

		Vehicle	100 mg	300 mg	700 mg				
		Cont1	Low	Med	High	P_Value	P_Value	P_Value	P_Value
Organ Name	Tumor Name	N=50	N=50	N=50	N=50	Dos Resp	C vs. L	C vs. M	C vs. H
Pituitary		(50)	(50)	(48)	(50)	.	.	.	.
	B-Adenoma	27	14	15	21	0.679	0.998	0.994	0.909
		[41]	[36]	[36]	[40]	.	.	.	.
Prostate		(49)	(50)	(50)	(50)	.	.	.	.
	M-Schwannoma, Malignant	0	0	1	0	0.496	.	0.506	.
		[36]	[34]	[34]	[35]	.	.	.	.
Seminal Vesicle		(49)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma	0	0	1	0	0.496	.	0.506	.
		[36]	[34]	[34]	[35]	.	.	.	.
Skin/Subcutis		(50)	(50)	(50)	(50)	.	.	.	.
	B-Basal Cell Tumor	0	0	0	1	0.252	.	.	0.494
		[36]	[34]	[34]	[35]	.	.	.	.
	B-Fibroma	1	0	0	0	1.000	1.000	1.000	1.000
		[36]	[34]	[34]	[35]	.	.	.	.
	B-Keratoacanthoma	0	0	0	1	0.252	.	.	0.494
		[36]	[34]	[34]	[35]	.	.	.	.
	B-Papilloma, Squamous Cell	1	0	0	0	1.000	1.000	1.000	1.000
		[36]	[34]	[34]	[35]	.	.	.	.
Skin/Subcutis	M-Carcinoma, Basal Cell	1	0	0	1	0.442	1.000	1.000	0.747
		[36]	[34]	[34]	[35]	.	.	.	.
	M-Carcinoma, Squamous Cell	1	0	0	0	1.000	1.000	1.000	1.000
		[36]	[34]	[34]	[35]	.	.	.	.
	M-Fibrosarcoma	0	0	2	0	0.500	.	0.253	.
		[36]	[34]	[34]	[35]	.	.	.	.
	M-Sarcoma	1	0	0	0	1.000	1.000	1.000	1.000
		[37]	[34]	[34]	[35]	.	.	.	.
	SQUAMOUS_CELL_CARCINOMA+PAPIL	2	0	0	1	0.641	1.000	1.000	0.875
		[36]	[34]	[34]	[35]	.	.	.	.
Spinal Cord		(50)	(50)	(50)	(50)	.	.	.	.
	B-Oligodendroglioma	0	0	0	1	0.252	.	.	0.494
		[36]	[34]	[34]	[35]	.	.	.	.
Stomach, Gl		(50)	(50)	(50)	(50)	.	.	.	.
	M-Schwannoma, Malignant	0	0	0	1	0.252	.	.	0.494
		[36]	[34]	[34]	[35]	.	.	.	.
Stomach, Nongl		(50)	(49)	(50)	(50)	.	.	.	.
Testis		(50)	(50)	(50)	(50)	.	.	.	.
	B-Interstitial Cell Tumor	2	4	7	6	0.097	0.360	0.084	0.125
		[36]	[34]	[34]	[35]	.	.	.	.

Table 3A (1) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Rats (Vehicle control1, low, medium and high dose goups)

Organ Name	Tumor Name	Vehicle Cont1 N=50	100 mg Low N=50	300 mg Med N=50	700 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Thymus		(49)	(47)	(50)	(50)	.	.	.	.
	B-Thymoma	4	0	1	0	0.983	1.000	0.972	1.000
		[37]	[34]	[34]	[35]	.	.	.	.
Thyroid		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, C-cell	7	8	6	6	0.679	0.519	0.740	0.694
		[37]	[34]	[35]	[35]	.	.	.	.
	B-Adenoma, Follicular Cell	9	7	13	4	0.895	0.824	0.249	0.963
		[36]	[34]	[34]	[35]	.	.	.	.
	C_CELL_ADENOMA+CARCINOMA	8	8	7	6	0.740	0.628	0.730	0.782
		[37]	[34]	[35]	[35]	.	.	.	.
	FOLLICULAR_CELL_ADENOMA+CARCI	10	8	13	6	0.819	0.816	0.337	0.910
		[36]	[34]	[34]	[35]	.	.	.	.
	M-Carcinoma, C-cell	1	0	1	0	0.748	1.000	0.759	1.000
		[36]	[34]	[34]	[35]	.	.	.	.
	M-Carcinoma, Follicular Cell	2	1	0	2	0.429	0.888	1.000	0.683
		[36]	[34]	[34]	[35]	.	.	.	.

Numbers with parentheses are number of the animals with organs examined and also usable

Numbers with brackets are survival-adjusted group size

Numbers are the tumor bearing animals

Table 3A (2): Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Rats (Vehicle control2, low, medium and high dose goups)

Organ Name	Tumor Name	Vehicle Cont2 N=50	100 mg Low N=50	300 mg Med N=50	700 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Adipose Tissue		(50)	(50)	(49)	(50)	.	.	.	.
	M-Schwannoma, Malignant	0	0	0	1	0.269	.	.	0.506
		[27]	[34]	[34]	[35]	.	.	.	.
Adrenal, Cortex		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma	3	0	0	0	1.000	1.000	1.000	1.000
		[27]	[34]	[34]	[35]	.	.	.	.
	M-Carcinoma	0	1	0	0	0.792	0.523	.	.
		[27]	[34]	[34]	[35]	.	.	.	.
Adrenal, Medull		(50)	(50)	(50)	(50)	.	.	.	.
	B-Pheochromocytoma	1	3	1	3	0.328	0.344	0.770	0.317
		[27]	[35]	[34]	[35]	.	.	.	.
Body, Whole/Cav		(50)	(50)	(50)	(50)	.	.	.	.
	B-Hemangioma	0	2	5	9	0.001	0.271	0.033	0.001
		[27]	[34]	[34]	[35]	.	.	.	.
	B-Hibernoma	1	0	0	0	1.000	1.000	1.000	1.000
		[27]	[34]	[34]	[35]	.	.	.	.
	M-Hibernoma	0	2	1	0	0.778	0.271	0.523	.
		[27]	[35]	[35]	[35]	.	.	.	.
	M-Histiocytic Sarcoma	0	0	0	3	0.018	.	.	0.125
		[27]	[34]	[34]	[35]	.	.	.	.
	M-Lymphosarcoma	0	0	5	0	0.491	.	0.037	.
		[27]	[34]	[36]	[35]	.	.	.	.
Brain		(50)	(50)	(50)	(50)	.	.	.	.
	B-Granular Cell Tumor	1	1	0	0	0.958	0.770	1.000	1.000
		[27]	[34]	[34]	[35]	.	.	.	.
	B-Meningioma	0	1	0	0	0.794	0.523	.	.
		[27]	[35]	[34]	[35]	.	.	.	.
Colon		(50)	(50)	(50)	(49)	.	.	.	.
Duodenum		(49)	(50)	(50)	(50)	.	.	.	.
	M-Leiomyosarcoma	0	1	0	0	0.792	0.523	.	.
		[27]	[34]	[34]	[35]	.	.	.	.
Epididymis		(50)	(50)	(50)	(50)	.	.	.	.
	B-Benign Mesothelioma	0	1	0	0	0.792	0.523	.	.
		[27]	[34]	[34]	[35]	.	.	.	.
Gl, Preputial		(50)	(50)	(50)	(50)	.	.	.	.
	M-Schwannoma, Malignant	0	1	0	0	0.792	0.523	.	.
		[27]	[34]	[34]	[35]	.	.	.	.

Table 3A (2) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Rats (Vehicle control2, low, medium and high dose goups)

Organ Name	Tumor Name	Vehicle Cont2 N=50	100 mg Low N=50	300 mg Med N=50	700 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Gl, Zymbal's		(50)	(50)	(50)	(50)	.	.	.	.
	M-Carcinoma	1	0	1	0	0.785	1.000	0.770	1.000
		[27]	[34]	[35]	[35]	.	.	.	.
Heart		(50)	(50)	(50)	(50)	.	.	.	.
	M-Endocardial Schwannoma	2	0	0	0	1.000	1.000	1.000	1.000
		[27]	[34]	[34]	[35]	.	.	.	.
Kidney		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Tubule Cell	0	0	0	1	0.269	.	.	0.506
		[27]	[34]	[34]	[35]	.	.	.	.
	M-Carcinoma, Tubule Cell	1	0	0	0	1.000	1.000	1.000	1.000
		[27]	[34]	[34]	[35]	.	.	.	.
Liver		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Hepatocellular	0	0	1	0	0.531	.	0.518	.
		[27]	[34]	[34]	[35]	.	.	.	.
Mammary, Male		(50)	(50)	(50)	(49)	.	.	.	.
	B-Fibroadenoma	1	0	0	0	1.000	1.000	1.000	1.000
		[27]	[34]	[34]	[35]	.	.	.	.
Mesentery		(50)	(50)	(50)	(50)	.	.	.	.
	M-Schwannoma, Malignant	0	0	0	1	0.269	.	.	0.506
		[27]	[34]	[34]	[35]	.	.	.	.
Pancreas		(50)	(50)	(50)	(49)	.	.	.	.
	B-Adenoma, Islet Cell	2	1	0	1	0.775	0.892	1.000	0.880
		[27]	[34]	[34]	[35]	.	.	.	.
Parathyroid		(49)	(48)	(49)	(48)	.	.	.	.
	B-Adenoma	0	0	1	0	0.531	.	0.518	.
		[27]	[34]	[34]	[35]	.	.	.	.
Pituitary		(50)	(50)	(48)	(50)	.	.	.	.
	B-Adenoma	16	14	15	21	0.302	0.792	0.698	0.249
		[30]	[36]	[36]	[40]	.	.	.	.
Prostate		(50)	(50)	(50)	(50)	.	.	.	.
	M-Sarcoma	1	0	0	0	1.000	1.000	1.000	1.000
		[27]	[34]	[34]	[35]	.	.	.	.
	M-Schwannoma, Malignant	0	0	1	0	0.531	.	0.518	.
		[27]	[34]	[34]	[35]	.	.	.	.

Table 3A (2) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Rats (Vehicle control2, low, medium and high dose goups)

Organ Name	Tumor Name	Vehicle Cont2 N=50	100 mg Low N=50	300 mg Med N=50	700 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Seminal Vesicle		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma	0	0	1	0	0.531	.	0.518	.
		[27]	[34]	[34]	[35]	.	.	.	.
Skin/Subcutis		(50)	(50)	(50)	(50)	.	.	.	.
	B-Basal Cell Tumor	0	0	0	1	0.269	.	.	0.506
		[27]	[34]	[34]	[35]	.	.	.	.
	B-Fibroma	2	0	0	0	1.000	1.000	1.000	1.000
		[27]	[34]	[34]	[35]	.	.	.	.
	B-Keratoacanthoma	2	0	0	1	0.683	1.000	1.000	0.884
		[27]	[34]	[34]	[35]	.	.	.	.
	B-Papilloma, Squamous Cell	1	0	0	0	1.000	1.000	1.000	1.000
		[27]	[34]	[34]	[35]	.	.	.	.
	B-Trichoepithelioma	1	0	0	0	1.000	1.000	1.000	1.000
		[27]	[34]	[34]	[35]	.	.	.	.
	M-Carcinoma, Basal Cell	1	0	0	1	0.468	1.000	1.000	0.759
		[27]	[34]	[34]	[35]	.	.	.	.
	M-Fibrosarcoma	1	0	2	0	0.726	1.000	0.527	1.000
		[27]	[34]	[34]	[35]	.	.	.	.
Skin/Subcutis	SQUAMOUS_CELL_CARCINOMA+PAPIL	3	0	0	1	0.829	1.000	1.000	0.945
		[27]	[34]	[34]	[35]	.	.	.	.
Spinal Cord		(50)	(50)	(50)	(50)	.	.	.	.
	B-Oligodendroglioma	0	0	0	1	0.269	.	.	0.506
		[27]	[34]	[34]	[35]	.	.	.	.
Stomach, Gl		(50)	(50)	(50)	(50)	.	.	.	.
	M-Schwannoma, Malignant	0	0	0	1	0.269	.	.	0.506
		[27]	[34]	[34]	[35]	.	.	.	.
Stomach, Nongl		(50)	(49)	(50)	(50)	.	.	.	.
	M-Sarcoma, NOS	1	0	0	0	1.000	1.000	1.000	1.000
		[27]	[34]	[34]	[35]	.	.	.	.
Testis		(50)	(50)	(50)	(50)	.	.	.	.
	B-Interstitial Cell Tumor	0	4	7	6	0.057	0.070	0.008	0.014
		[27]	[34]	[34]	[35]	.	.	.	.
Thymus		(50)	(47)	(50)	(50)	.	.	.	.
	B-Thymoma	0	0	1	0	0.531	.	0.518	.
		[27]	[34]	[34]	[35]	.	.	.	.

Table 3A (2) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Rats (Vehicle control2, low, medium and high dose goups)

Organ Name	Tumor Name	Vehicle Cont2 N=50	100 mg Low N=50	300 mg Med N=50	700 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Thyroid		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, C-cell	4	8	6	6	0.589	0.225	0.431	0.385
		[27]	[34]	[35]	[35]	.	.	.	.
	B-Adenoma, Follicular Cell	2	7	13	4	0.582	0.102	0.003	0.350
		[27]	[34]	[34]	[35]	.	.	.	.
	C_CELL_ADENOMA+CARCINOMA	4	8	7	6	0.586	0.225	0.317	0.385
		[27]	[34]	[35]	[35]	.	.	.	.
	FOLLICULAR_CELL_ADENOMA+CARCI	3	8	13	6	0.478	0.121	0.007	0.241
		[27]	[34]	[34]	[35]	.	.	.	.
	M-Carcinoma, C-cell	0	0	1	0	0.531	.	0.518	.
		[27]	[34]	[34]	[35]	.	.	.	.
	M-Carcinoma, Follicular Cell	1	1	0	2	0.293	0.770	1.000	0.500
		[27]	[34]	[34]	[35]	.	.	.	.

Numbers with parentheses are number of the animals with organs examined and also usable

Numbers with brackets are survival-adjusted group size

Numbers are the tumor bearing animals

Table 3A (3): Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Rats (Combined vehicle controls, low, medium and high dose groups)

Organ Name	Tumor Name	Combined Conts N=100	100 mg Low N=50	300 mg Med N=50	700 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Adipose Tissue		(100)	(50)	(49)	(50)	.	.	.	.
	M-Schwannoma, Malignant	0	0	0	1	0.211	.	.	0.333
		[63]	[34]	[34]	[35]	.	.	.	.
Adrenal, Cortex		(100)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma	5	0	0	0	1.000	1.000	1.000	1.000
		[63]	[34]	[34]	[35]	.	.	.	.
	M-Carcinoma	0	1	0	0	0.621	0.349	.	.
		[63]	[34]	[34]	[35]	.	.	.	.
Adrenal, Medull		(100)	(50)	(50)	(50)	.	.	.	.
	B-Pheochromocytoma	3	3	1	3	0.319	0.343	0.816	0.312
		[63]	[35]	[34]	[35]	.	.	.	.
	M-Complex Malignant Pheochro	1	0	0	0	1.000	1.000	1.000	1.000
		[63]	[34]	[34]	[35]	.	.	.	.
Body, Whole/Cav		(100)	(50)	(50)	(50)	.	.	.	.
	B-Hemangioma	3	2	5	9	0.001	0.567	0.089	0.002
		[63]	[34]	[34]	[35]	.	.	.	.
	B-Hibernoma	1	0	0	0	1.000	1.000	1.000	1.000
		[63]	[34]	[34]	[35]	.	.	.	.
	M-Hibernoma	0	2	1	0	0.617	0.120	0.349	.
		[63]	[35]	[35]	[35]	.	.	.	.
	M-Histiocytic Sarcoma	1	0	0	3	0.030	1.000	1.000	0.105
		[63]	[34]	[34]	[35]	.	.	.	.
	M-Lymphosarcoma	0	0	5	0	0.294	.	0.005	.
		[63]	[34]	[36]	[35]	.	.	.	.
Brain		(100)	(50)	(50)	(50)	.	.	.	.
	B-Granular Cell Tumor	2	1	0	0	0.947	0.724	1.000	1.000
		[63]	[34]	[34]	[35]	.	.	.	.
	B-Meningioma	0	1	0	0	0.623	0.349	.	.
		[63]	[35]	[34]	[35]	.	.	.	.
Duodenum		(98)	(50)	(50)	(50)	.	.	.	.
	M-Carcinoma	1	0	0	0	1.000	1.000	1.000	1.000
		[63]	[34]	[34]	[35]	.	.	.	.
	M-Leiomyosarcoma	0	1	0	0	0.621	0.349	.	.
		[63]	[34]	[34]	[35]	.	.	.	.
Epididymis		(100)	(50)	(50)	(50)	.	.	.	.
	B-Benign Mesothelioma	0	1	0	0	0.621	0.349	.	.
		[63]	[34]	[34]	[35]	.	.	.	.
	M-Schwannoma, Malignant	1	0	0	0	1.000	1.000	1.000	1.000
		[63]	[34]	[34]	[35]	.	.	.	.

Table 3A (3) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Rats (Combined vehicle controls, low, medium and high dose groups)

Organ Name	Tumor Name	Combined Conts N=100	100 mg Low N=50	300 mg Med N=50	700 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Gl, Preputial		(100)	(50)	(50)	(50)	.	.	.	.
	M-Schwannoma, Malignant	0	1	0	0	0.621	0.349	.	.
		[63]	[34]	[34]	[35]	.	.	.	.
Gl, Zymbal's		(100)	(50)	(50)	(50)	.	.	.	.
	M-Carcinoma	1	0	1	0	0.664	1.000	0.574	1.000
		[63]	[34]	[35]	[35]	.	.	.	.
Heart		(99)	(50)	(50)	(50)	.	.	.	.
	M-Endocardial Schwannoma	2	0	0	0	1.000	1.000	1.000	1.000
		[63]	[34]	[34]	[35]	.	.	.	.
Jejunum		(98)	(50)	(50)	(49)	.	.	.	.
	M-Carcinoma	1	0	0	0	1.000	1.000	1.000	1.000
		[63]	[34]	[34]	[35]	.	.	.	.
Kidney		(100)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Tubule Cell	0	0	0	1	0.211	.	.	0.333
		[63]	[34]	[34]	[35]	.	.	.	.
	B-Lipoma	1	0	0	0	1.000	1.000	1.000	1.000
		[63]	[34]	[34]	[35]	.	.	.	.
	M-Carcinoma, Tubule Cell	1	0	0	0	1.000	1.000	1.000	1.000
		[63]	[34]	[34]	[35]	.	.	.	.
Liver		(100)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Hepatocellular	0	0	1	0	0.416	.	0.344	.
		[63]	[34]	[34]	[35]	.	.	.	.
Mammary, Male		(99)	(50)	(50)	(49)	.	.	.	.
	B-Fibroadenoma	1	0	0	0	1.000	1.000	1.000	1.000
		[63]	[34]	[34]	[35]	.	.	.	.
Mesentery		(100)	(50)	(50)	(50)	.	.	.	.
	M-Schwannoma, Malignant	0	0	0	1	0.211	.	.	0.333
		[63]	[34]	[34]	[35]	.	.	.	.
Muscle, Bi Fem		(100)	(50)	(50)	(50)	.	.	.	.
	B-Granular Cell Tumor	1	0	0	0	1.000	1.000	1.000	1.000
		[63]	[34]	[34]	[35]	.	.	.	.
Pancreas		(100)	(50)	(50)	(49)	.	.	.	.
	B-Adenoma, Islet Cell	2	1	0	1	0.601	0.724	1.000	0.704
		[63]	[34]	[34]	[35]	.	.	.	.
Parathyroid		(95)	(48)	(49)	(48)	.	.	.	.
	B-Adenoma	0	0	1	0	0.416	.	0.344	.
		[63]	[34]	[34]	[35]	.	.	.	.

Table 3A (3) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Rats (Combined vehicle controls, low, medium and high dose groups)

Organ Name	Tumor Name	Combined Conts N=100	100 mg Low N=50	300 mg Med N=50	700 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Pituitary		(100)	(50)	(48)	(50)	.	.	.	.
	B-Adenoma	43	14	15	21	0.706	0.981	0.958	0.638
		[71]	[36]	[36]	[40]	.	.	.	.
Prostate		(99)	(50)	(50)	(50)	.	.	.	.
	M-Sarcoma	1	0	0	0	1.000	1.000	1.000	1.000
		[63]	[34]	[34]	[35]	.	.	.	.
	M-Schwannoma, Malignant	0	0	1	0	0.416	.	0.344	.
		[63]	[34]	[34]	[35]	.	.	.	.
Seminal Vesicle		(99)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma	0	0	1	0	0.416	.	0.344	.
		[63]	[34]	[34]	[35]	.	.	.	.
Skin/Subcutis		(100)	(50)	(50)	(50)	.	.	.	.
	B-Basal Cell Tumor	0	0	0	1	0.211	.	.	0.333
		[63]	[34]	[34]	[35]	.	.	.	.
	B-Fibroma	3	0	0	0	1.000	1.000	1.000	1.000
		[63]	[34]	[34]	[35]	.	.	.	.
	B-Keratoacanthoma	2	0	0	1	0.544	1.000	1.000	0.707
		[63]	[34]	[34]	[35]	.	.	.	.
	B-Papilloma, Squamous Cell	2	0	0	0	1.000	1.000	1.000	1.000
		[63]	[34]	[34]	[35]	.	.	.	.
	B-Trichoepithelioma	1	0	0	0	1.000	1.000	1.000	1.000
		[63]	[34]	[34]	[35]	.	.	.	.
	M-Carcinoma, Basal Cell	2	0	0	1	0.544	1.000	1.000	0.707
		[63]	[34]	[34]	[35]	.	.	.	.
	M-Carcinoma, Squamous Cell	1	0	0	0	1.000	1.000	1.000	1.000
		[63]	[34]	[34]	[35]	.	.	.	.
	M-Fibrosarcoma	1	0	2	0	0.592	1.000	0.272	1.000
		[63]	[34]	[34]	[35]	.	.	.	.
	M-Sarcoma	1	0	0	0	1.000	1.000	1.000	1.000
		[63]	[34]	[34]	[35]	.	.	.	.
	SQUAMOUS_CELL_CARCINOMA+PAPIL	5	0	0	1	0.872	1.000	1.000	0.915
		[63]	[34]	[34]	[35]	.	.	.	.
Spinal Cord		(100)	(50)	(50)	(50)	.	.	.	.
	B-Oligodendroglioma	0	0	0	1	0.211	.	.	0.333
		[63]	[34]	[34]	[35]	.	.	.	.
Stomach, Gl		(100)	(50)	(50)	(50)	.	.	.	.
	M-Schwannoma, Malignant	0	0	0	1	0.211	.	.	0.333
		[63]	[34]	[34]	[35]	.	.	.	.

Table 3A (3) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Rats (Combined vehicle controls, low, medium and high dose groups)

Organ Name	Tumor Name	Combined Conts N=100	100 mg Low N=50	300 mg Med N=50	700 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Stomach, Nongl		(100)	(49)	(50)	(50)	.	.	.	.
	M-Sarcoma, NOS	1	0	0	0	1.000	1.000	1.000	1.000
		[63]	[34]	[34]	[35]	.	.	.	.
Testis		(100)	(50)	(50)	(50)	.	.	.	.
	B-Interstitial Cell Tumor	2	4	7	6	0.022	0.111	0.008	0.017
		[63]	[34]	[34]	[35]	.	.	.	.
Thymus		(99)	(47)	(50)	(50)	.	.	.	.
	B-Thymoma	4	0	1	0	0.950	1.000	0.881	1.000
		[63]	[34]	[34]	[35]	.	.	.	.
Thyroid		(100)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, C-cell	11	8	6	6	0.604	0.310	0.575	0.517
		[63]	[34]	[35]	[35]	.	.	.	.
	B-Adenoma, Follicular Cell	11	7	13	4	0.681	0.435	0.021	0.801
		[63]	[34]	[34]	[35]	.	.	.	.
	C_CELL_ADENOMA+CARCINOMA	12	8	7	6	0.646	0.378	0.508	0.588
		[63]	[34]	[35]	[35]	.	.	.	.
	FOLLICULAR_CELL_ADENOMA+CARCI	13	8	13	6	0.583	0.447	0.048	0.654
		[63]	[34]	[34]	[35]	.	.	.	.
	M-Carcinoma, C-cell	1	0	1	0	0.660	1.000	0.571	1.000
		[63]	[34]	[34]	[35]	.	.	.	.
	M-Carcinoma, Follicular Cell	3	1	0	2	0.433	0.822	1.000	0.537
		[63]	[34]	[34]	[35]	.	.	.	.

Numbers with parentheses are number of the animals with organs examined and also usable

Numbers with brackets are survival-adjusted group size

Numbers are the tumor bearing animals

Table 3B (1): Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Female Rats (Vehicle control1, low, medium and high dose groups)

Organ Name	Tumor Name	Vehicle Cont1 N=50	100 mg Low N=50	300 mg Med N=50	700 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Adrenal, Cortex		(50)	(50)	(50)	(49)	.	.	.	.
	B-Adenoma	0	1	1	0	0.613	0.512	0.500	.
		[34]	[39]	[35]	[33]	.	.	.	.
Adrenal, Medull		(50)	(50)	(50)	(49)	.	.	.	.
	B-Pheochromocytoma	0	0	1	0	0.486	.	0.500	.
		[34]	[38]	[35]	[33]	.	.	.	.
Body, Whole/Cav		(50)	(50)	(50)	(50)	.	.	.	.
	B-Hibernoma	0	0	0	1	0.241	.	.	0.500
		[34]	[38]	[35]	[34]	.	.	.	.
	M-Hemangiosarcoma	1	0	0	0	1.000	1.000	1.000	1.000
		[34]	[38]	[35]	[33]	.	.	.	.
	M-Hibernoma	1	1	0	0	0.941	0.759	1.000	1.000
		[35]	[39]	[35]	[33]	.	.	.	.
	M-Histiocytic Sarcoma	2	0	1	1	0.567	1.000	0.880	0.870
		[35]	[38]	[35]	[33]	.	.	.	.
Cervix		(50)	(50)	(50)	(50)	.	.	.	.
	B-Granular Cell Tumor	0	1	0	0	0.757	0.512	.	.
		[34]	[38]	[35]	[33]	.	.	.	.
	B-Polyp, Endometrial Stromal	0	3	0	1	0.569	0.134	.	0.494
		[34]	[39]	[35]	[33]	.	.	.	.
	M-Carcinoma, Squamous Cell	0	0	1	0	0.486	.	0.500	.
		[34]	[38]	[35]	[33]	.	.	.	.
	M-Schwannoma, Malignant	1	2	0	0	0.936	0.518	1.000	1.000
		[35]	[39]	[35]	[33]	.	.	.	.
Eye		(50)	(50)	(50)	(50)	.	.	.	.
	M-Melanoma, Amelanotic	1	0	0	0	1.000	1.000	1.000	1.000
		[34]	[38]	[35]	[33]	.	.	.	.
Gl, Clitoral		(50)	(50)	(50)	(50)	.	.	.	.
	B-Papilloma, Squamous	0	0	0	1	0.236	.	.	0.494
		[34]	[38]	[35]	[33]	.	.	.	.
Gl, Harderian		(50)	(50)	(50)	(50)	.	.	.	.
	M-Schwannoma, Malignant	0	0	0	1	0.236	.	.	0.494
		[34]	[38]	[35]	[33]	.	.	.	.
Kidney		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Tubule Cell	0	0	1	0	0.486	.	0.500	.
		[34]	[38]	[35]	[33]	.	.	.	.
	B-Lipoma	1	0	0	0	1.000	1.000	1.000	1.000
		[34]	[38]	[35]	[33]	.	.	.	.
	B-Papilloma, Transitional C	0	0	1	0	0.489	.	0.506	.

Table 3B (1) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Female Rats (Vehicle control1, low, medium and high dose groups)

Organ Name	Tumor Name	Vehicle Cont1 N=50	100 mg Low N=50	300 mg Med N=50	700 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
		[34]	[38]	[36]	[33]	.	.	.	.
	M-Carcinoma, Transitional Ce	0	1	0	0	0.757	0.512	.	.
		[34]	[38]	[35]	[33]	.	.	.	.
	M-Carcinoma, Tubule Cell	0	0	0	1	0.241	.	.	0.500
		[34]	[38]	[35]	[34]	.	.	.	.
Kidney	M-Sarcoma	0	0	0	1	0.241	.	.	0.494
		[34]	[38]	[35]	[34]	.	.	.	.
Mammary, Female		(50)	(50)	(49)	(50)	.	.	.	.
	B-Adenoma	2	2	1	0	0.941	0.701	0.875	1.000
		[35]	[39]	[35]	[33]	.	.	.	.
	B-Cystadenoma	0	1	0	0	0.759	0.512	.	.
		[34]	[39]	[35]	[33]	.	.	.	.
	B-Fibroadenoma	6	9	3	2	0.979	0.320	0.922	0.966
		[35]	[39]	[35]	[33]	.	.	.	.
	M-Carcinoma	4	1	1	1	0.890	0.977	0.973	0.973
		[34]	[39]	[35]	[34]	.	.	.	.
Mesentery		(50)	(50)	(50)	(50)	.	.	.	.
	M-Leiomyosarcoma	0	0	1	0	0.486	.	0.500	.
		[34]	[38]	[35]	[33]	.	.	.	.
	M-Schwannoma, Malignant	0	0	0	1	0.241	.	.	0.500
		[34]	[38]	[35]	[34]	.	.	.	.
Ovary		(50)	(50)	(50)	(50)	.	.	.	.
	B-Granulosa/Theca Cell Tumor	0	1	0	1	0.301	0.518	.	0.494
		[34]	[39]	[35]	[33]	.	.	.	.
	B-Papillary Cystadenoma	1	0	0	0	1.000	1.000	1.000	1.000
		[34]	[38]	[35]	[33]	.	.	.	.
Pancreas		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Islet Cell	1	0	0	0	1.000	1.000	1.000	1.000
		[34]	[38]	[35]	[33]	.	.	.	.
	M-Carcinoma, Islet Cell	0	1	0	0	0.757	0.512	.	.
		[34]	[38]	[35]	[33]	.	.	.	.
Parathyroid		(41)	(46)	(47)	(44)	.	.	.	.
	B-Adenoma	0	1	0	0	0.759	0.512	.	.
		[34]	[39]	[35]	[33]	.	.	.	.
Pituitary		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma	29	33	34	27	0.560	0.210	0.150	0.598
		[40]	[42]	[39]	[37]	.	.	.	.

Table 3B (1) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Female Rats (Vehicle control1, low, medium and high dose groups)

Organ Name	Tumor Name	Vehicle Cont1 N=50	100 mg Low N=50	300 mg Med N=50	700 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Skin/Subcutis		(50)	(50)	(50)	(50)	.	.	.	.
	B-Keratoacanthoma	0	0	1	0	0.489	.	0.506	.
		[34]	[38]	[36]	[33]	.	.	.	.
	B-Papilloma, Squamous Cell	0	0	1	1	0.173	.	0.500	0.494
		[34]	[38]	[35]	[33]	.	.	.	.
	M-Fibrosarcoma	0	1	0	0	0.759	0.518	.	.
		[34]	[39]	[35]	[33]	.	.	.	.
	SQUAMOUS_CELL_PAPILLOMA+KERAT	0	0	2	1	0.182	.	0.253	0.494
		[34]	[38]	[36]	[33]	.	.	.	.
Stomach, Gl		(49)	(50)	(50)	(50)	.	.	.	.
Stomach, Nongl		(48)	(50)	(50)	(50)	.	.	.	.
	M-Carcinoma, Squamous Cell	0	1	0	0	0.759	0.512	.	.
		[34]	[39]	[35]	[33]	.	.	.	.
Thymus		(49)	(49)	(48)	(49)	.	.	.	.
	B-Thymoma	4	3	2	2	0.797	0.793	0.894	0.888
Thymus	B-Thymoma	[34]	[38]	[35]	[34]	.	.	.	.
	M-Carcinoma	0	1	0	0	0.759	0.518	.	.
		[34]	[39]	[35]	[33]	.	.	.	.
	M-Schwannoma, Malignant	0	0	0	1	0.236	.	.	0.494
		[34]	[38]	[35]	[33]	.	.	.	.
Thyroid		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, C-cell	6	3	6	5	0.411	0.933	0.639	0.726
		[34]	[39]	[36]	[33]	.	.	.	.
	B-Adenoma, Follicular Cell	4	3	3	1	0.902	0.803	0.784	0.971
		[34]	[39]	[35]	[33]	.	.	.	.
	FOLLICULAR_CELL_ADENOMA+CARCI	5	5	3	1	0.967	0.661	0.860	0.985
		[35]	[39]	[35]	[33]	.	.	.	.
	M-Carcinoma, Follicular Cell	1	2	0	0	0.936	0.509	1.000	1.000
		[35]	[39]	[35]	[33]	.	.	.	.
Uterus		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma	0	0	0	1	0.236	.	.	0.494
		[34]	[38]	[35]	[33]	.	.	.	.
	B-Polyp, Endometrial Stromal	5	5	5	5	0.476	0.661	0.631	0.631
		[34]	[38]	[35]	[34]	.	.	.	.
	M-Carcinoma	1	1	1	1	0.493	0.765	0.753	0.747
		[34]	[39]	[35]	[33]	.	.	.	.
	M-Schwannoma, Malignant	1	0	0	0	1.000	1.000	1.000	1.000
		[34]	[38]	[35]	[33]	.	.	.	.

Table 3B (1) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Female Rats (Vehicle control1, low, medium and high dose groups)

Organ Name	Tumor Name	Vehicle Cont1 N=50	100 mg Low N=50	300 mg Med N=50	700 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Vagina		(50)	(50)	(50)	(50)	.	.	.	.
	B-Granular Cell Tumor	0	0	1	1	0.173	.	0.500	0.494
		[34]	[38]	[35]	[33]	.	.	.	.
	B-Stromal Polyp	0	1	0	0	0.757	0.512	.	.
		[34]	[38]	[35]	[33]	.	.	.	.
	M-Schwannoma, Malignant	1	1	0	0	0.941	0.759	1.000	1.000
		[35]	[39]	[35]	[33]	.	.	.	.

Numbers with parentheses are number of the animals with organs examined and also usable

Numbers with brackets are survival-adjusted group size

Numbers are the tumor bearing animals

Table 3B (2): Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Female Rats (Vehicle control2, low, medium and high dose groups)

		Vehicle	100 mg	300 mg	700 mg				
		Cont2	Low	Med	High	P_Value	P_Value	P_Value	P_Value
Organ Name	Tumor Name	N=50	N=50	N=50	N=50	Dos Resp	C vs. L	C vs. M	C vs. H
Adrenal, Cortex		(50)	(50)	(50)	(49)	.	.	.	.
	B-Adenoma	0	1	1	0	0.619	0.531	0.519	.
		[33]	[39]	[35]	[33]	.	.	.	.
	M-Carcinoma	1	0	0	0	1.000	1.000	1.000	1.000
		[33]	[38]	[35]	[33]	.	.	.	.
Adrenal, Medull		(50)	(50)	(50)	(49)	.	.	.	.
	B-Pheochromocytoma	1	0	1	0	0.741	1.000	0.772	1.000
		[33]	[38]	[35]	[33]	.	.	.	.
Body, Whole/Cav		(50)	(50)	(50)	(50)	.	.	.	.
	B-Hemangioma	1	0	0	0	1.000	1.000	1.000	1.000
		[33]	[38]	[35]	[33]	.	.	.	.
	B-Hibernoma	0	0	0	1	0.243	.	.	0.519
		[33]	[38]	[35]	[34]	.	.	.	.
	M-Hemangiosarcoma	1	0	0	0	1.000	1.000	1.000	1.000
		[33]	[38]	[35]	[33]	.	.	.	.
	M-Hibernoma	0	1	0	0	0.764	0.531	.	.
		[33]	[39]	[35]	[33]	.	.	.	.
	M-Histiocytic Sarcoma	1	0	1	1	0.372	1.000	0.778	0.766
		[33]	[38]	[35]	[33]	.	.	.	.
	M-Lymphosarcoma	1	0	0	0	1.000	1.000	1.000	1.000
		[33]	[38]	[35]	[33]	.	.	.	.
Brain		(50)	(50)	(50)	(50)	.	.	.	.
	B-Granular Cell Tumor	4	0	0	0	1.000	1.000	1.000	1.000
		[34]	[38]	[35]	[33]	.	.	.	.
Cervix		(49)	(50)	(50)	(50)	.	.	.	.
	B-Fibroma	1	0	0	0	1.000	1.000	1.000	1.000
		[34]	[38]	[35]	[33]	.	.	.	.
	B-Granular Cell Tumor	0	1	0	0	0.763	0.531	.	.
		[33]	[38]	[35]	[33]	.	.	.	.
	B-Polyp, Endometrial Stromal	0	3	0	1	0.575	0.150	.	0.513
		[33]	[39]	[35]	[33]	.	.	.	.
	M-Carcinoma, Squamous Cell	0	0	1	0	0.489	.	0.519	.
		[33]	[38]	[35]	[33]	.	.	.	.
	M-Leiomyosarcoma	1	0	0	0	1.000	1.000	1.000	1.000
		[34]	[38]	[35]	[33]	.	.	.	.
	M-Schwannoma, Malignant	1	2	0	0	0.942	0.556	1.000	1.000
		[33]	[39]	[35]	[33]	.	.	.	.

Table 3B (2) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Female Rats (Vehicle control2, low, medium and high dose groups)

Organ Name	Tumor Name	Vehicle Cont2 N=50	100 mg Low N=50	300 mg Med N=50	700 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Gl, Clitoral		(50)	(50)	(50)	(50)	.	.	.	.
	B-Papilloma, Squamous	0	0	0	1	0.237	.	.	0.513
		[33]	[38]	[35]	[33]	.	.	.	.
Gl, Harderian		(50)	(50)	(50)	(50)	.	.	.	.
	M-Schwannoma, Malignant	0	0	0	1	0.237	.	.	0.513
		[33]	[38]	[35]	[33]	.	.	.	.
Kidney		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Tubule Cell	0	0	1	0	0.489	.	0.519	.
		[33]	[38]	[35]	[33]	.	.	.	.
	B-Papilloma, Transitional C	0	0	1	0	0.493	.	0.525	.
		[33]	[38]	[36]	[33]	.	.	.	.
	M-Carcinoma, Transitional Ce	0	1	0	0	0.763	0.531	.	.
		[33]	[38]	[35]	[33]	.	.	.	.
	M-Carcinoma, Tubule Cell	0	0	0	1	0.243	.	.	0.519
		[33]	[38]	[35]	[34]	.	.	.	.
	M-Sarcoma	0	0	0	1	0.243	.	.	0.513
		[33]	[38]	[35]	[34]	.	.	.	.
Mammary, Female		(50)	(50)	(49)	(50)	.	.	.	.
	B-Adenoma	1	2	1	0	0.873	0.547	0.772	1.000
		[33]	[39]	[35]	[33]	.	.	.	.
	B-Cystadenoma	0	1	0	0	0.764	0.531	.	.
		[33]	[39]	[35]	[33]	.	.	.	.
	B-Fibroadenoma	9	9	3	2	0.997	0.733	0.992	0.997
		[34]	[39]	[35]	[33]	.	.	.	.
	B-Fibroma	1	0	0	0	1.000	1.000	1.000	1.000
		[34]	[38]	[35]	[33]	.	.	.	.
	M-Carcinoma	4	1	1	1	0.886	0.980	0.976	0.976
		[35]	[39]	[35]	[34]	.	.	.	.
Mesentery		(50)	(50)	(50)	(50)	.	.	.	.
	M-Leiomyosarcoma	0	0	1	0	0.489	.	0.519	.
		[33]	[38]	[35]	[33]	.	.	.	.
	M-Schwannoma, Malignant	0	0	0	1	0.243	.	.	0.519
		[33]	[38]	[35]	[34]	.	.	.	.
Ovary		(50)	(50)	(50)	(50)	.	.	.	.
	B-Granulosa/Theca Cell Tumor	1	1	0	1	0.527	0.788	1.000	0.766
		[34]	[39]	[35]	[33]	.	.	.	.
	M-Malignant Granulosa/Theca	2	0	0	0	1.000	1.000	1.000	1.000
		[33]	[38]	[35]	[33]	.	.	.	.

Table 3B (2) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Female Rats (Vehicle control2, low, medium and high dose groups)

		Vehicle	100 mg	300 mg	700 mg				
		Cont2	Low	Med	High	P_Value	P_Value	P_Value	P_Value
Organ Name	Tumor Name	N=50	N=50	N=50	N=50	Dos Resp	C vs. L	C vs. M	C vs. H
Pancreas		(50)	(50)	(50)	(50)	.	.	.	.
	M-Carcinoma, Islet Cell	0	1	0	0	0.763	0.531	.	.
		[33]	[38]	[35]	[33]	.	.	.	.
Parathyroid		(47)	(46)	(47)	(44)	.	.	.	.
	B-Adenoma	0	1	0	0	0.764	0.531	.	.
		[33]	[39]	[35]	[33]	.	.	.	.
Pituitary		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma	32	33	34	27	0.770	0.633	0.544	0.913
		[40]	[42]	[39]	[37]	.	.	.	.
Rectum		(50)	(50)	(50)	(50)	.	.	.	.
	B-Granular Cell Tumor	1	0	0	0	1.000	1.000	1.000	1.000
		[33]	[38]	[35]	[33]	.	.	.	.
Skin/Subcutis		(50)	(50)	(50)	(50)	.	.	.	.
	B-Keratoacanthoma	0	0	1	0	0.493	.	0.525	.
		[33]	[38]	[36]	[33]	.	.	.	.
Skin/Subcutis	B-Papilloma, Squamous Cell	1	0	1	1	0.372	1.000	0.772	0.766
		[33]	[38]	[35]	[33]	.	.	.	.
	M-Carcinoma, Basal Cell	1	0	0	0	1.000	1.000	1.000	1.000
		[34]	[38]	[35]	[33]	.	.	.	.
	M-Fibrosarcoma	1	1	0	0	0.943	0.788	1.000	1.000
		[34]	[39]	[35]	[33]	.	.	.	.
	SQUAMOUS_CELL_CARCINOMA+PAPIL	1	0	2	1	0.348	1.000	0.538	0.766
		[33]	[38]	[36]	[33]	.	.	.	.
Stomach, Nongl		(50)	(50)	(50)	(50)	.	.	.	.
	M-Carcinoma, Squamous Cell	0	1	0	0	0.764	0.531	.	.
		[33]	[39]	[35]	[33]	.	.	.	.
Thymus		(50)	(49)	(48)	(49)	.	.	.	.
	B-Thymoma	2	3	2	2	0.586	0.559	0.721	0.712
		[33]	[38]	[35]	[34]	.	.	.	.
	M-Carcinoma	0	1	0	0	0.764	0.537	.	.
		[33]	[39]	[35]	[33]	.	.	.	.
	M-Schwannoma, Malignant	0	0	0	1	0.237	.	.	0.513
		[33]	[38]	[35]	[33]	.	.	.	.
Thyroid		(49)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, C-cell	6	3	6	5	0.411	0.948	0.692	0.771
		[34]	[39]	[36]	[33]	.	.	.	.
	B-Adenoma, Follicular Cell	2	3	3	1	0.759	0.559	0.537	0.889
		[33]	[39]	[35]	[33]	.	.	.	.

Table 3B (2) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Female Rats (Vehicle control2, low, medium and high dose groups)

Organ Name	Tumor Name	Vehicle Cont2 N=50	100 mg Low N=50	300 mg Med N=50	700 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
	FOLLICULAR_CELL_ADENOMA+CARCI	2	5	3	1	0.855	0.282	0.537	0.889
		[33]	[39]	[35]	[33]	.	.	.	.
	M-Carcinoma, Follicular Cell	0	2	0	0	0.814	0.279	.	.
		[33]	[39]	[35]	[33]	.	.	.	.
Uterus		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma	0	0	0	1	0.237	.	.	0.513
		[33]	[38]	[35]	[33]	.	.	.	.
	B-Polyp, Endometrial Stromal	5	5	5	5	0.476	0.708	0.680	0.680
		[34]	[38]	[35]	[34]	.	.	.	.
	M-Carcinoma	0	1	1	1	0.283	0.531	0.519	0.513
		[33]	[39]	[35]	[33]	.	.	.	.
Vagina		(50)	(50)	(50)	(50)	.	.	.	.
	B-Granular Cell Tumor	0	0	1	1	0.176	.	0.519	0.513
		[33]	[38]	[35]	[33]	.	.	.	.
	B-Stromal Polyp	0	1	0	0	0.763	0.531	.	.
		[33]	[38]	[35]	[33]	.	.	.	.
	M-Schwannoma, Malignant	0	1	0	0	0.764	0.531	.	.
		[33]	[39]	[35]	[33]	.	.	.	.

Numbers with parentheses are number of the animals with organs examined and also usable

Numbers with brackets are survival-adjusted group size

Numbers are the tumor bearing animals

Table 3B (3): Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Female Rats (Combined vehicle controls, low, medium and high dose groups)

Organ Name	Tumor Name	Combined Conts N=100	100 mg Low N=50	300 mg Med N=50	700 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Adrenal, Cortex		(100)	(50)	(50)	(49)	.	.	.	.
	B-Adenoma	0	1	1	0	0.475	0.353	0.342	.
		[67]	[39]	[35]	[33]	.	.	.	.
	M-Carcinoma	1	0	0	0	1.000	1.000	1.000	1.000
		[67]	[38]	[35]	[33]	.	.	.	.
Adrenal, Medull		(100)	(50)	(50)	(49)	.	.	.	.
	B-Pheochromocytoma	1	0	1	0	0.633	1.000	0.569	1.000
		[67]	[38]	[35]	[33]	.	.	.	.
Bile Duct		(100)	(49)	(50)	(49)	.	.	.	.
Body, Whole/Cav		(100)	(50)	(50)	(50)	.	.	.	.
	B-Hemangioma	1	0	0	0	1.000	1.000	1.000	1.000
		[67]	[38]	[35]	[33]	.	.	.	.
	B-Hibernoma	0	0	0	1	0.195	.	.	0.342
		[67]	[38]	[35]	[34]	.	.	.	.
	M-Hemangiosarcoma	2	0	0	0	1.000	1.000	1.000	1.000
		[67]	[38]	[35]	[33]	.	.	.	.
	M-Hibernoma	1	1	0	0	0.850	0.583	1.000	1.000
		[68]	[39]	[35]	[33]	.	.	.	.
	M-Histiocytic Sarcoma	3	0	1	1	0.564	1.000	0.820	0.808
		[68]	[38]	[35]	[33]	.	.	.	.
	M-Lymphosarcoma	1	0	0	0	1.000	1.000	1.000	1.000
		[67]	[38]	[35]	[33]	.	.	.	.
Brain		(100)	(50)	(50)	(50)	.	.	.	.
	B-Granular Cell Tumor	4	0	0	0	1.000	1.000	1.000	1.000
		[68]	[38]	[35]	[33]	.	.	.	.
Cervix		(99)	(50)	(50)	(50)	.	.	.	.
	B-Fibroma	1	0	0	0	1.000	1.000	1.000	1.000
		[68]	[38]	[35]	[33]	.	.	.	.
	B-Granular Cell Tumor	0	1	0	0	0.613	0.353	.	.
		[67]	[38]	[35]	[33]	.	.	.	.
	B-Polyp, Endometrial Stromal	0	3	0	1	0.416	0.044	.	0.336
		[67]	[39]	[35]	[33]	.	.	.	.
	M-Carcinoma, Squamous Cell	0	0	1	0	0.393	.	0.342	.
		[67]	[38]	[35]	[33]	.	.	.	.
	M-Leiomyosarcoma	1	0	0	0	1.000	1.000	1.000	1.000
		[68]	[38]	[35]	[33]	.	.	.	.
	M-Schwannoma, Malignant	2	2	0	0	0.927	0.451	1.000	1.000
		[68]	[39]	[35]	[33]	.	.	.	.

Table 3B (3) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Female Rats (Combined vehicle controls, low, medium and high dose groups)

Organ Name	Tumor Name	Combined Conts N=100	100 mg Low N=50	300 mg Med N=50	700 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Eye		(100)	(50)	(50)	(50)	.	.	.	.
	M-Melanoma, Amelanotic	1	0	0	0	1.000	1.000	1.000	1.000
		[67]	[38]	[35]	[33]	.	.	.	.
Gl, Clitoral		(100)	(50)	(50)	(50)	.	.	.	.
	B-Papilloma, Squamous	0	0	0	1	0.191	.	.	0.336
		[67]	[38]	[35]	[33]	.	.	.	.
Gl, Harderian		(100)	(50)	(50)	(50)	.	.	.	.
	M-Schwannoma, Malignant	0	0	0	1	0.191	.	.	0.336
		[67]	[38]	[35]	[33]	.	.	.	.
Kidney		(100)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Tubule Cell	0	0	1	0	0.393	.	0.342	.
		[67]	[38]	[35]	[33]	.	.	.	.
	B-Lipoma	1	0	0	0	1.000	1.000	1.000	1.000
		[67]	[38]	[35]	[33]	.	.	.	.
	B-Papilloma, Transitional C	0	0	1	0	0.397	.	0.347	.
		[67]	[38]	[36]	[33]	.	.	.	.
	M-Carcinoma, Transitional Ce	0	1	0	0	0.613	0.353	.	.
		[67]	[38]	[35]	[33]	.	.	.	.
	M-Carcinoma, Tubule Cell	0	0	0	1	0.195	.	.	0.342
		[67]	[38]	[35]	[34]	.	.	.	.
	M-Sarcoma	0	0	0	1	0.195	.	.	0.336
		[67]	[38]	[35]	[34]	.	.	.	.
LN, Axillary		(96)	(50)	(48)	(50)	.	.	.	.
LN, Mandibular		(98)	(49)	(50)	(50)	.	.	.	.
LN, Mesenteric		(98)	(50)	(50)	(50)	.	.	.	.
Mammary, Female		(100)	(50)	(49)	(50)	.	.	.	.
	B-Adenoma	3	2	1	0	0.909	0.574	0.814	1.000
		[69]	[39]	[35]	[33]	.	.	.	.
	B-Cystadenoma	0	1	0	0	0.615	0.353	.	.
		[67]	[39]	[35]	[33]	.	.	.	.
	B-Fibroadenoma	15	9	3	2	0.994	0.497	0.981	0.994
		[69]	[39]	[35]	[33]	.	.	.	.
	B-Fibroma	1	0	0	0	1.000	1.000	1.000	1.000
		[68]	[38]	[35]	[33]	.	.	.	.
	M-Carcinoma	8	1	1	1	0.951	0.983	0.979	0.979
		[70]	[39]	[35]	[34]	.	.	.	.

Table 3B (3) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Female Rats (Combined vehicle controls, low, medium and high dose groups)

Organ Name	Tumor Name	Combined Conts N=100	100 mg Low N=50	300 mg Med N=50	700 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Mesentery		(100)	(50)	(50)	(50)	.	.	.	.
	M-Leiomyosarcoma	0	0	1	0	0.393	.	0.342	.
		[67]	[38]	[35]	[33]	.	.	.	.
	M-Schwannoma, Malignant	0	0	0	1	0.195	.	.	0.342
		[67]	[38]	[35]	[34]	.	.	.	.
Ovary		(100)	(50)	(50)	(50)	.	.	.	.
	B-Granulosa/Theca Cell Tumor	1	1	0	1	0.390	0.589	1.000	0.561
		[68]	[39]	[35]	[33]	.	.	.	.
	B-Papillary Cystadenoma	1	0	0	0	1.000	1.000	1.000	1.000
		[67]	[38]	[35]	[33]	.	.	.	.
	M-Malignant Granulosa/Theca	2	0	0	0	1.000	1.000	1.000	1.000
		[67]	[38]	[35]	[33]	.	.	.	.
Pancreas		(100)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Islet Cell	1	0	0	0	1.000	1.000	1.000	1.000
		[67]	[38]	[35]	[33]	.	.	.	.
	M-Carcinoma, Islet Cell	0	1	0	0	0.613	0.353	.	.
		[67]	[38]	[35]	[33]	.	.	.	.
Parathyroid		(88)	(46)	(47)	(44)	.	.	.	.
	B-Adenoma	0	1	0	0	0.615	0.353	.	.
		[67]	[39]	[35]	[33]	.	.	.	.
Pituitary		(100)	(50)	(50)	(50)	.	.	.	
	B-Adenoma	61	33	34	27	0.603	0.360	0.267	0.800
		[80]	[42]	[39]	[37]	.	.	.	.
Rectum		(100)	(50)	(50)	(50)	.	.	.	.
	B-Granular Cell Tumor	1	0	0	0	1.000	1.000	1.000	1.000
		[67]	[38]	[35]	[33]	.	.	.	.
Skin/Subcutis		(100)	(50)	(50)	(50)	.	.	.	.
	B-Keratoacanthoma	0	0	1	0	0.397	.	0.347	.
		[67]	[38]	[36]	[33]	.	.	.	.
	B-Papilloma, Squamous Cell	1	0	1	1	0.260	1.000	0.569	0.561
		[67]	[38]	[35]	[33]	.	.	.	.
	M-Carcinoma, Basal Cell	1	0	0	0	1.000	1.000	1.000	1.000
		[68]	[38]	[35]	[33]	.	.	.	.
	M-Fibrosarcoma	1	1	0	0	0.850	0.589	1.000	1.000
		[68]	[39]	[35]	[33]	.	.	.	.
	SQUAMOUS_CELL_CARCINOMA+PAPIL	1	0	2	1	0.229	1.000	0.276	0.561
		[67]	[38]	[36]	[33]	.	.	.	.

Table 3B (3) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Female Rats (Combined vehicle controls, low, medium and high dose groups)

Organ Name	Tumor Name	Combined Conts N=100	100 mg Low N=50	300 mg Med N=50	700 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Stomach, Nongl		(98)	(50)	(50)	(50)	.	.	.	.
	M-Carcinoma, Squamous Cell	0	1	0	0	0.615	0.353	.	.
		[67]	[39]	[35]	[33]	.	.	.	.
Thymus		(99)	(49)	(48)	(49)	.	.	.	.
	B-Thymoma	6	3	2	2	0.727	0.677	0.828	0.819
		[68]	[38]	[35]	[34]	.	.	.	.
	M-Carcinoma	0	1	0	0	0.615	0.358	.	.
		[67]	[39]	[35]	[33]	.	.	.	.
	M-Schwannoma, Malignant	0	0	0	1	0.191	.	.	0.336
		[67]	[38]	[35]	[33]	.	.	.	.
Thyroid		(99)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, C-cell	12	3	6	5	0.501	0.952	0.649	0.745
		[68]	[39]	[36]	[33]	.	.	.	.
	B-Adenoma, Follicular Cell	6	3	3	1	0.850	0.677	0.651	0.948
		[68]	[39]	[35]	[33]	.	.	.	.
	FOLLICULAR_CELL_ADENOMA+CARCI	7	5	3	1	0.924	0.430	0.724	0.966
		[68]	[39]	[35]	[33]	.	.	.	.
	M-Carcinoma, Follicular Cell	1	2	0	0	0.842	0.283	1.000	1.000
		[68]	[39]	[35]	[33]	.	.	.	.
Uterus		(100)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma	0	0	0	1	0.191	.	.	0.336
		[67]	[38]	[35]	[33]	.	.	.	.
	B-Polyp, Endometrial Stromal	10	5	5	5	0.488	0.668	0.634	0.634
		[68]	[38]	[35]	[34]	.	.	.	.
	M-Carcinoma	1	1	1	1	0.328	0.583	0.569	0.561
		[68]	[39]	[35]	[33]	.	.	.	.
	M-Schwannoma, Malignant	1	0	0	0	1.000	1.000	1.000	1.000
		[67]	[38]	[35]	[33]	.	.	.	.
Vagina		(100)	(50)	(50)	(50)	.	.	.	.
	B-Granular Cell Tumor	0	0	1	1	0.113	.	0.342	0.336
		[67]	[38]	[35]	[33]	.	.	.	.
	B-Stromal Polyp	0	1	0	0	0.613	0.353	.	.
		[67]	[38]	[35]	[33]	.	.	.	.
	M-Schwannoma, Malignant	1	1	0	0	0.850	0.583	1.000	1.000
		[68]	[39]	[35]	[33]	.	.	.	.

Numbers with parentheses are number of the animals with organs examined and also usable

Numbers with brackets are survival-adjusted group size

Numbers are the tumor bearing animals

Table 4A: Intercurrent Mortality Rate
Male Mice

	VEHICLE_CONTROL1		VEHICLE_CONTROL2		LOW		MEDIUM		HIGH	
Week	NO.OF DEATH	PERCENT	NO.OF DEATH	PERCENT	NO.OF DEATH	PERCENT	NO.OF DEATH	PERCENT	NO.OF DEATH	PERCENT
0-52	6	12.0%	5	10.0%	2	4.0%	9	18.0%	6	12.0%
53-78	8	28.0%	5	20.0%	9	22.0%	7	32.0%	11	34.0%
79-92	5	38.0%	9	38.0%	9	40.0%	10	52.0%	8	50.0%
93-96	4	46.0%	3	44.0%	2	44.0%	6	64.0%	2	54.0%
Term. Sac.	27	100.0%	28	100.0%	28	100.0%	18	100.0%	23	100.0%

Table 4B: Intercurrent Mortality Rate
Female Mice

	VEHICLE_CONTROL1		VEHICLE_CONTROL2		LOW		MEDIUM		HIGH	
Week	NO.OF DEATH	PERCENT	NO.OF DEATH	PERCENT	NO.OF DEATH	PERCENT	NO.OF DEATH	PERCENT	NO.OF DEATH	PERCENT
0-52	1	2.0%	8	16.0%	13	26.0%	4	8.0%	10	20.0%
53-78	7	16.0%	13	42.0%	7	40.0%	11	30.0%	13	46.0%
79-92	10	36.0%	5	52.0%	9	58.0%	7	44.0%	6	58.0%
93-101	7	50.0%	7	66.0%	5	68.0%	6	56.0%	3	64.0%
Term. Sac.	25	100.0%	17	100.0%	16	100.0%	22	100.0%	18	100.0%

Table 5A: Intercurrent Mortality Comparison
Male Mice

Test	P-Value (across four groups)	P-Value (vehicle_control1 vs low)	P-Value (vehicle_control1 vs medium)	P-Value (vehicle_control1 vs high)
Dose Response	0.0003	0.4864	0.8788	0.0002
Homogeneity	0.3211	0.3394	0.7330	0.0473

Test	P-Value (across four groups)	P-Value (vehicle_control2 vs low)	P-Value (vehicle_control2 vs medium)	P-Value (vehicle_control2 vs high)
Dose Response	0.0005	0.8368	0.5903	<.0001
Homogeneity	0.4044	0.8811	0.6384	0.0172

Test	P-Value (across four groups)	P-Value (combined vehicle_controls vs low)	P-Value (combined vehicle_controls vs medium)	P-Value (combined vehicle_controls vs high)
Dose Response	0.0002	0.6058	0.6316	<.0001
Homogeneity	0.2952	0.5252	0.9340	0.0089

Table 5B: Intercurrent Mortality Comparison
Female Mice

Test	P-Value (across four groups)	P-Value (vehicle_control1 vs low)	P-Value (vehicle_control1 vs medium)	P-Value (vehicle_control1 vs high)
Dose Response	0.8488	<.0001	0.5922	0.0947
Homogeneity	0.7429	0.0663	0.2956	0.0780

Test	P-Value (across four groups)	P-Value (vehicle_control2 vs low)	P-Value (vehicle_control2 vs medium)	P-Value (vehicle_control2 vs high)
Dose Response	0.6109	0.0371	0.3431	0.8748
Homogeneity	0.8886	0.8384	0.2039	0.5523

Test	P-Value (across four groups)	P-Value (combined vehicle_controls vs low)	P-Value (combined vehicle_controls vs medium)	P-Value (combined vehicle_controls vs high)
Dose Response	0.9467	0.0002	0.8373	0.2736
Homogeneity	0.8763	0.3359	0.9037	0.4952

Table 6A (1): Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons
Male Mice (Vehicle control1, low, medium and high dose groups)

Organ Name	Tumor Name	Vehicle Cont1 N=50	100 mg Low N=50	300 mg Med N=50	1000 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Adrenal, Cortex		(49)	(50)	(49)	(50)	.	.	.	.
	B-Adenoma, Cortical Cell	0	1	1	0	0.521	0.514	0.485	.
		[32]	[24]	[23]	[19]	.	.	.	.
	B-Adenoma, Subcapsular Cell	0	1	1	0	0.519	0.521	0.485	.
		[32]	[25]	[23]	[19]	.	.	.	.
Body, Whole/Cav		(50)	(50)	(50)	(50)	.	.	.	.
	B-Hemangioma	1	1	1	1	0.378	0.768	0.746	0.732
		[32]	[24]	[23]	[19]	.	.	.	.
	M-Hemangiosarcoma	3	2	2	2	0.409	0.847	0.813	0.794
		[33]	[25]	[23]	[19]	.	.	.	.
	M-Histiocytic Sarcoma	1	0	0	0	1.000	1.000	1.000	1.000
		[33]	[24]	[23]	[19]	.	.	.	.
	M-Lymphosarcoma	4	2	3	2	0.510	0.919	0.774	0.880
		[33]	[24]	[23]	[19]	.	.	.	.
	M-Mast Cell Tumor	0	1	0	0	0.674	0.514	.	.
		[32]	[24]	[23]	[19]	.	.	.	.
	M-Mesothelioma, Malignant	1	0	0	0	1.000	1.000	1.000	1.000
		[33]	[24]	[23]	[19]	.	.	.	.
Gl, Harderian		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma	6	3	1	3	0.525	0.936	0.992	0.903
		[33]	[25]	[23]	[19]	.	.	.	.
Gl, Subling Sal		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma	0	0	0	1	0.194	.	.	0.479
		[32]	[24]	[23]	[19]	.	.	.	.
Kidney		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Tubular	0	0	1	3	0.007	.	0.493	0.105
		[32]	[24]	[23]	[19]	.	.	.	.
	M-Carcinoma, Tubular	0	0	0	2	0.036	.	.	0.226
		[32]	[24]	[23]	[19]	.	.	.	.
	CARCINOMA+ADENOMA	0	0	1	5	<.0001	.	0.493	0.021
		[32]	[24]	[23]	[19]	.	.	.	.
	B-Adenoma, Hepatocellular	14	15	11	3	0.995	0.609	0.818	0.999
		[34]	[26]	[25]	[20]	.	.	.	.
	M-Carcinoma, Hepatocellular	5	6	5	1	0.906	0.560	0.594	0.982
		[33]	[25]	[23]	[19]	.	.	.	.
Lung		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Bronchiolar-Alveo	10	7	8	7	0.314	0.880	0.751	0.820

Table 6A (1) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Mice (Vehicle control1, low, medium and high dose groups)

Organ Name	Tumor Name	Vehicle Cont1 N=50	100 mg Low N=50	300 mg Med N=50	1000 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Lung	B-Adenoma, Bronchiolar-Alveo	[34]	[25]	[24]	[20]	.	.	.	.
	M-Carcinoma, Bronchiolar-Alv	7	6	3	3	0.745	0.779	0.943	0.943
		[34]	[26]	[23]	[20]	.	.	.	.
Rectum		(50)	(50)	(50)	(50)	.	.	.	.
	M-Leiomyosarcoma	0	1	0	0	0.677	0.521	.	.
		[32]	[25]	[23]	[19]	.	.	.	.
Seminal Vesicle		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma	1	0	0	0	1.000	1.000	1.000	1.000
		[32]	[24]	[23]	[19]	.	.	.	.
Skin/Subcutis		(50)	(49)	(50)	(49)	.	.	.	.
	M-Sarcoma	2	1	0	0	0.965	0.891	1.000	1.000
		[33]	[24]	[23]	[19]	.	.	.	.
Spleen		(50)	(50)	(50)	(50)	.	.	.	.
	M-Mesothelioma, Malignant	1	0	0	0	1.000	1.000	1.000	1.000
		[33]	[24]	[23]	[19]	.	.	.	.
Testis		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Interstitial Cell	2	1	2	1	0.496	0.891	0.682	0.864
		[32]	[24]	[23]	[19]	.	.	.	.
Thymus		(37)	(42)	(37)	(35)	.	.	.	.
Thyroid		(49)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Follicular Cell	1	0	0	0	1.000	1.000	1.000	1.000
		[32]	[24]	[23]	[19]	.	.	.	.

Numbers with parentheses are number of the animals with organs examined and also usable

Numbers with brackets are survival-adjusted group size

Numbers are the tumor bearing animals

Table 6A (2): Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons
Male Mice (Vehicle control2, low, medium and high dose groups)

Organ Name	Tumor Name	Vehicle Cont2 N=50	100 mg Low N=50	300 mg Med N=50	1000 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Adrenal, Cortex		(50)	(50)	(49)	(50)	.	.	.	.
	B-Adenoma, Cortical Cell	0	1	1	0	0.528	0.500	0.471	.
		[31]	[24]	[23]	[19]	.	.	.	.
	B-Adenoma, Subcapsular Cell	1	1	1	0	0.741	0.753	0.716	1.000
		[32]	[25]	[23]	[19]	.	.	.	.
Body, Whole/Cav		(50)	(50)	(50)	(50)	.	.	.	.
	B-Hemangioma	1	1	1	1	0.386	0.754	0.732	0.718
		[31]	[24]	[23]	[19]	.	.	.	.
	M-Hemangiosarcoma	1	2	2	2	0.207	0.510	0.467	0.448
		[31]	[25]	[23]	[19]	.	.	.	.
	M-Histiocytic Sarcoma	1	0	0	0	1.000	1.000	1.000	1.000
		[31]	[24]	[23]	[19]	.	.	.	.
	M-Lymphosarcoma	7	2	3	2	0.765	0.984	0.931	0.971
		[33]	[24]	[23]	[19]	.	.	.	.
	M-Mast Cell Tumor	0	1	0	0	0.680	0.500	.	.
		[31]	[24]	[23]	[19]	.	.	.	.
Gallbladder		(46)	(48)	(50)	(49)	.	.	.	.
	B-Adenoma	1	0	0	0	1.000	1.000	1.000	1.000
		[31]	[24]	[23]	[19]	.	.	.	.
Gl, Harderian		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma	8	3	1	3	0.691	0.974	0.998	0.951
		[33]	[25]	[23]	[19]	.	.	.	.
Gl, Subling Sal		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma	0	0	0	1	0.196	.	.	0.466
		[31]	[24]	[23]	[19]	.	.	.	.
Kidney		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Tubular	0	0	1	3	0.008	.	0.478	0.096
		[31]	[24]	[23]	[19]	.	.	.	.
	M-Carcinoma, Tubular	0	0	0	2	0.037	.	.	0.214
		[31]	[24]	[23]	[19]	.	.	.	.
	CARCINOMA+ADENOMA	0	0	1	5	<.0001	.	0.478	0.019
		[31]	[24]	[23]	[19]	.	.	.	.
Liver		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Hepatocellular	11	15	11	3	0.987	0.288	0.539	0.993
		[32]	[26]	[25]	[20]	.	.	.	.
	M-Carcinoma, Hepatocellular	2	6	5	1	0.779	0.140	0.170	0.853
		[31]	[25]	[23]	[19]	.	.	.	.
Lung		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Bronchiolar-Alveo	8	7	8	7	0.234	0.698	0.511	0.612
		[32]	[25]	[24]	[20]	.	.	.	.

Table 6A (2) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Mice (Vehicle control2, low, medium and high dose groups)

Organ Name	Tumor Name	Vehicle Cont2 N=50	100 mg Low N=50	300 mg Med N=50	1000 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
	M-Carcinoma, Bronchiolar-Alv	9	6	3	3	0.867	0.887	0.977	0.977
Pituitary		(49)	(49)	(50)	(49)	.	.	.	.
	B-Adenoma	1	0	0	0	1.000	1.000	1.000	1.000
		[31]	[24]	[23]	[19]	.	.	.	.
Rectum		(50)	(50)	(50)	(50)	.	.	.	.
	M-Leiomyosarcoma	0	1	0	0	0.684	0.507	.	.
		[31]	[25]	[23]	[19]	.	.	.	.
Skin/Subcutis		(50)	(49)	(50)	(49)	.	.	.	.
	M-Leiomyosarcoma	1	0	0	0	1.000	1.000	1.000	1.000
		[31]	[24]	[23]	[19]	.	.	.	.
	M-Sarcoma	0	1	0	0	0.680	0.500	.	.
		[31]	[24]	[23]	[19]	.	.	.	.
Spleen		(49)	(50)	(50)	(50)	.	.	.	.
Testis		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Interstitial Cell	2	1	2	1	0.505	0.880	0.659	0.853
		[31]	[24]	[23]	[19]	.	.	.	.

Numbers with parentheses are number of the animals with organs examined and also usable

Numbers with brackets are survival-adjusted group size

Numbers are the tumor bearing animals

Table 6A (3): Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Mice (Combined vehicle controls, low, medium and high dose groups)

Organ Name	Tumor Name	Combined Conts N=100	100 mg Low N=50	300 mg Med N=50	1000 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Adrenal, Cortex		(99)	(50)	(49)	(50)	.	.	.	.
	B-Adenoma, Cortical Cell	0	1	1	0	0.371	0.340	0.314	.
		[63]	[24]	[23]	[19]	.	.	.	.
	B-Adenoma, Subcapsular Cell	1	1	1	0	0.563	0.570	0.527	1.000
		[64]	[25]	[23]	[19]	.	.	.	.
Body, Whole/Cav		(100)	(50)	(50)	(50)	.	.	.	.
	B-Hemangioma	2	1	1	1	0.340	0.716	0.691	0.674
		[63]	[24]	[23]	[19]	.	.	.	.
	M-Hemangiosarcoma	4	2	2	2	0.269	0.680	0.629	0.599
		[64]	[25]	[23]	[19]	.	.	.	.
	M-Histiocytic Sarcoma	2	0	0	0	1.000	1.000	1.000	1.000
		[64]	[24]	[23]	[19]	.	.	.	.
	M-Lymphosarcoma	11	2	3	2	0.720	0.971	0.878	0.950
		[65]	[24]	[23]	[19]	.	.	.	.
	M-Mast Cell Tumor	0	1	0	0	0.512	0.340	.	.
		[63]	[24]	[23]	[19]	.	.	.	.
	M-Mesothelioma, Malignant	1	0	0	0	1.000	1.000	1.000	1.000
		[64]	[24]	[23]	[19]	.	.	.	.
Gallbladder		(91)	(48)	(50)	(49)	.	.	.	.
	B-Adenoma	1	0	0	0	1.000	1.000	1.000	1.000
		[63]	[24]	[23]	[19]	.	.	.	.
Gl, Harderian		(100)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma	14	3	1	3	0.738	0.969	0.997	0.942
		[65]	[25]	[23]	[19]	.	.	.	.
Gl, Subling Sal		(100)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma	0	0	0	1	0.147	.	.	0.309
		[63]	[24]	[23]	[19]	.	.	.	.
Kidney		(100)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Tubular	0	0	1	3	0.002	.	0.320	0.028
		[63]	[24]	[23]	[19]	.	.	.	.
	M-Carcinoma, Tubular	0	0	0	2	0.021	.	.	0.094
		[63]	[24]	[23]	[19]	.	.	.	.
	CARCINOMA+ADENOMA	0	0	1	5	<.0001	.	0.320	0.002
		[63]	[24]	[23]	[19]	.	.	.	.
Liver		(100)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Hepatocellular	25	15	11	3	0.984	0.385	0.675	0.999
		[66]	[26]	[25]	[20]	.	.	.	.
	M-Carcinoma, Hepatocellular	7	6	5	1	0.773	0.270	0.316	0.952
		[64]	[25]	[23]	[19]	.	.	.	.

Table 6A (3) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Male Mice (Combined vehicle controls, low, medium and high dose groups)

Organ Name	Tumor Name	Combined Conts N=100	100 mg Low N=50	300 mg Med N=50	1000 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Lung		(100)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Bronchiolar-Alveo	18	7	8	7	0.245	0.811	0.624	0.725
		[66]	[25]	[24]	[20]	.	.	.	.
	M-Carcinoma, Bronchiolar-Alv	16	6	3	3	0.847	0.860	0.975	0.971
		[66]	[26]	[23]	[20]	.	.	.	.
Pituitary		(99)	(49)	(50)	(49)	.	.	.	.
	B-Adenoma	1	0	0	0	1.000	1.000	1.000	1.000
		[63]	[24]	[23]	[19]	.	.	.	.
Rectum		(100)	(50)	(50)	(50)	.	.	.	.
	M-Leiomyosarcoma	0	1	0	0	0.515	0.346	.	.
		[63]	[25]	[23]	[19]	.	.	.	.
Seminal Vesicle		(100)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma	1	0	0	0	1.000	1.000	1.000	1.000
		[63]	[24]	[23]	[19]	.	.	.	.
Skin/Subcutis		(100)	(49)	(50)	(49)	.	.	.	.
	M-Leiomyosarcoma	1	0	0	0	1.000	1.000	1.000	1.000
		[63]	[24]	[23]	[19]	.	.	.	.
	M-Sarcoma	2	1	0	0	0.884	0.716	1.000	1.000
		[64]	[24]	[23]	[19]	.	.	.	.
Spleen		(99)	(50)	(50)	(50)	.	.	.	.
	M-Mesothelioma, Malignant	1	0	0	0	1.000	1.000	1.000	1.000
		[64]	[24]	[23]	[19]	.	.	.	.
Testis		(100)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Interstitial Cell	4	1	2	1	0.499	0.881	0.629	0.849
		[63]	[24]	[23]	[19]	.	.	.	.
Thymus		(84)	(42)	(37)	(35)	.	.	.	.
Thyroid		(99)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Follicular Cell	1	0	0	0	1.000	1.000	1.000	1.000
		[63]	[24]	[23]	[19]	.	.	.	.

Numbers with parentheses are number of the animals with organs examined and also usable

Numbers with brackets are survival-adjusted group size

Numbers are the tumor bearing animals

Table 6B (1): Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Female Mice (Vehicle control1, low, medium and high dose groups)

Organ Name	Tumor Name	Vehicle Cont1 N=50	100 mg Low N=50	300 mg Med N=50	1000 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Adrenal, Cortex		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Cortical Cell	1	1	0	0	0.911	0.675	1.000	1.000
		[26]	[20]	[23]	[17]	.	.	.	.
	B-Adenoma, Subcapsular Cell	1	1	1	0	0.787	0.675	0.726	1.000
		[25]	[20]	[23]	[17]	.	.	.	.
	M-Carcinoma, Subcap Cell	0	0	1	0	0.471	.	0.480	.
		[25]	[20]	[23]	[17]	.	.	.	.
Body, Whole/Cav		(50)	(50)	(50)	(50)	.	.	.	.
	B-Hemangioma	0	0	0	1	0.200	.	.	0.424
		[25]	[20]	[23]	[17]	.	.	.	.
	M-Hemangiosarcoma	3	3	1	4	0.184	0.530	0.932	0.353
		[25]	[21]	[24]	[19]	.	.	.	.
	M-Histiocytic Sarcoma	1	6	3	6	0.035	0.023	0.277	0.023
		[25]	[22]	[23]	[18]	.	.	.	.
	M-Lymphosarcoma	9	12	16	6	0.752	0.116	0.053	0.705
		[27]	[23]	[26]	[19]	.	.	.	.
	M-Mast Cell Tumor	1	0	0	0	1.000	1.000	1.000	1.000
		[25]	[20]	[23]	[17]	.	.	.	.
Bone, Other		(50)	(50)	(50)	(50)	.	.	.	.
	M-Osteosarcoma	0	1	0	0	0.709	0.427	.	.
		[25]	[21]	[23]	[17]	.	.	.	.
Brain		(50)	(50)	(50)	(50)	.	.	.	.
	M-Astrocytoma	0	1	0	0	0.709	0.427	.	.
		[25]	[21]	[23]	[17]	.	.	.	.
Cervix		(50)	(50)	(50)	(49)	.	.	.	.
	B-Leiomyoma	0	1	1	0	0.562	0.427	0.480	.
		[25]	[20]	[23]	[17]	.	.	.	.
	M-Leiomyosarcoma	0	0	0	1	0.200	.	.	0.424
		[25]	[20]	[23]	[17]	.	.	.	.
Gallbladder		(48)	(46)	(48)	(49)	.	.	.	.
	B-Adenoma	0	0	1	0	0.471	.	0.480	.
		[25]	[20]	[23]	[17]	.	.	.	.
Gl, Harderian		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma	4	3	0	1	0.884	0.645	1.000	0.940
		[25]	[20]	[23]	[17]	.	.	.	.

Table 6B (1) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Female Mice (Vehicle control1, low, medium and high dose groups)

Organ Name	Tumor Name	Vehicle Cont1 N=50	100 mg Low N=50	300 mg Med N=50	1000 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Liver		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Hepatocellular	0	0	1	0	0.471	.	0.480	.
		[25]	[20]	[23]	[17]	.	.	.	.
Lung		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Bronchiolar-Alveo	4	5	6	5	0.280	0.334	0.317	0.334
		[26]	[22]	[23]	[20]	.	.	.	.
	M-Carcinoma, Bronchiolar-Alv	1	2	0	1	0.533	0.389	1.000	0.672
		[25]	[20]	[23]	[17]	.	.	.	.
Muscle, Other		(50)	(50)	(50)	(50)	.	.	.	.
	M-Osteosarcoma	0	0	0	1	0.209	.	.	0.433
		[25]	[20]	[23]	[18]	.	.	.	.
Ovary		(50)	(50)	(50)	(49)	.	.	.	.
	B-Cystadenoma	2	0	1	1	0.481	1.000	0.859	0.809
		[26]	[20]	[23]	[17]	.	.	.	.
	B-Granulosa Cell Tumor	1	1	1	1	0.421	0.675	0.733	0.672
		[25]	[20]	[23]	[17]	.	.	.	.
	M-Granulosa Cell Tumor	0	1	0	0	0.706	0.427	.	.
		[25]	[20]	[23]	[17]	.	.	.	.
Pancreas		(50)	(50)	(49)	(50)	.	.	.	.
	B-Adenoma, Islet Cell	1	0	0	0	1.000	1.000	1.000	1.000
		[25]	[20]	[23]	[17]	.	.	.	.
Pituitary		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma	3	2	1	1	0.763	0.716	0.932	0.898
		[25]	[20]	[23]	[17]	.	.	.	.
Rectum		(50)	(50)	(48)	(49)	.	.	.	.
	M-Carcinoma	0	0	1	0	0.471	.	0.487	.
		[25]	[20]	[23]	[17]	.	.	.	.
Skin/Subcutis		(50)	(50)	(50)	(50)	.	.	.	.
	M-Sarcoma	1	0	0	1	0.362	1.000	1.000	0.665
		[25]	[20]	[23]	[17]	.	.	.	.
Spleen		(50)	(50)	(50)	(50)	.	.	.	.
	M-Osteosarcoma	0	0	0	1	0.200	.	.	0.424
		[25]	[20]	[23]	[17]	.	.	.	.
Thymus		(40)	(47)	(48)	(47)	.	.	.	.
	B-Thymoma, Benign	1	0	0	1	0.362	1.000	1.000	0.665

Table 6B (1) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Female Mice (Vehicle control1, low, medium and high dose groups)

Organ Name	Tumor Name	Vehicle Cont1 N=50	100 mg Low N=50	300 mg Med N=50	1000 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
		[25]	[20]	[23]	[17]	.	.	.	.
	M-Thymoma, Malignant	0	0	1	0	0.471	.	0.480	.
		[25]	[20]	[23]	[17]	.	.	.	.
Thyroid		(49)	(50)	(49)	(50)	.	.	.	.
	B-Adenoma, Follicular Cell	2	0	1	0	0.863	1.000	0.859	1.000
		[26]	[20]	[23]	[17]	.	.	.	.
Uterus		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Endometrial	1	0	0	0	1.000	1.000	1.000	1.000
		[25]	[20]	[23]	[17]	.	.	.	.
	B-Leiomyoma	1	0	1	0	0.729	1.000	0.733	1.000
		[25]	[20]	[24]	[17]	.	.	.	.
	B-Polyp, Endometrial Stromal	4	2	2	1	0.828	0.818	0.881	0.940
		[25]	[20]	[23]	[17]	.	.	.	.
	M-Carcinoma, Endometrial	0	0	0	2	0.042	.	.	0.184
		[25]	[20]	[23]	[18]	.	.	.	.
	M-Leiomyosarcoma	0	0	0	1	0.209	.	.	0.424
		[25]	[20]	[23]	[18]	.	.	.	.
	M-Osteosarcoma	1	0	0	0	1.000	1.000	1.000	1.000
		[25]	[20]	[23]	[17]	.	.	.	.
	M-Sarcoma, Endometrial Strom	1	0	1	0	0.723	1.000	0.733	1.000
		[25]	[20]	[23]	[17]	.	.	.	.

Numbers with parentheses are number of the animals with organs examined and also usable

Numbers with brackets are survival-adjusted group size

Numbers are the tumor bearing animals

Table 6B (2): Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Female Mice (Vehicle control2, low, medium and high dose groups)

Organ Name	Tumor Name	Vehicle Cont2 N=50	100 mg Low N=50	300 mg Med N=50	1000 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Adrenal, Cortex		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Cortical Cell	0	1	0	0	0.723	0.492	.	.
		[23]	[20]	[23]	[17]	.	.	.	.
	B-Adenoma, Subcapsular Cell	1	1	1	0	0.802	0.746	0.791	1.000
		[23]	[20]	[23]	[17]	.	.	.	.
	M-Carcinoma, Subcap Cell	0	0	1	0	0.482	.	0.539	.
		[23]	[20]	[23]	[17]	.	.	.	.
Body, Whole/Cav		(50)	(50)	(50)	(50)	.	.	.	.
	B-Hemangioma	1	0	0	1	0.370	1.000	1.000	0.737
		[23]	[20]	[23]	[17]	.	.	.	.
	M-Hemangiosarcoma	1	3	1	4	0.076	0.306	0.797	0.177
		[23]	[21]	[24]	[19]	.	.	.	.
	M-Histiocytic Sarcoma	7	6	3	6	0.304	0.688	0.971	0.663
		[26]	[22]	[23]	[18]	.	.	.	.
	M-Lymphosarcoma	14	12	16	6	0.949	0.745	0.606	0.986
		[26]	[23]	[26]	[19]	.	.	.	.
	M-Mast Cell Tumor	1	0	0	0	1.000	1.000	1.000	1.000
		[23]	[20]	[23]	[17]	.	.	.	.
Bone, Other		(50)	(50)	(50)	(50)	.	.	.	.
	M-Osteosarcoma	0	1	0	0	0.726	0.492	.	.
		[23]	[21]	[23]	[17]	.	.	.	.
Brain		(50)	(50)	(50)	(50)	.	.	.	.
	M-Astrocytoma	0	1	0	0	0.726	0.492	.	.
		[23]	[21]	[23]	[17]	.	.	.	.
Cervix		(50)	(50)	(50)	(49)	.	.	.	.
	B-Leiomyoma	0	1	1	0	0.579	0.492	0.539	.
		[23]	[20]	[23]	[17]	.	.	.	.
	M-Leiomyosarcoma	1	0	0	1	0.366	1.000	1.000	0.737
		[24]	[20]	[23]	[17]	.	.	.	.
Gallbladder		(49)	(46)	(48)	(49)	.	.	.	.
	B-Adenoma	0	0	1	0	0.482	.	0.539	.
		[23]	[20]	[23]	[17]	.	.	.	.
Gl, Harderian		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma	2	3	0	1	0.754	0.484	1.000	0.868
		[23]	[20]	[23]	[17]	.	.	.	.
Liver		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Hepatocellular	0	0	1	0	0.482	.	0.539	.
		[23]	[20]	[23]	[17]	.	.	.	.

Table 6B (2) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Female Mice (Vehicle control2, low, medium and high dose groups)

Organ Name	Tumor Name	Vehicle Cont2 N=50	100 mg Low N=50	300 mg Med N=50	1000 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Lung		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Bronchiolar-Alveo	5	5	6	5	0.374	0.591	0.608	0.613
		[25]	[22]	[23]	[20]	.	.	.	.
	M-Carcinoma, Bronchiolar-Alv	6	2	0	1	0.943	0.963	1.000	0.992
		[25]	[20]	[23]	[17]	.	.	.	.
Mammary, Female		(45)	(42)	(42)	(47)	.	.	.	.
	M-Carcinoma	2	0	0	0	1.000	1.000	1.000	1.000
		[23]	[20]	[23]	[17]	.	.	.	.
Muscle, Other		(50)	(50)	(50)	(50)	.	.	.	.
	M-Osteosarcoma	0	0	0	1	0.214	.	.	0.492
		[23]	[20]	[23]	[18]	.	.	.	.
Ovary		(50)	(50)	(50)	(49)	.	.	.	.
	B-Cystadenoma	1	0	1	1	0.335	1.000	0.791	0.737
		[23]	[20]	[23]	[17]	.	.	.	.
	B-Granulosa Cell Tumor	0	1	1	1	0.237	0.492	0.539	0.483
		[23]	[20]	[23]	[17]	.	.	.	.
	B-Luteoma	1	0	0	0	1.000	1.000	1.000	1.000
		[23]	[20]	[23]	[17]	.	.	.	.
	M-Granulosa Cell Tumor	0	1	0	0	0.723	0.492	.	.
		[23]	[20]	[23]	[17]	.	.	.	.
Pituitary		(50)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma	2	2	1	1	0.675	0.681	0.907	0.868
		[23]	[20]	[23]	[17]	.	.	.	.
Rectum		(50)	(50)	(48)	(49)	.	.	.	.
	M-Carcinoma	0	0	1	0	0.482	.	0.546	.
		[23]	[20]	[23]	[17]	.	.	.	.
Skin/Subcutis		(50)	(50)	(50)	(50)	.	.	.	.
	M-Sarcoma	1	0	0	1	0.370	1.000	1.000	0.737
		[23]	[20]	[23]	[17]	.	.	.	.
Spleen		(50)	(50)	(50)	(50)	.	.	.	.
	M-Osteosarcoma	0	0	0	1	0.205	.	.	0.483
		[23]	[20]	[23]	[17]	.	.	.	.
Thymus		(47)	(47)	(48)	(47)	.	.	.	.
	B-Thymoma, Benign	0	0	0	1	0.205	.	.	0.483
		[23]	[20]	[23]	[17]	.	.	.	.
	M-Thymoma, Malignant	0	0	1	0	0.482	.	0.539	.

Table 6B (2) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Female Mice (Vehicle control2, low, medium and high dose groups)

Organ Name	Tumor Name	Vehicle Cont2 N=50	100 mg Low N=50	300 mg Med N=50	1000 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Thyroid		(50)	(50)	(49)	(50)	.	.	.	.
	B-Adenoma, Follicular Cell	0	0	1	0	0.482	.	0.539	.
		[23]	[20]	[23]	[17]	.	.	.	.
Ureter		(48)	(49)	(49)	(47)	.	.	.	.
Uterus		(50)	(50)	(50)	(50)	.	.	.	.
	B-Leiomyoma	0	0	1	0	0.488	.	0.546	.
		[23]	[20]	[24]	[17]	.	.	.	.
	B-Polyp, Endometrial Stromal	2	2	2	1	0.642	0.668	0.759	0.868
		[24]	[20]	[23]	[17]	.	.	.	.
	M-Carcinoma, Endometrial	0	0	0	2	0.044	.	.	0.237
		[23]	[20]	[23]	[18]	.	.	.	.
	M-Leiomyosarcoma	0	0	0	1	0.214	.	.	0.483
		[23]	[20]	[23]	[18]	.	.	.	.
	M-Sarcoma, Endometrial Strom	2	0	1	0	0.873	1.000	0.907	1.000
		[24]	[20]	[23]	[17]	.	.	.	.
Vagina		(50)	(48)	(48)	(50)	.	.	.	.
	M-Carcinoma, Squamous Cell	1	0	0	0	1.000	1.000	1.000	1.000
		[24]	[20]	[23]	[17]	.	.	.	.

Numbers with parentheses are number of the animals with organs examined and also usable

Numbers with brackets are survival-adjusted group size

Numbers are the tumor bearing animals

Table 6B (3): Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Female Mice (Combined vehicle controls, low, medium and high dose groups)

Organ Name	Tumor Name	Combined Conts N=100	100 mg Low N=50	300 mg Med N=50	1000 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Adrenal, Cortex		(100)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Cortical Cell	1	1	0	0	0.800	0.506	1.000	1.000
		[49]	[20]	[23]	[17]	.	.	.	.
	B-Adenoma, Subcapsular Cell	2	1	1	0	0.775	0.656	0.717	1.000
		[48]	[20]	[23]	[17]	.	.	.	.
	M-Carcinoma, Subcap Cell	0	0	1	0	0.370	.	0.340	.
		[48]	[20]	[23]	[17]	.	.	.	.
Body, Whole/Cav		(100)	(50)	(50)	(50)	.	.	.	.
	B-Hemangioma	1	0	0	1	0.291	1.000	1.000	0.500
		[48]	[20]	[23]	[17]	.	.	.	.
	M-Hemangiosarcoma	4	3	1	4	0.107	0.358	0.887	0.197
		[48]	[21]	[24]	[19]	.	.	.	.
	M-Histiocytic Sarcoma	8	6	3	6	0.104	0.197	0.780	0.186
		[50]	[22]	[23]	[18]	.	.	.	.
	M-Lymphosarcoma	23	12	16	6	0.820	0.351	0.184	0.925
		[53]	[23]	[26]	[19]	.	.	.	.
	M-Mast Cell Tumor	2	0	0	0	1.000	1.000	1.000	1.000
		[48]	[20]	[23]	[17]	.	.	.	.
Bone, Other		(100)	(50)	(50)	(50)	.	.	.	.
	M-Osteosarcoma	0	1	0	0	0.560	0.296	.	.
		[48]	[21]	[23]	[17]	.	.	.	.
Brain		(100)	(50)	(50)	(50)	.	.	.	.
	M-Astrocytoma	0	1	0	0	0.560	0.296	.	.
		[48]	[21]	[23]	[17]	.	.	.	.
Cervix		(100)	(50)	(50)	(49)	.	.	.	.
	B-Leiomyoma	0	1	1	0	0.415	0.296	0.340	.
		[48]	[20]	[23]	[17]	.	.	.	.
	M-Leiomyosarcoma	1	0	0	1	0.291	1.000	1.000	0.500
		[48]	[20]	[23]	[17]	.	.	.	.
Gallbladder		(97)	(46)	(48)	(49)	.	.	.	.
	B-Adenoma	0	0	1	0	0.370	.	0.340	.
		[48]	[20]	[23]	[17]	.	.	.	.
Gl, Harderian		(100)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma	6	3	0	1	0.862	0.532	1.000	0.919
		[48]	[20]	[23]	[17]	.	.	.	.
Liver		(100)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Hepatocellular	0	0	1	0	0.370	.	0.340	.
		[48]	[20]	[23]	[17]	.	.	.	.

Table 6B (3) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Female Mice (Combined vehicle controls, low, medium and high dose groups)

Organ Name	Tumor Name	Combined Conts N=100	100 mg Low N=50	300 mg Med N=50	1000 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Lung		(100)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Bronchiolar-Alveo	9	5	6	5	0.277	0.403	0.398	0.414
		[50]	[22]	[23]	[20]	.	.	.	.
	M-Carcinoma, Bronchiolar-Alv	7	2	0	1	0.878	0.804	1.000	0.942
		[49]	[20]	[23]	[17]	.	.	.	.
Mammary, Female		(91)	(42)	(42)	(47)	.	.	.	.
	M-Carcinoma	2	0	0	0	1.000	1.000	1.000	1.000
		[48]	[20]	[23]	[17]	.	.	.	.
Muscle, Other		(100)	(50)	(50)	(50)	.	.	.	.
	M-Osteosarcoma	0	0	0	1	0.165	.	.	0.299
		[48]	[20]	[23]	[18]	.	.	.	.
Ovary		(100)	(50)	(50)	(49)	.	.	.	.
	B-Cystadenoma	3	0	1	1	0.464	1.000	0.812	0.751
		[49]	[20]	[23]	[17]	.	.	.	.
	B-Granulosa Cell Tumor	1	1	1	1	0.260	0.506	0.566	0.500
		[48]	[20]	[23]	[17]	.	.	.	.
	B-Luteoma	1	0	0	0	1.000	1.000	1.000	1.000
		[48]	[20]	[23]	[17]	.	.	.	.
	M-Granulosa Cell Tumor	0	1	0	0	0.556	0.296	.	.
		[48]	[20]	[23]	[17]	.	.	.	.
Pancreas		(100)	(50)	(49)	(50)	.	.	.	.
	B-Adenoma, Islet Cell	1	0	0	0	1.000	1.000	1.000	1.000
		[48]	[20]	[23]	[17]	.	.	.	.
Pituitary		(100)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma	5	2	1	1	0.749	0.671	0.924	0.882
		[48]	[20]	[23]	[17]	.	.	.	.
Rectum		(100)	(50)	(48)	(49)	.	.	.	.
	M-Carcinoma	0	0	1	0	0.370	.	0.346	.
		[48]	[20]	[23]	[17]	.	.	.	.
Skin/Subcutis		(100)	(50)	(50)	(50)	.	.	.	.
	M-Sarcoma	2	0	0	1	0.402	1.000	1.000	0.645
		[49]	[20]	[23]	[17]	.	.	.	.
Spleen		(100)	(50)	(50)	(50)	.	.	.	.
	M-Osteosarcoma	0	0	0	1	0.157	.	.	0.292
		[48]	[20]	[23]	[17]	.	.	.	.

Table 6B (3) Continued: Tumor Rates and P-Values for Dose Response Relationship and Pair-wise Comparisons Female Mice (Combined vehicle controls, low, medium and high dose groups)

Organ Name	Tumor Name	Combined Conts N=100	100 mg Low N=50	300 mg Med N=50	1000 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
Thymus		(87)	(47)	(48)	(47)	.	.	.	.
	B-Thymoma, Benign	1	0	0	1	0.291	1.000	1.000	0.500
Thymus	B-Thymoma, Benign	[48]	[20]	[23]	[17]	.	.	.	.
	M-Thymoma, Malignant	0	0	1	0	0.370	.	0.340	.
		[48]	[20]	[23]	[17]	.	.	.	.
Thyroid		(99)	(50)	(49)	(50)	.	.	.	.
	B-Adenoma, Follicular Cell	2	0	1	0	0.756	1.000	0.712	1.000
		[49]	[20]	[23]	[17]	.	.	.	.
Uterus		(100)	(50)	(50)	(50)	.	.	.	.
	B-Adenoma, Endometrial	1	0	0	0	1.000	1.000	1.000	1.000
		[48]	[20]	[23]	[17]	.	.	.	.
	B-Leiomyoma	1	0	1	0	0.613	1.000	0.575	1.000
		[48]	[20]	[24]	[17]	.	.	.	.
	B-Polyp, Endometrial Stromal	6	2	2	1	0.773	0.742	0.831	0.916
		[49]	[20]	[23]	[17]	.	.	.	.
	M-Carcinoma, Endometrial	0	0	0	2	0.026	.	.	0.087
		[48]	[20]	[23]	[18]	.	.	.	.
	M-Leiomyosarcoma	0	0	0	1	0.165	.	.	0.292
		[48]	[20]	[23]	[18]	.	.	.	.
	M-Osteosarcoma	1	0	0	0	1.000	1.000	1.000	1.000
		[48]	[20]	[23]	[17]	.	.	.	.
	M-Sarcoma, Endometrial Strom	3	0	1	0	0.856	1.000	0.812	1.000
		[49]	[20]	[23]	[17]	.	.	.	.
Vagina		(100)	(48)	(48)	(50)	.	.	.	.
	M-Carcinoma, Squamous Cell	1	0	0	0	1.000	1.000	1.000	1.000
		[49]	[20]	[23]	[17]	.	.	.	.

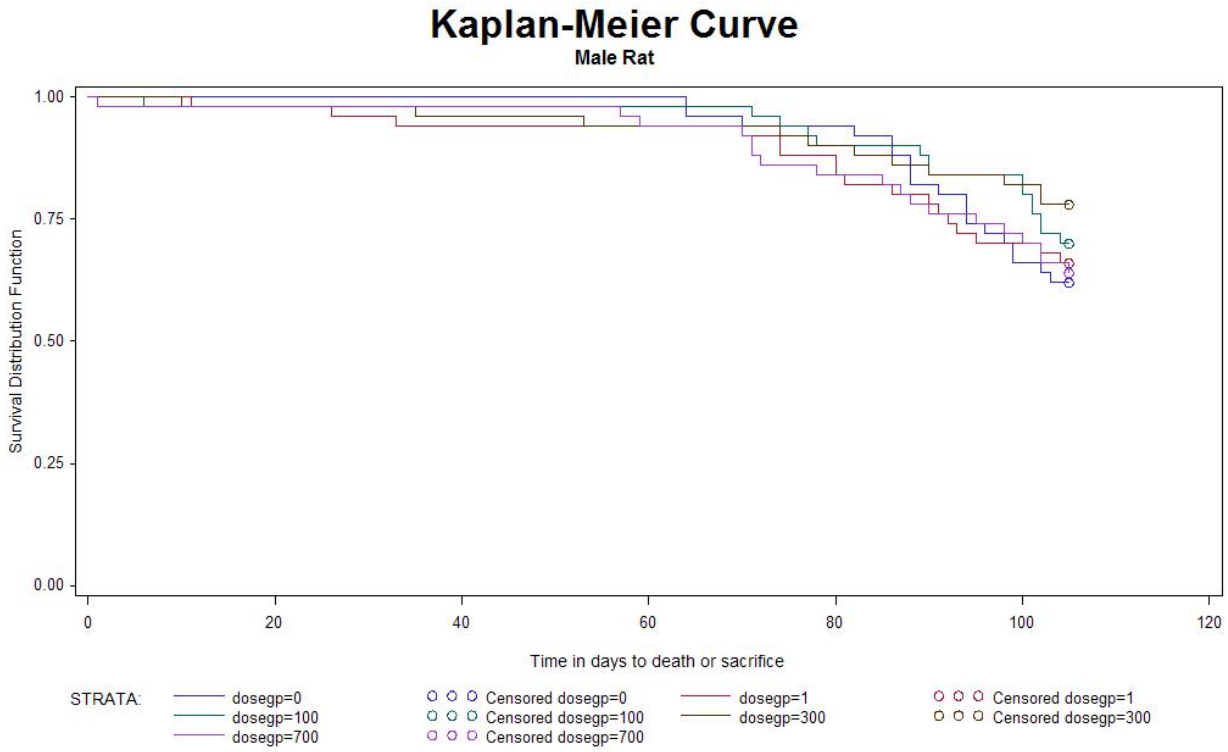
Numbers with parentheses are number of the animals with organs examined and also usable

Numbers with brackets are survival-adjusted group size

Numbers are the tumor bearing animals

Figure 1A: Kaplan-Meier Survival Functions for Male Rats

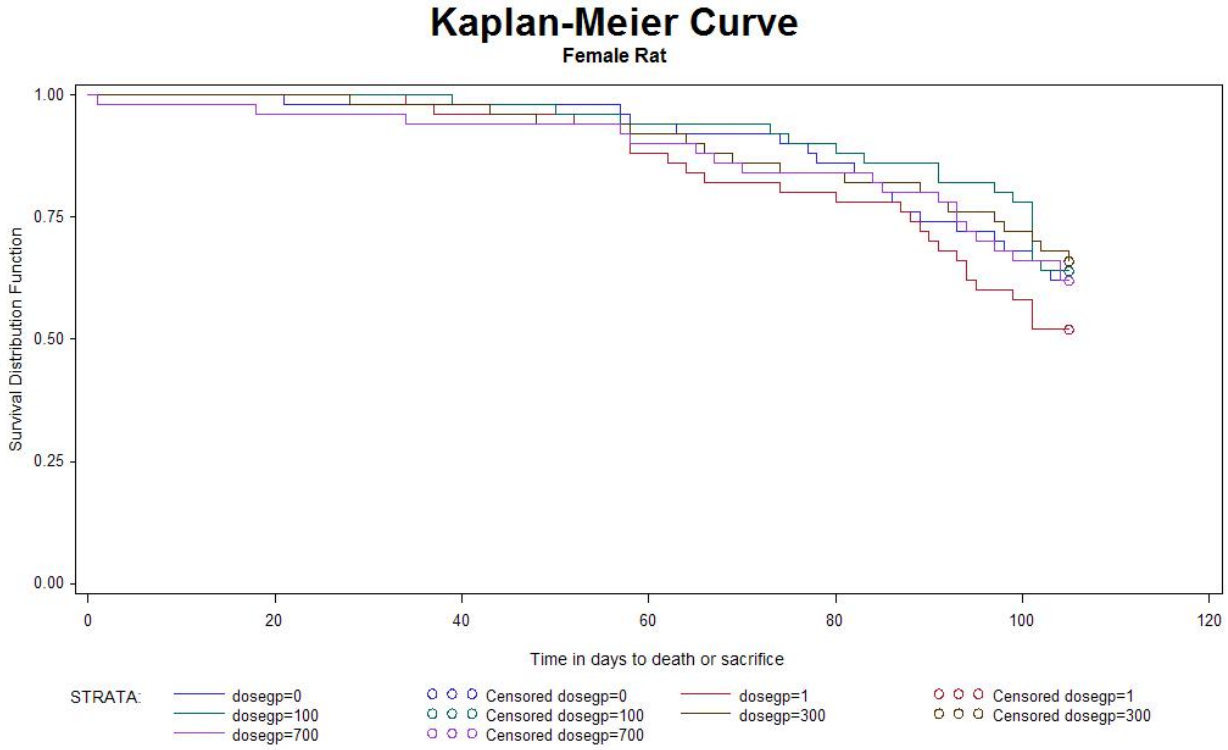
Male Rats (two vehicle controls, low, medium and high dose groups)



X-Axis: Weeks, Y-Axis: Survival rates

Figure 1B: Kaplan-Meier Survival Functions for Female Rats

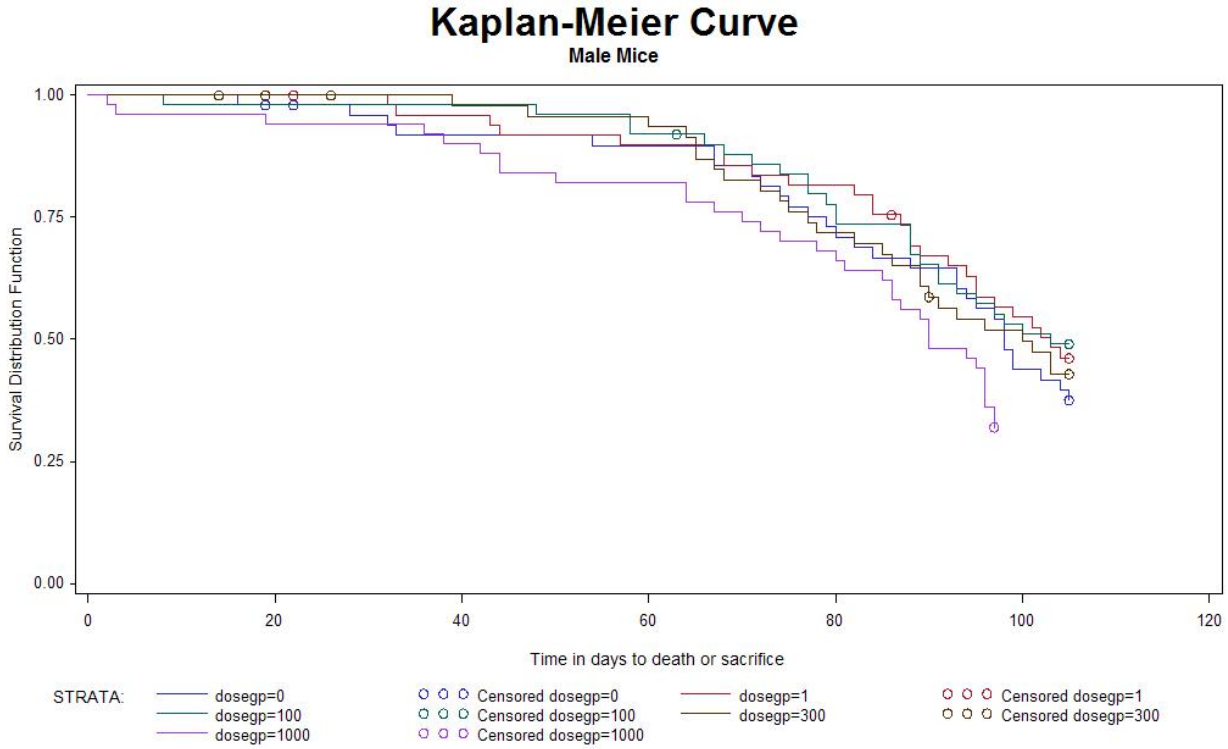
Female Rats (two vehicle controls, low, medium and high dose groups)



X-Axis: Weeks, Y-Axis: Survival rates

Figure 2A: Kaplan-Meier Survival Functions for Male Mice

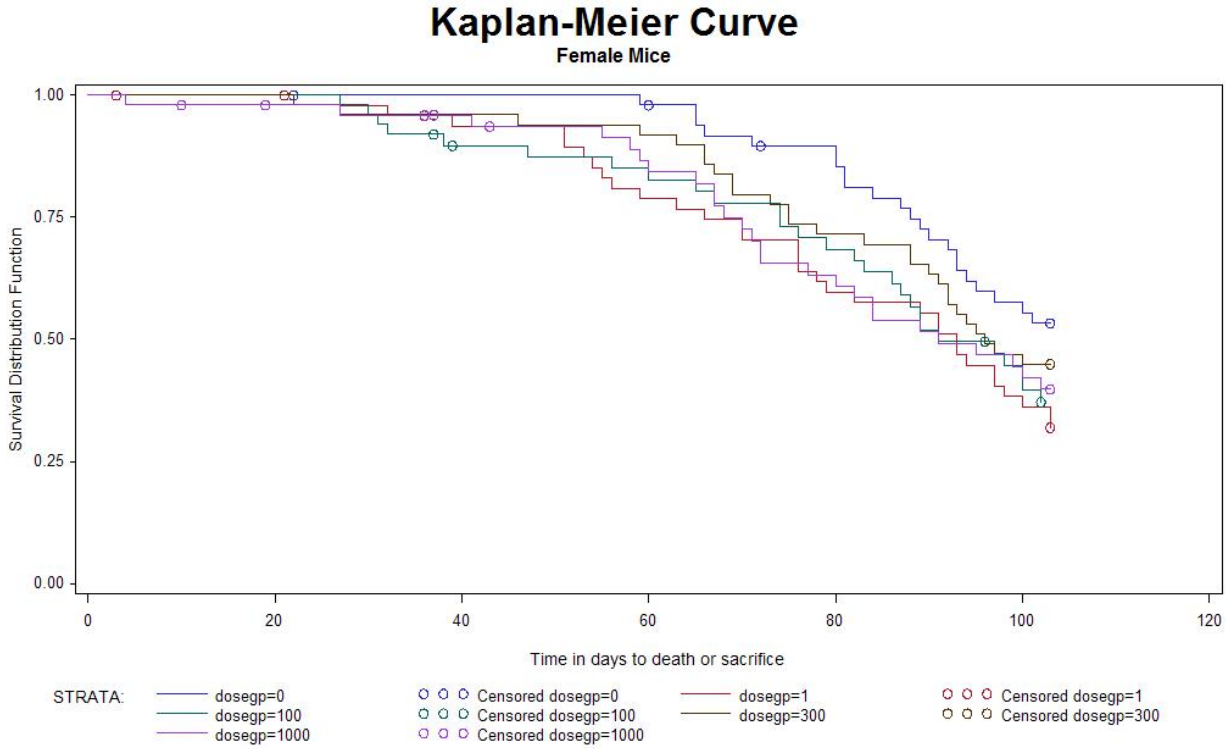
Male Mice (two vehicle controls, low, medium and high dose groups)



X-Axis: Weeks, Y-Axis: Survival rates

Figure 2B: Kaplan-Meier Survival Functions for Female Mice

Female Mice (two vehicle controls, low, medium and high dose groups)



X-Axis: Weeks, Y-Axis: Survival rates

7. References:

1. 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.

2. Bieler, G. S. and Williams, R. L. (1993). "Ratio estimates, the delta method, and quantal response tests for increased carcinogenicity". *Biometrics* 49, 793-801.

3. Cox D. R. (1972) "Regression models and life tables", *Journal of the Royal Statistical Society*, B, 34, 187-220.

4. Gehan (1965) "A generalized Wilcoxon test for comparing arbitrarily singly censored samples", *Biometrika*, 52, 203-223.

5. Haseman, J (1983), "A re-examination of false-positive rates for carcinogenesis studies", *Fundamental and Applied Toxicology*, 3: 334-339.

6. Lin, K.K. and Rahman, M.A. (1998), "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.

7. Rahman, M.A. and Lin, K.K. (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.

8. Peto, R., M.C. Pike, N.E. Day, R.G. Gray, P.N. Lee, S. Parish, J. Peto, Richards, and J.Wahrendorf (1980), "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.

9. Tarone RE (1975), "Test for trend in life table analysis", *Biometrika*, 62: 679-82.

10. U.S. Department of Health and Human Services, "Guidance for Industry: Statistical Aspects of the Design, Analysis, and Interpretation of Chronic Rodent Carcinogenicity Studies of Pharmaceuticals", Center for Drug E valuation and Research, Food and Drug Administration, Sliver Spring, Maryland, 2001.

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

/s/

MIN MIN
10/08/2013

KARL K LIN
10/15/2013
Concur with reivew

STATISTICS FINLING CHECKLIST FOR NDA 204629
Empagliflozin tablet

Filing meeting, April 23, 2013
Dongmei Liu, statistical reviewer

NDA Number: 204629 **Applicant: Boehringer Ingelheim** **Stamp Date: Mar. 5, 2013**

Drug Name: Empagliflozin **NDA/BLA Type: NDA** **PDUFA goal: Mar. 5, 2014**

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

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

IS THE STATISTICAL SECTION OF THE APPLICATION FILEABLE? ___Yes_____

Please identify and list any potential review issues to be forwarded to the Applicant for the 74-day letter. --- None at this time.

Content Parameter (possible review concerns for 74-day letter)	Yes	No	NA	Comment
Designs utilized are appropriate for the indications requested.	X			
Endpoints and methods of analysis are specified in the protocols/statistical analysis plans.	X			
Interim analyses (if present) were pre-specified in the protocol and appropriate adjustments in significance level made. DSMB meeting minutes and data are available.			X	No interim analysis for efficacy.
Appropriate references for novel statistical methodology (if present) are included.			X	Analysis used standard method.
Safety data organized to permit analyses across clinical trials in the NDA/BLA.	X			
Investigation of effect of dropouts on statistical analyses as described by applicant appears adequate.	X			

1. Drug information

- Proposed trade name: (b) (4)
- Generic name (components): empagliflozin
- Proposed indication: improve glycaemic control in adult patients with T2DM as an adjunct to diet and exercise
- Dosage form: 25 mg once daily
- Route of administration: oral tablet
- Applicant: BIPI
- Stamp date: March 5, 2013
- PDUFA date: March 5, 2014

2. Clinical studies

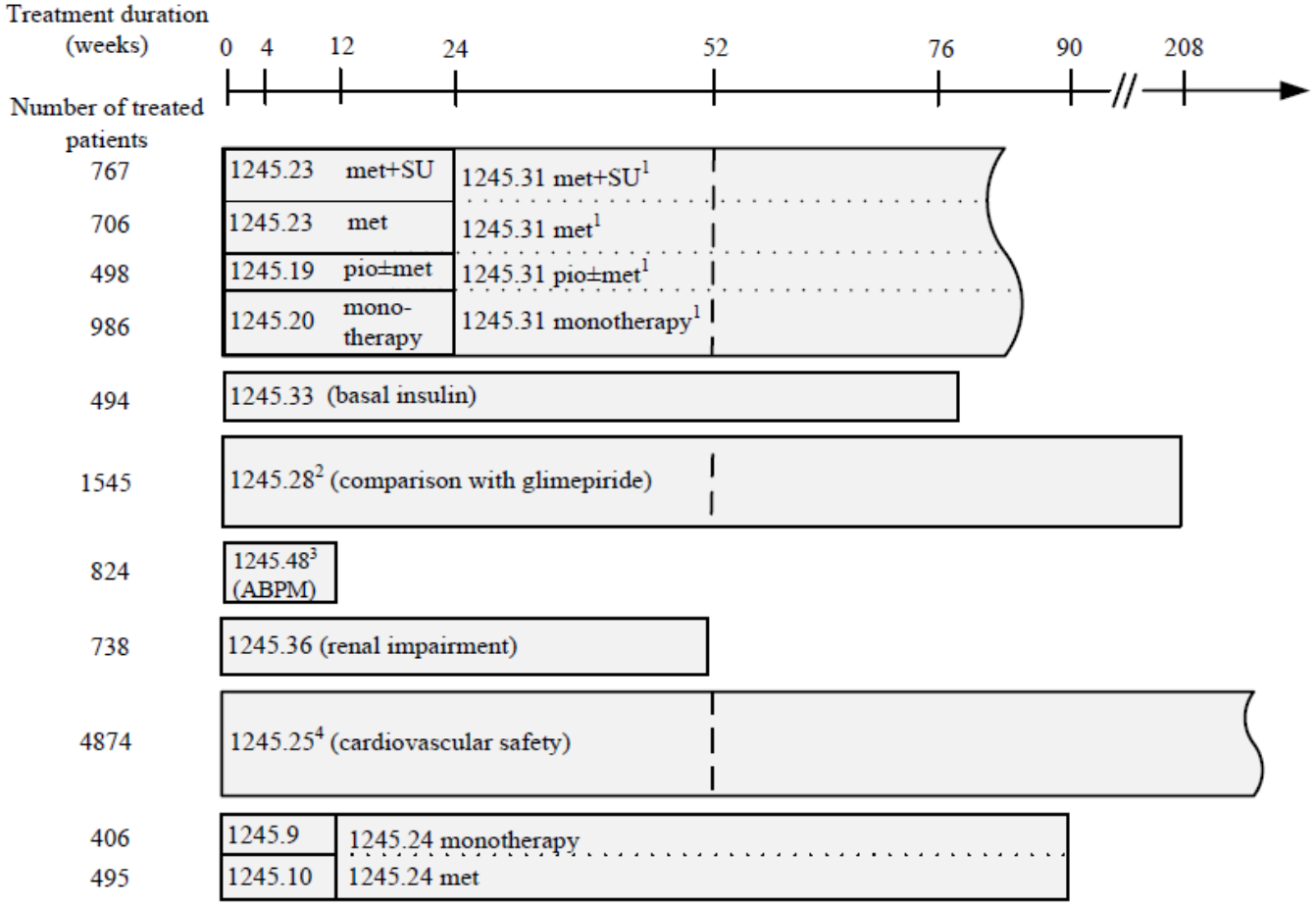


Figure 1.3.2: 1 Graphical overview of trials important for the evaluation of efficacy and safety of empagliflozin

3. Efficacy endpoints for pivotal studies

- Primary efficacy endpoint:
    - Change from baseline in HbA1c after 24 weeks of treatment
- Secondary efficacy endpoints (with some variation among trials)
    - Change from baseline in FPG, body weight, SBP, DBP after 24 weeks of treatment
    - HbA1c < 7% after 24 weeks of treatment

4. Statistical methods

- Change from baseline in HbA1c, FPG and body weight after 24 weeks of treatment
    - was analyzed using an analysis of variance (ANOVA) model with treatment, renal function at baseline, and region as fixed factors, and baseline variable of interest at covariates.
- Non inferiority test on the primary efficacy endpoint
    - Efficacy was considered confirmed if the upper bound of the 2-sided 95% CI for the estimated treatment difference for the mean change in HbA1c was below or equal to 0.3%.
- Missing data on HbA1c was imputed by last observation carried forward (LOCF) method. Sensitivity analysis and supportive analysis were done, which included analysis by repeated measurement model, completer analysis, analysis using available data exactly as observed, etc.

5. Data quality

Datasets were provided as SAS XPORT transport files version 5. I only had chance to test out one analysis dataset from trial 1245-0020. No problem has been detected so far.

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

/s/

DONGMEI LIU
04/23/2013

JON T SAHLROOT
04/29/2013

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

NDA Number: 204629

Applicant: Boehringer Ingelheim Pharmaceuticals Inc.

Stamp Date: 03/05/2013

Drug Name: Empagliflozin tablets, 10 mg or 25 mg

NDA Type: Standard review

Electronic submission, \\CDSESUB1\EVSPROD\NDA204629\204629.enx

This statistical safety filing review relates only contents of the submission relevant to the cardiovascular (CV) meta-analysis. A separate statistical filing review by Dr. Liu covers the overall efficacy and safety data contained in the application.

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

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

IS THE STATISTICAL SECTION OF THE APPLICATION FILEABLE? Yes

Potential review issues to be forwarded to the Applicant: Meeting minutes of data safety monitoring board (DSMB) and DSMB charter for the dedicated cardiovascular outcomes trial, 1245.25, were not included in the submission.

Content Parameter (possible review concerns for 74-day letter)	Yes	No	NA	Comment
Designs utilized are appropriate for the indications requested.	X			Cardiovascular outcomes trial was initiated in response to the guidance for type 2 diabetes drugs. Meta-analyses of CV safety use data from pivotal trials for efficacy and interim data from CV outcome trial.
Endpoints and methods of analysis are specified in the protocols/statistical analysis plans.	X			In the pivotal trials for efficacy, cardiovascular events were defined and

File name: StatisticsFilingChecklist_NDA204629

				adjudicated in a consistent manner.
Interim analyses (if present) were pre-specified in the protocol and appropriate adjustments in significance level made. DSMB meeting minutes and data are available.		X		DSMB meeting minutes are not available for the cardiovascular outcomes trial.
Appropriate references for novel statistical methodology (if present) are included.			X	
Safety data organized to permit analyses across clinical trials in the NDA/BLA.	X			See below*
Investigation of effect of dropouts on statistical analyses as described by applicant appears adequate.			X	

*The Applicant submitted analysis CV datasets for the 8 individual trials included in the meta-analysis, but no integrated CV datasets were submitted. Due to the formats of the individual trial data and the extent of the number of trials, compilation of the individual trial data into an integrated database by a reviewer would require significant resources that would not be subject to the same quality control measures implemented by the Applicant. Therefore, a request for integrated datasets along with descriptions of the desired data format was sent to the Applicant on 3/25/13. In response, the integrated datasets were submitted to the application on 4/12/13 and can be found at

EDR location: \\Cdsesub1\evsprod\NDA204629\0001\m5\datasets.

See discussion of preliminary findings using these datasets in the summary of the meta-analysis section that follows.

Background

Empagliflozin is a novel, oral, and selective inhibitor of sodium-dependent glucose co-transporter 2 (SGLT-2). The Applicant seeks for empagliflozin to be "indicated as an adjunct to diet and exercise to improve glycemic control in adults with type 2 diabetes mellitus". According to the 2008 FDA diabetes guidance, the cardiovascular meta-analysis was conducted to rule out an excess risk margin of 1.8 as measured by the hazard ratio. (b) (4) Separate alpha spending functions were used to test the 1.8 and 1.3 risk margins.

Brief summary of cardiovascular meta-analysis

The cardiovascular meta-analysis was based on data from completed randomized double blind controlled phase II and III trials with treatment duration longer than 12 weeks as well as randomized double-blind phase III trials that were ongoing at the time of the CV meta-analysis and had a pre-planned interim analysis. Table 1 and 2, extracted from the study report, provide descriptions of the completed and ongoing trials included in the meta-analysis. According to the study report, trial 1245.23 comprised two separate studies (metformin or metformin plus sulfonylurea as background therapy), under one study number. Note also that ongoing trial 1245.31 represents the extension phases of trials 1245.19, 1245.20, and 1245.23. There was no

plan to assess the effect of dropouts from the main trials to the extension trial. Ongoing trials also include one Phase 3 randomized, double-blind, placebo-controlled cardiovascular outcome trial 1245.25. Interim data from this trial were included in the meta-analysis. A supportive meta-analysis will be performed at the end of this trial.

Table 1: Completed Phase II and Phase III Trials Included in CV Meta-analysis

Trial ID Clinical phase Reference	Objective / Trial design	Parallel treatment groups (excluding open label arms)/ Treatment duration	Number of randomised patients (excluding open label arms)
1245.19 Phase III [U12-1516-02]	Empa with pioglitazone background therapy Randomised, double-blind, placebo-controlled	Placebo Empa 10 mg Empa 25 mg 24 weeks	499
1245.20 Phase III [U12-1517-01]	Empa as monotherapy Randomised, double-blind, placebo- and active-controlled, uncontrolled open-label group for poorly controlled patients (HbA_{1c} >10%)	Placebo Empa 10 mg Empa 25 mg Sitagliptin 100 mg 24 weeks	899
1245.23 Phase III [U12-1518-01]	Empa with metformin or metformin plus sulfonylurea background therapy (two separate studies). Randomised, double-blind, placebo-controlled; uncontrolled open-label group for poorly controlled patients (HbA_{1c} >10%)	Placebo Empa 10 mg Empa 25 mg 24 weeks	638 (metformin) 669 (metformin plus sulfonylurea)
1245.36 Phase III	Empa in patients with various degrees of renal impairment Randomised, double-blind, placebo-controlled	Placebo Empa 10 mg Empa 25 mg 52 weeks	741
1245.33 Phase II	Empa with basal insulin ± metformin ± sulfonylurea background therapy Randomised, double-blind, placebo-controlled	Placebo Empa 10 mg Empa 25 mg 78 weeks	494

Table 2: Ongoing Phase III Trials at the Time of the CV Meta-analysis but with Formal Interim Data Included in the Meta-analysis

Trial ID Clinial phase Reference	Title	Parallel treatment groups / Treatment duration	Number of randomised patients
1245.25 Phase III -	Cardiovascular safety/outcome trial Randomised, double-blind, placebo-controlled	Placebo Empa 10 mg Empa 25 mg Event-driven (estimated: 8 years)	7000 planned to be randomised overall. (b) (4) (b) (4)
1245.28 Phase III (interim CTR)	Empa versus glimepiride; with metformin background therapy Randomised, double-blind, active-controlled	Empa 25 mg Glimepiride (maximal tolerated dose between 1 mg and 4 mg) 208 weeks	1549
1245.31 Phase III [U12-1521-01] (interim CTR)	Long-term safety and efficacy of empa [extension of 1245.19, 20, and 23] Randomised, double-blind, placebo- and active-controlled	Placebo; Empa 10 mg Empa 25 mg Sitagliptin 100 mg Overall: minimum of 76 weeks (incl. 24 weeks of preceding trial) The interim analysis was performed after the DBL of the last preceding trial (1245.19)	Continued in extension trial: 305 (1245.19/pioglitazone) 615 (1245.20/drug-naive) 463 (1245.23/metformin) 473 (1245.23/metformin plus sulfonylurea)

The primary endpoint for the meta-analysis was a composite endpoint comprising CV death, nonfatal MI, nonfatal stroke, or hospitalization due to unstable angina. All CV events were adjudicated prospectively by an independent Clinical Event Committee. Approximately 10000 randomized patients were included in the meta-analysis, the majority of patients (6200) received empagliflozin and 3800 patients received comparator medication. Database lock for the interim meta-analysis was August 31, 2012. Because the number of events at this date exceeded the number planned for the final analysis (that is, 182 adjudicated events), this interim meta-analysis was regarded as the final meta-analysis for the 1.8 risk margin. According to the study report, there were a total of (b) (4) primary CV events ((b) (4) in the empagliflozin treated and (b) (4) in the comparator treated arms) as of the database lock date, resulting in hazard ratio (b) (4), 95% CI (b) (4)

***Reviewer's Comment**: The reviewer was able to use the requested integrated safety data to verify the overall number of reported primary events in both treatment arms as well as the number of subjects in both treatment arms. The reviewer was also able to verify the number of subjects in both treatment arms for each trial as shown in Table 7.5.1:2 of the study report. Replication of the Applicant's hazard ratio and results of other analyses was not attempted at the time of this filing review and will be addressed in the statistical safety review.*

Brief summary of cardiovascular outcome trial at interim analysis

Trial 1245.25 is a phase 3 multicenter cardiovascular outcomes trial. Approximately 7000 patients were planned to be enrolled. As of the interim cut-off date for the trial, there were (b) (4) events (b) (4) in empagliflozin and (b) (4) in placebo) positively adjudicated by committee from this trial, and included in the CV meta-analysis. No interim report for this trial was submitted to the

NDA. Case narratives of death, serious adverse events and events leading to discontinuation are provided in the submission. As noted in the pre-NDA meeting package, the analyses were performed by a team independent of the 1245.25 trial team so that the 1245.25 trial team remained blinded to the results.

Table 3: Ongoing CV Outcomes Trial Design and Interim Results for 1.8 Risk Margin

Study number/patient population	Design	Treatment arms/Sample size ITT	Primary endpoint/Analysis	Interim results
1245.25 Patients with type 2 diabetes mellitus	Multinational, double-blind, 3 parallel groups comparing 2 daily doses of empagliflozin (10 mg and 25 mg) to placebo as add-on to standard of care	- Placebo (1513) - Empagliflozin (3046)	Primary endpoint for the trial: Time to first occurrence of any of the following events as adjudicated by the CEC: CV death, nonfatal MI, or nonfatal stroke. Analyze findings with Cox proportional hazards model	Cardiovascular events (primary endpoint*) The hazard ratio is (b) (4) with a 95% CI of (b) (4) for empagliflozin compared with placebo.
*Primary endpoint for meta-analysis, that is, composite of CV death, nonfatal MI, nonfatal stroke or hospitalization for unstable angina.				

Reviewer's Comment: The datasets submitted (as well as the integrated datasets) contain an additional (b) (4) ***) enrolled in this trial, but whose randomization dates exceed the trial's interim analysis cut-off date. Therefore, these subjects are not included in the table above and Applicant's results in the study report.***

Janelle K. Charles 04-25-2013
Reviewing Statistician Date

Mat Soukup 04-25-2013
Supervisor/Team Leader Date

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

/s/

JANELLE K CHARLES
04/25/2013

MATTHEW J SOUKUP
04/25/2013
concur