CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

206619Orig1s000

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 - Addendum
Clinical Studies

NDA/BLA Serial Number:	206619 0034 (35) dated 9/16/2014 & 0023 (24) dated 7/30/2014
Drug Name:	Viekira Pak [Ombitasvir/Paritaprevir/Ritonavir tablets co-packaged with Dasabuvir]
Indication(s):	Treatment of chronic Hepatitis C Virus infection, including patients with cirrhosis
Applicant:	AbbVie, Inc.
Date(s):	Original application Received April 21, 2014
	PDUFA date December 21, 2014
Review Priority:	Priority 8-month clock
Biometrics Division:	Division of Biometrics IV
Statistical Reviewer:	Joy Mele, M.S.
Concurring Reviewers:	Greg Soon, Ph.D., Biometrics IV Team Leader
	Dionne Price, Ph.D., Biometrics IV Division Director
Medical Division:	Division of Antiviral Products
Clinical Team:	Russell Fleischer, PA-C, MPH Clinical Reviewer
	Linda Lewis, M.D. Clinical Team Leader
Project Manager:	Katherine Schumann, M.S.

Keywords: Clinical trials, non-inferiority, historical controls, HCV, Ombitasvir, Paritaprevir, Dasabuvir

1. Introduction

1.1 Overview

The applicant (AbbVie) has submitted an NDA for a combination product of three direct acting agents (DAA) to be marketed as Viekira Pak (called 3DAA in this review). This product is composed of a fixed dose tablet of ombitasvir [ABT-267 12.5 mg], paritaprevir [ABT-450 75 mg] and ritonavir [ABT-538 50 mg] co-packaged with dasabuvir [ABT-333 250 mg]. An indication for the treatment of genotype 1 (GT1) chronic hepatitis C infection (HCV) in patients with or without cirrhosis is being sought by the applicant. A statistical review dated September 19, 2014 by this reviewer was completed of the original application and is available in the CDER document room.

The results of two additional studies were submitted after the original submission and these results are briefly reviewed in Section 2 of this review. These studies are small studies in special populations (Table 1.1.1). Study M12-999 is a single arm study in 34 patients who had a liver transplantation. Study 14-004 is a two-arm study comparing 12 weeks of 3DAA plus ribavirin (RBV) to 24 weeks of treatment in patients co-infected with HCV and HIV.

Table 1.1.1 Phase 2 studies of 3DAA used in special populations

Study	Design	Treatment Period	Follow-up Period	Randomized Arms (ITT N)	HCV Study Population
M12-999	OL, MC	24 weeks	48 weeks	3DAA+RBV (34)	GT 1 liver transplant patients
M14-004	OL, R, MC	12 or 24 weeks	48 weeks	3DAA+RBV 12 wk (31) 3DAA+RBV 24 wk (32)	GT 1 patients with HIV-1 co-infection

OL=open label, R=randomized, MC=multicenter

Endpoints and statistical methods were consistent with what was defined for the Phase 3 trials and are not described here; details may be found in the statistical review of the original application dated September 19, 2014.

In Section 3 of this review, statistical comments regarding labeling are provided.

1.2 Data Sources and Quality

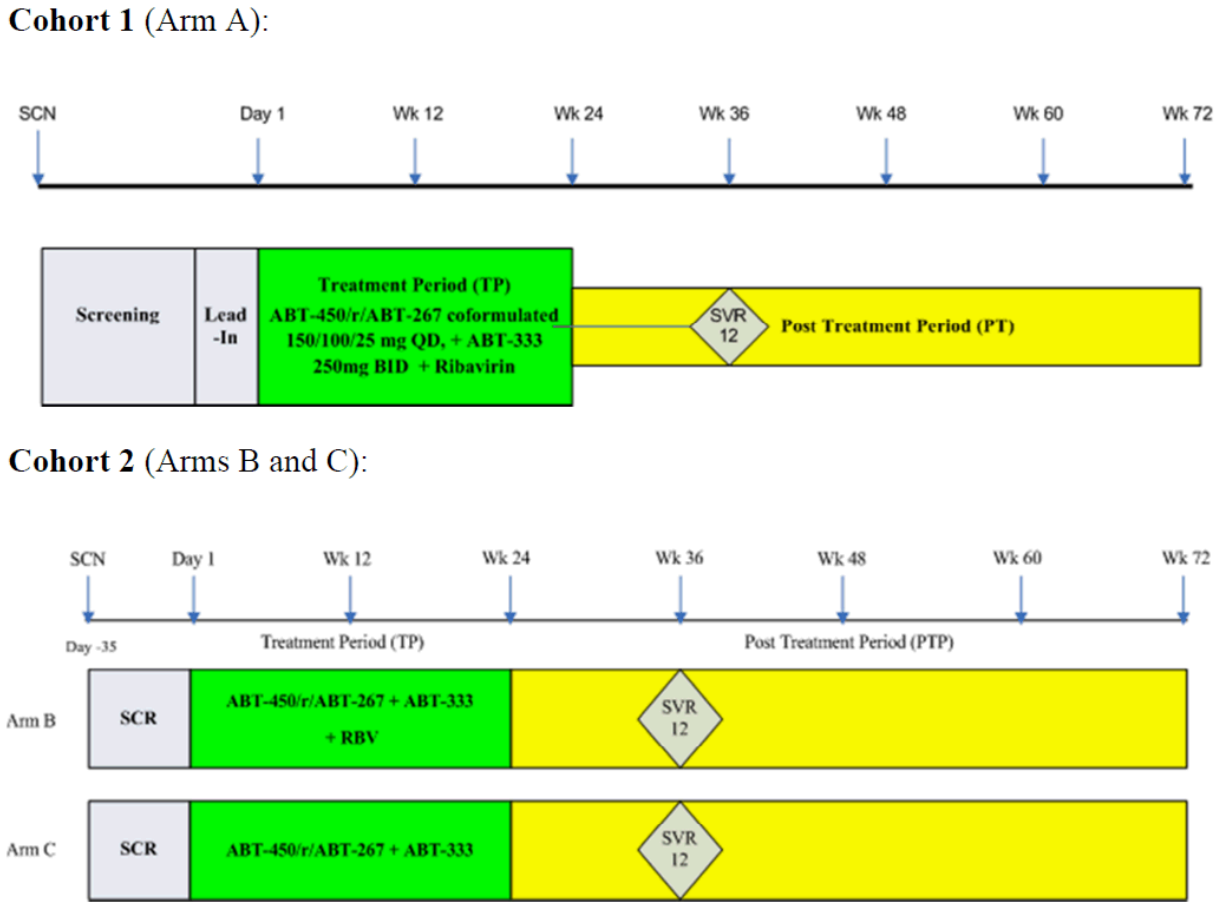
The data was provided as tabulation files and as analysis files. The application may be accessed in Global Submit at the following link: \\CDSESUB1\evsprod\NDA206619\206619.enx .

2. Statistical Evaluation of Efficacy

2.1 Study M12-999

Study M12-999 is an open-label, multi-center study in adult liver transplant recipients with recurrent HCV genotype 1 infection. The applicant planned to conduct two parts for Study M12-999; Cohort 1 and Cohort 2. Enrollment and follow-up for Cohort 1 is complete and is the subject of this review; Cohort 2 is currently enrolling (Figure 2.1.1.1).

Figure 2.1.1.1 Applicant's trial schematic



A total of 34 HCV genotype 1-infected subjects with fibrosis ≤ F2 (Metavir) who were treatment-naïve after transplantation but may have received previous HCV treatment (pegIFN or IFN with or without RBV) prior to liver transplantation, enrolled under Cohort 1, were treated with 3DAA plus ribavirin (RBV) for 24 weeks. Immunosupressant treatment with CNIs, cyclosporine or tacrolimus at a stable dose was allowed. One subject discontinued treatment after 18 weeks due to memory impairment, anxiety and rash; this patient achieved an SVR_{12}.

The majority of patients were male, white, GT1a and non-CC. The average time since transplantation was about 4 years.

Table 2.1.1.1 Study 12-999 Cohort 1 Demographics

	3DAA+RBV 24 weeks N=34
% male	79%
% white	85%
Age (years) Mean (SD) Min-Max	 60 (7) 30-71
HCV genotype 1a 1b	 85% 15%
IL28B GT CC	24%
Baseline HCV RNA log 10 Mean (SD)	 6.6 (0.5)
Baseline fibrosis F0-F1 F2	 56% 44%
Donor type Living Deceased	 12% 88%
Immunosupressive med Cyclosporine Tacrolimus	 15% 85%
Time since liver transplantation (mos) Mean (SD) Median Min-Max % > 5 years post transplantation	 48 (33) 40 13-136 26%
Baseline Cr clearance (mL min) mean	90.5

The SVR_{12} rate was 97% with 95% CI from 85% to 99.9% (Table 2.1.1.2).

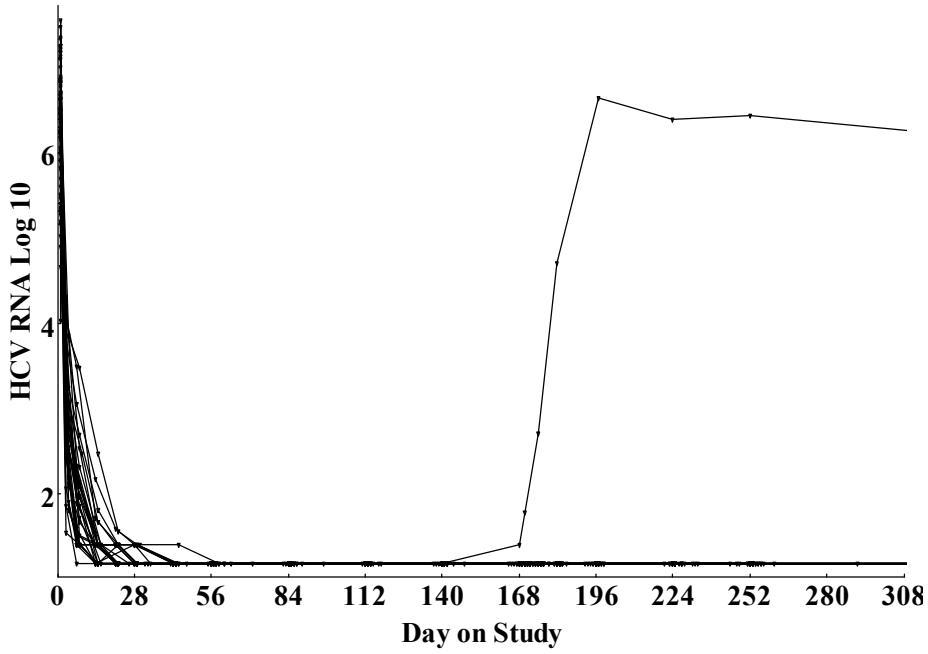
Table 2.1.1.2 Study 12-999 Cohort 1 efficacy results

	3DAA+RBV 24 weeks N=34
Overall Outcome for patients not achieving SVR_{12} On-treatment failure Relapse Missing	33/34 97% 95% CI 85%, 99.9%[1] 0 1 post-trt day 3 0
GT 1a GT 1b	28/29 97% 5/5 100%

[1]Clopper Pearson exact confidence interval

Only one patient did not achieve an SVR_{12} after 24 weeks of treatment due to a relapse on Day 3 as illustrated in Figure 2.1.1.2. The patient is a female, 54 years, black, baseline HCV RNA of 18,400,000 (log 10 7.3), 1a, and non-cc.

Figure 2.1.1.2 Study M12-999 Cohort 1 HCV RNA by patient



Although this study is uncontrolled, the evidence supports treatment with 3DAA+RBV for 24 weeks in a liver transplant population with a 97% SVR_{12} which is consistent with the rates observed in the Phase 3 trials.

2.2 Study M14-004

Study M14-004 is a Phase 2/3 randomized, open label, study in HCV GT1 human immunodeficiency virus type 1 (HIV-1) co-infected adults. Pegylated interferon/ribavirin (pegIFN/RBV) treatment alone has shown to have limited effectiveness in this population.

This trial consists of three parts:

- Part 1a: Phase 2 trial with patients with an unquantifiable plasma HIV-1 RNA (HIV-1 RNA < 40 copies/mL with the Abbott RealTi*m*e HIV-1 Assay) and a CD4+ T-cell count ≥ 200 cells/mm3 or CD4+ T-cell % ≥ 14% while on a stable atazanavir (ATV) or raltegravir (RAL) containing HIV-1 antiretroviral treatment (ART) randomized stratified on HCV treatment history and presence of cirrhosis to 3DAA+RBV administered for 12 weeks or for 24 weeks.
- Part 1b: Phase 2 trial with patients on a stable QD darunavir (DRV)-containing HIV-1 ART regimen will be randomized to either maintain DRV QD or to switch to DRV BID administration during a pretreatment period. All patients will receive 3DAA+RBV for 12 weeks.
- Part 2: Phase 3 trial not yet initiated.

The report submitted by the applicant contains data collected until post-treatment week (PTW) 12 from Part 1a. A total of 62 patients were randomized at USA sites to treatment (31:12-week treatment and 32:24-week treatment) in Part 1a. One patient on the 12-week treatment discontinued treatment early due to withdrawing consent.

Table 2.1.2 on the following page shows the randomized treatment groups were well-balanced based on several baseline demographic variables. More than 90% of patients were male, about ¾ were white and the mean age was 51 years. About 89% were genotype 1a and about two-thirds were treatment naïve. About 19% had cirrhosis. The boxplots for CD4 count and HCV RNA at baseline show a similar distribution for the treatment groups (Figure 2.1.2.1).

Figure 2.1.2.1 Boxplots of baseline CD4 counts (left) & HCV RNA (right)

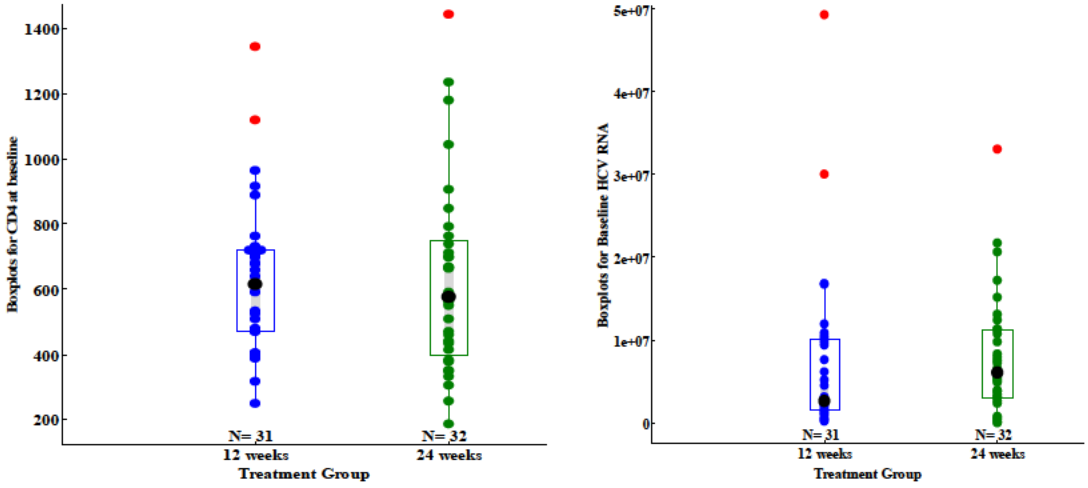


Table 2.1.2.1 Study 14-004 Part 1a Demographics

	3DAA+RBV 12 weeks N=31	3DAA+RBV 24 weeks N=32
% male	94%	91%
% white	77%	75%
Age (years)		
Mean (SD)	51 (6)	51 (8)
Min-Max	38-66	31-69
HCV genotype		
1a	87%	91%
1b	13%	9%
IL28B GT CC	16%	22%
HCV treatment naïve	65%	69%
HCV treatment exp	35%	31%
Null responder	16%	16%
Partial responder	16%	6%
Relapser	3%	9%
Baseline HCV RNA log 10		
Mean (SD)	6.5 (0.6)	6.6 (0.8)
Cirrhotic	19%	19%
Baseline CD4+T-cell count		
Mean (SD)	633 (236)	625 (296)
Median	614	575
HIV therapy		
Atazanavir	52%	38%
Raltegravir	48%	62%

The primary endpoint in this trial is the percentage of subjects achieving SVR12 within each treatment group. A two-sided 95% confidence interval was to be constructed using the Wilson score method for each proportion. As a secondary analysis, the percentages of subjects with SVR12 in the 24-week arm were to be compared to that of the 12-week arm using Fisher's exact test.

The applicant did not name a threshold for non-inferiority or superiority for comparisons to an historical control as was done for the Phase 3 trials so the criteria for showing efficacy was not prespecified. However, given that the trial is a randomized trial, a comparison of the two arms can give evidence in favor of one arm or the other. In addition, a margin of 10% for non-inferiority as done for the Phase 3 trials may be reasonable for comparing the arms.

In addition to comparing SVR12, the applicant planned to compare CD4 counts. That data is summarized on the following page. Other endpoints were also planned but they are not summarized here because they were not proposed for labeling.

At the end of treatment for all patients, there was a median drop in CD4 count of 47 in the 12 week arm and 60 in the 24 week arm (Figure 2.1.2.1). The boxplot suggests that either the median or the mean and standard deviation could represent the change in CD4 count at the end of treatment. After treatment is complete the CD4 count increases (Table 2.1.2.2 on the following page) with about 70% of patients in both arms returning to baseline. A graph (Figure 2.1.2.2) on the following page) of the individual patient changes in CD4 count shows the variability amongst patients and no notable treatment arm differences.

Table 2.1.2.2 Study 14-004 CD4+ T-cell count results

	3DAA+RBV 12 weeks N=31	3DAA+RBV 24 weeks N=32
Baseline		
Mean (SD)	633 (236)	625 (296)
Median	614	576
Change from baseline		
Last value on treatment (ITT)	N=31	N=32
Mean (SD)	-110 (173)	-94 (133)
Median	-47	-60
EOT (Completers)	N=27	N=29
Mean (SD)	-108 (181)	-89 (129)
Median	-47	-62
% of patients returned to baseline CD4 count or higher at end of study	23/31 (74%)	22/32 (69%)
N (%) patients with any CD4+ on treatment <200	0/31 (0%)	2/32 (6%)

Figure 2.1.2.1 Boxplots of change from baseline in CD4+ T-cell count at end of treatment by treatment arm (ITT)

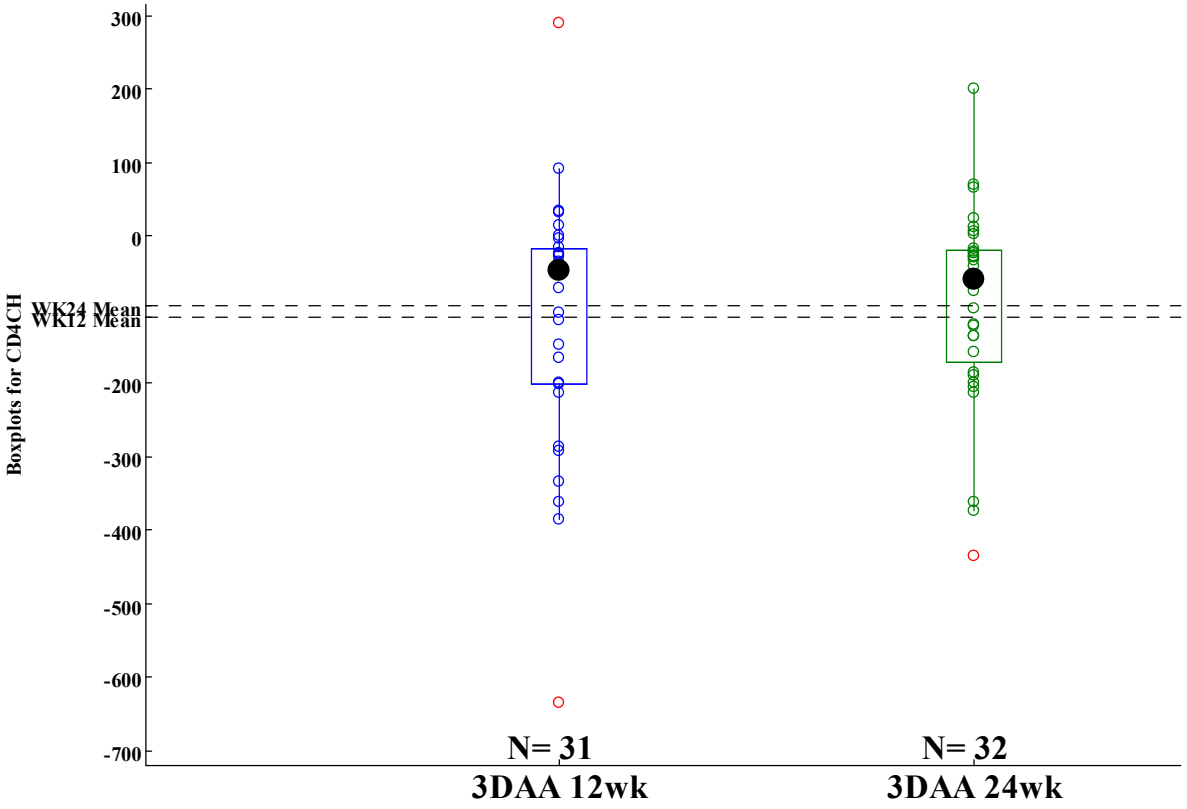
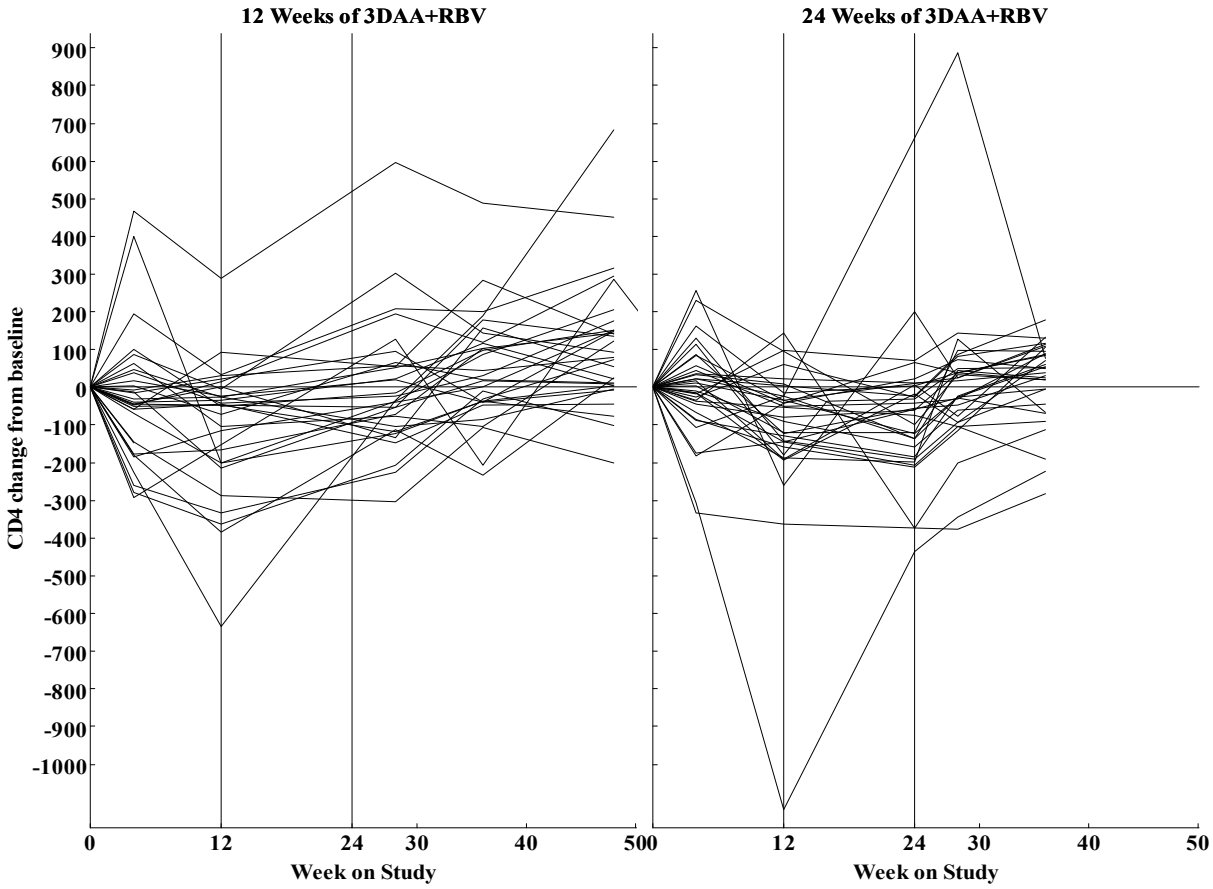


Figure 2.1.2.2 Study 14-004 CD4 change from baseline by patient and treatment



For SVR_{12}, the results for the two arms are not significantly different but favor the 12 week arm by about 3%. The confidence intervals for the arms have lower bounds of 79% and 75% but without a pre-specified threshold, it is not clear how to interpret these values. Two patients in the 12 week arm did not achieve an SVR; one relapsed at post-treatment week 2 and the other withdrew consent and discontinued treatment on Day 71 with a recorded HCV RNA of 15. Three patients in the 24 week arm did not achieve an SVR; one patient was an on treatment failure at treatment week 16 and two patients were recorded as relapses but were also suspected re-infections (these two cases are discussed in the FDA virologist's review).

Table 2.1.2.3 Study 14-004 SVR_{12} results

	3DAA+RBV 12 weeks N=31	3DAA+RBV 24 weeks N=32	24 wks-12 wks Trt diff (95%CI)
Overall	29/31 93.5% 95% CI 79%, 99%	29/32 91% 95% CI 75%, 98%	**-3% (-16%, +10%)**
Outcome for patients not achieving SVR_{12}			
On-treatment failure	0	1 TW16	
Relapse	1 PTW2	2 PTW 8,12[1]	
Missing	1 TW10	0[2]	
GT 1a	25/27 92.5%	26/29 90%	-3% (-18%, +12%)
GT 1b	4/4 100%	3/3 100%	0%
Cirrhotic	5/6 83%	5/6 83%	0% (-42%, +42%)
Not Cirrhotic	24/25 96%	24/26 92%	-4% (-17%, +9%)
Baseline HCV RNA By median			
< 5,140,000	18/18 100%	12/13 92%	-8% (-22%, +7%)
≥ 5,140,000	11/13 85%	17/19 90%	+5% (-19%, +29%)
Trt experience			
Naive	19/20 95%	20/22 91%	-4% (-19%, +11%)
Experienced	10/11 91%	9/10 90%	-1% (-26%, +24%)
Baseline CD4 count By median			
< 609	13/14 93%	16/17 94%	+1% (-16, +19%)
≥ 609	16/17 94%	13/15 87%	-7 (-28%, +13%)

[1] These two patients were identified by both the applicant and the FDA virologist as re-infections
[2] One patient had no viral load recorded at PTW12 but had SVR_8 and based on the plan to carry-forward data was counted as achieving an SVR_{12}

The one relapse in the 12 week arm (subject M14004-33471-105901) had the following characteristics: null responder, cirrhotic , non-cc, 58 years, male, baseline platelets of 141, BMI 23.5, baseline HCV RNA 16,700,000, baseline albumin>35, child pugh A, alpha feto protein 5.6. The applicant had recommended 12 weeks of treatment for this type of patient as of October 22, 2014. It appears though that this cirrhotic patient may have benefited from 24 weeks of treatment considering other risk factors.

The applicant concludes that the efficacy in this study of HCV/HIVsubjects is consistent with the efficacy in the Phase 3 studies and then recommends that these patients be treated as proposed for the HCV GT1 mono-infected patients. This reviewer thinks the data is insufficient to assume that subgroups of patients defined by genotype subtype and cirrhosis or not should be treated as recommended for HCV mono-infected patients. For example, the proposed labeling recommends

that GT1b patients without cirrhosis be treated with 3DAA for 12 weeks and that GT1b patients with cirrhosis be treated with 3DAA plus RBV for 12 weeks; with only 7 GT1b patients in this study, it is not at all clear how these HCV/HIV patients should be treated.

Overall the data supports treating HCV/HIV co-infected non-cirrhotic patients with 12 weeks of treatment with 3DAA+RBV with a 96% (24/25) SVR rate.

3. Labeling Comments

3.1 Statistical comments on the Clinical Trial section of the labeling

This reviewer's recommendations for Section 14 of the labeling are briefly summarized here. At the time of this review, a final labeling had not been agreed on between the FDA medical division and the applicant AbbVie. Several versions of Section 14.0 are documented in communications with the applicant.

Statistical comments on Section 14 included:

1. Reorganize by indication instead of by each study with the following headings:

   14.2 (b) (4)
   14.3 (b) (4)
   14.4 (b) (4)

2. Include treatment differences (b) (4) for comparisons of the randomized arms
   a. (b) (4)
   b. (b) (4)
3. Remove (b) (4) from the individual tables (b) (4).
4. Summarize demographics for all trials in Section 14.1.
5. Remove originally proposed (b) (4)
6. Include in a section regarding durability of response the SVR_{12} rate for the Phase 2 study summarized to demonstrate the correlation between the SVR_{12} and (b) (4) rates.
7. The proposed labeling for the co-infection study M14-004 states that "Median declines in the CD4+ T-cell counts of 47.0 and 62.0 cells/mm^3 were observed at the end of 12 and 24 weeks treatment respectively, (b) (4)." As shown in Table 2.1.2.2 and Figure 2.1.2.1 of this review, (b) (4) This reviewer would recommend additional details be included.

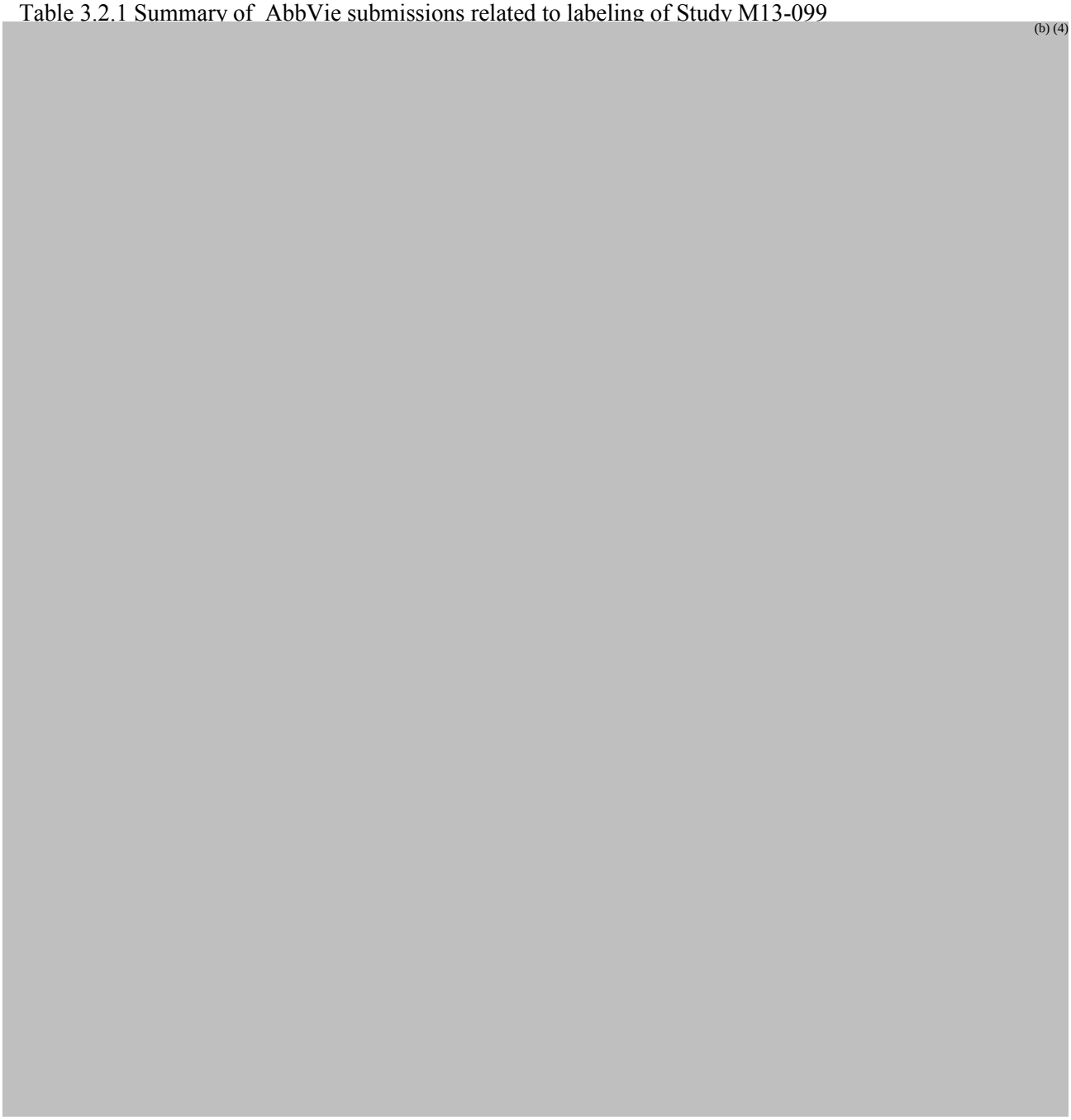
The labeling for the cirrhotic trial M13-099 involved several correspondences with the applicant and several internal discussions. These discussions are well documented in meeting minutes and correspondences. The statistical comments regarding labeling of the results of M13-099 are summarized in the following section (Section 3.2) of this review. At the writing of this review, no final decision regarding labeling the cirrhotic trial had been reached.

3.2 Statistical comments for labeling Study M13-099, a trial of GT 1a patients with cirrhosis

The table below summarizes the labeling that AbbVie has proposed for GT 1a patients with compensated cirrhosis since the original submission on April 21, 2014 to the submission on October 17, 2014 and the rationale for their proposals.

Table 3.2.1 Summary of AbbVie submissions related to labeling of Study M13-099

(b) (4)



The FDA medical division has consistently responded that all compensated cirrhotic GT 1a patients should be treated with 24 weeks of treatment.

This reviewer addressed length of treatment for 1a cirrhotics in the original statistical review. Here the additional results from the applicant are briefly reviewed and results of additional analyses by this reviewer are presented and discussed.

AbbVie performed several analyses that led to their proposal (b) (4)

The analyses were problematic and not acceptable alone as sufficient evidence (b) (4)

(b) (4)

This reviewer performed many subgroup analyses (b) (4)

The primary endpoint of SVR_{12} was used for all subgroup analyses by this reviewer because:

1. As already mentioned, the protocol specified that SVR_{12} would be used to compare treatment durations
    a. SVR_{12} is named as the primary outcome and is appropriate for comparing groups because all outcomes are counted within each randomized arm
    b. When comparing durations, lost-to-follow-ups should be considered because these losses may be treatment duration related. Ignoring losses can lead to biased results that are uninterpretable.

c. If relapse is the best measure to distinguish treatment durations, one would expect that the other outcomes would be similar in both groups begin compared and then relapse would be the outcome driving the difference. This is actually the case for the GT1a cirrhotic data where the predominant event in the Week 12 arm is relapse. Analyzing this outcome as part of a composite outcome ensures an unbiased comparison of the groups.

2. Relapse is a descriptive secondary endpoint and not an appropriate measure for making inferences
   a. Relapse is computed subsetting on an outcome (compliance and cure) which can lead to uninterpretable results because the groups being compared are no longer randomized groups
   b. Summarizing relapse numbers is important in conjunction with the analysis of the primary outcome, SVR_{12}

This reviewer's analyses of subgroups for Study M13-099 are summarized in Table 3.2.2 below. The results consistently show a higher SVR_{12} rate for 24 weeks of treatment compared to 12 weeks of treatment. The only exception is the treatment naïve CC population; however, the group is very small (a total of 35 patients in both groups) and the wide confidence interval suggests that a difference in favor of 24 weeks as large as 12% is possible.

Three subgroups based on platelets less than144 x 10^9/L, BMI less than 28 kg/m^2 and albumin greater than 39.5 g/L showed statistically significant results based on a nominal p-value; an adjustment for multiple comparisons may render these results non-significant. Results for several subgroups had small lower boundaries for the confidence interval on the difference suggesting that perhaps a larger trial would show a difference. Those groups include all 1a patients, treatment experienced patients, patients with high HCV RNA baseline, patients with alpha-fetoprotein greater than 11.4 ng/mL and patients with no history of IV drug use.

Table 3.2.2 Study M13-099 Reviewer's SVR_{12} results by subgroups for the 1a population

	3DAA+RBV 12 weeks	3DAA+RBV 24 weeks	24 wks-12 wks Trt diff (95%CI)
Overall 1a	124/140 89%	115/121 95%	+6 % (-0.6%, +13%)
TN	59/64 92%	53/56 95%	+2.5% (-7%, +12%)
CC	19/19 100%	15/16 94%	-6% (-31%, +12%)
Non-CC	40/45 89%	38/40 95%	+6% (-5%, +17%)
TE	65/76 85.5%	62/65 95%	+9.7% (-0.4%, +20%)
Null	40/50 80%	39/42 93%	+13% (-1.8%, +28%)
Partial	11/11 100%	10/10 100%	0% (-29%, +28%)
Relapser	14/15 93%	13/13 100%	+7% (-19%, +31%)
Baseline HCV RNA			
By median			
< 3,630,000	66/74 89%	52/56 93%	+4% (-6%, +13%)
≥ 3,630,000	58/66 88%	63/65 97%	+10% (-0.7%, +20%)
By tertiles			
<2,120,000	45/48 90%	36/38 95%	+5% (-6%, +16%)
2,120,000-5,690,000	43/47 91%	39/41 95%	+4% (-7%, +14%)
>5,690,000	38/45 84%	40/42 95%	+11% (-2%, +23%)
Sex			
Male	89/101 88%	84/89 94%	+6% (-2%, +14%)
Female	35/39 90%	31/32 97%	+7% (-4%, +18%)

	3DAA+RBV 12 weeks	3DAA+RBV 24 weeks	24 wks-12 wks Trt diff (95%CI)
Median Age years			
<57	67/74 91%	53/56 95%	+4% (-4%, +13%)
≥ 57	57/66 86%	62/65 95%	+9% (-1%, +19%)
Median BMI kg/m^2			
<28	58/67 87%	65/67 97%	**+10% (+1%, +20%)**
≥ 28	66/73 90%	50/54 93%	+2% (-8%, +12%)
Platelet counts			
< 90 x 10^9/L	14/18 78%	17/18 94%	+16% (-8%, +40%)
≥90 x 10^9/L	110/122 90%	98/103 95%	+5% (-2%, +12%)
By baseline median			
< 144 x 10^9/L	62/71 87%	56/58 97%	**+9% (+0.4%, +19%)**
≥ 144 x 10^9/L	62/69 90%	59/63 94%	+4% (-6%, +14%)
α-fetoprotein			
< 20 ng/mL	97/104 93%	81/84 96%	+3% (-4%, +10%)
≥ 20 ng/mL	27/36 75%	34/37 92%	+17% (-0.5%, +34%)
By baseline median			
< 11.4 ng/mL	69/74 93%	52/54 96%	+3% (-5.5%, +12%)
≥ 11.4 ng/mL	55/66 83%	63/67 94%	+10% (-0.5%, +21%)
Albumin			
< 35 g/L	12/16 75%	14/16 88%	+12.5% (-15%, +40%)
≥ 35 g/L	112/124 90%	101/105 96%	+6% (-0.7%, +13%)
By baseline median			
< 39.5 g/L	63/70 90%	58/63 92%	+2% (-8%, +12%)
≥ 39.5 g/L	61/70 87%	57/58 98%	**+11% (+2%, +20%)**
By baseline tertiles			
< 38 g/L	36/41 88%	34/38 89.5%	+2% (-13%, +17%)
38 to <41 g/L	42/47 89%	35/36 97%	+8% (-4%, +20%)
≥ 41 g/L	46/52 88.5%	46/47 98%	+9% (-1%, +20%)
Any 1 of the 3 AbbVie criteria	38/49 78%	45/50 90%	+12% (-2%, +27%)
None	86/91 95%	70/71 98%	+4% (-1%, +10%)
Hx of diabetes			
No	108/120 90%	97/102 95%	+5% (-2%, +12%)
Yes	16/20 80%	18/19 95%	+16% (-6%, +38%)
Hx of IV drug use			
No	66/72 92%	72/73 99%	+7% (-0.8%, +15%)
Yes	57/67 85%	43/48 90%	+4% (-9%, +17%)


(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/

JOY D MELE
10/31/2014

GUOXING SOON
10/31/2014

DIONNE L PRICE
10/31/2014
concur

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

Statistical Review and Evaluation
Clinical Studies

NDA/BLA Serial Number:	206619 0002 (3) dated 3/18/2014 & 0003 (4) dated 4/21/2014
Drug Name:	Viekira Pak [Ombitasvir/Paritaprevir/Ritonavir tablets co-packaged with Dasabuvir]
Indication(s):	Treatment of chronic Hepatitis C Virus infection, including patients with cirrhosis
Applicant:	AbbVie, Inc.
Date(s):	Received April 21, 2014
	PDUFA date December 21, 2014
Review Priority:	Priority 8-month clock
Biometrics Division:	Division of Biometrics IV
Statistical Reviewer:	Joy Mele, M.S.
Concurring Reviewers:	Greg Soon, Ph.D., Biometrics IV Team Leader
	Dionne Price, Ph.D., Biometrics IV Division Director
Medical Division:	Division of Antiviral Products
Clinical Team:	Russell Fleischer, PA-C, MPH Clinical Reviewer
	Linda Lewis, M.D. Clinical Team Leader
Project Manager:	Katherine Schumann, M.S.

Keywords: Clinical trials, non-inferiority, historical controls, HCV, Ombitasvir, Paritaprevir, Dasabuvir

Table of Contents

1 EXECUTIVE SUMMARY

The applicant (AbbVie) has submitted an NDA for a combination product of three direct acting agents [1] (DAA) to be marketed as Viekira Pak (called 3DAA in this review). This product is composed of a fixed dose tablet of ombitasvir [ABT-267 12.5 mg], paritaprevir [ABT-450 75 mg] and ritonavir [ABT-538 50 mg] co-packaged with dasabuvir [ABT-333 250 mg]. An indication for the treatment of genotype 1 (GT1) chronic hepatitis C infection (HCV) in patients with or without cirrhosis is being sought by the applicant. The duration of treatment and whether to add ribavirin (RBV) to the regimen is dependent on the patient population.

The results of six Phase 3 trials and one 14-arm Phase 2 study were submitted to demonstrate the efficacy and safety of the 3DAA regimen (Table 1.1). All the Phase 3 trials were randomized, multicenter trials; however, the primary efficacy comparison in all 6 trials is to the historical sustained virologic response (SVR) rate of telaprevir plus pegylated interferon (pegIFN) with RBV therapy based on predefined thresholds for non-inferiority and superiority. In all trials the results for the DAA+RBV and DAA arms were shown to be superior to the historical control based on the lower bound of the confidence interval about the SVR_{12} (SVR 12 weeks after stopping treatment) being larger than pre-defined thresholds.

Table 1.1 Phase 2 and Phase 3 studies of 3DAA

Study	Design	Treatment Period	Follow-up Period	Randomized Arms (ITT N)	Study Population
Phase 2 Study Section 3.1.2.1 of review					
M11-652	OL, R, MC	8, 12 and 24 weeks	48 weeks	14 arms (see Table 3.1.1)	GT1 Non-cirrhotic TN & TE null responders
Phase 3 Studies by Design					
Studies comparing 3DAA+RBV versus 3DAA+Placebo Section 3.1.3.1 of review					
M13-389 PEARL II	OL, R, MC	12 weeks	48 weeks	3DAA+RBV (88) 3DAA (91)	GT 1b non-cirrhotic TE
M13-961 PEARL III	DB, R, MC	12 weeks	48 weeks	3DAA+RBV (210) 3DAA (209)	GT 1b non-cirrhotic TN
M14-002 PEARL IV	DB, R, MC	12 weeks	48 weeks	3DAA+RBV (100) 3DAA (205)	GT 1a non-cirrhotic TN
Studies comparing 3DAA+RBV versus Placebo Section 3.1.3.2 of review					
M11-646 SAPPHIRE I	DB, R, MC	12 weeks	48 weeks	3DAA+RBV (473) Placebo (158)	GT 1 non-cirrhotic TN
M13-098 SAPPHIRE II	DB, R, MC	12 weeks	48 weeks	3DAA+RBV (297) Placebo (97)	GT 1 non-cirrhotic TE
Study comparing duration of treatment for 3DAA+RBV in cirrhotic patients Section 3.1.3.3 of review					
M13-099 TURQUISE II	OL, R, MC	12 weeks and 24 weeks	48 weeks	3DAA+RBV 12 wks (208) 24 wks (172)	GT 1 cirrhotic TE & TN

OL=open label, DB=double-blind, R=randomized, MC=multicenter, Trt=treatment, FU=follow-up off treatment, TE=treatment experienced, TN=treatment naïve

The randomized arms in the six Phase 3 clinical studies allow one to address questions regarding the need for the addition of RBV in a non-cirrhotic population (Studies M13-389, M13-961 and M14-002), regarding the safety of the 3DAA+RBV regimen (Studies M11-646 and M13-098) and regarding the duration of treatment for cirrhotic patients (Study M13-099).

[1] DAA's are small molecules that target the hepatitis C virus.

For the 1a population without cirrhosis, the results for Study M14-002 demonstrate a statistically significantly higher SVR_{12} rate (about 7% higher, Table 1.2) for the full regimen including RBV compared to the 3DAA alone. For the 1a population with cirrhosis, 24 weeks of treatment was seen to be more beneficial than 12 weeks of treatment and although the difference was not statistically different, the difference was seen across multiple subgroups and there were no serious safety issues recommending against longer treatment.

Table 1.2 SVR_{12} rates for the 1a population

M14-002 GT 1a w/o cirrhosis	**3DAA+RBV[1]** **97/100 97% (94%, 100%)**	3DAA 185/205 90% (86%, 94%)	Trt diff (95%CI) +6.8% (+1%, +12%)
M13-099 GT 1a w/ cirrhosis	**3DAA+RBV/24 weeks** **115/121 95% (97%, 100%)**	3DAA+RBV/12 weeks 124/140 89% (81%, 94%)	Trt diff (95%CI) +6% (-1%, +13%)

[1]The regimens in the labeling proposed by FDA are bolded.

By contrast, the 1b population required less aggressive treatment with non-cirrhotics adequately treated with 3DAA without RBV and with cirrhotics treated with 12 weeks of treatment. The SVR_{12} rates for the 1b population were very high at 99% to 100% (Table 1.3).

Table 1.3 SVR_{12} rates for the 1b population

M13-389 & M13-961 GT 1b w/o cirrhosis	3DAA+RBV 298/301 99% (97%, 100%)	**3DAA[1]** **304/304 100% (99%, 100%)**	Trt diff (95%CI) -1% (-2%, +1%)
M13-099 GT 1b w/ cirrhosis	3DAA+RBV/24 weeks 51/51 100% (92%, 100%)	**3DAA+RBV/12 weeks** **67/68 98.5% (91%, 100%)**	Trt diff (95%CI) +1% (-6%, +8%)

[1]The regimens in the labeling proposed by FDA are bolded.

Two safety issues were named as secondary efficacy endpoints: 1) percentage of patients with hemoglobin below lower limit of normal (LLN) for 3DAA+RBV versus 3DAA in Study M14-002 and 2) percentage of patients with ALT less than or equal to the ULN for 3DAA+RBV versus placebo in Studies M11-646 and M13-098. As would be expected from the effect of RBV, a significant decrease in hemoglobin is seen in the 3DAA+RBV arm compared to the 3DAA arm (Table 1.4); hemoglobin values over time are illustrated in Figure 3.1.3.1.3 on page 29. In the two placebo-controlled trials, ALT was normalized in about 98% of patients taking 3DAA+RBV (median decrease of about 47 U/L) compared to about 15% of placebo patients (Table 1.4); ALT values over time are illustrated in Figure 3.1.3.2.3 on page 36.

Table 1.4 Results for changes in hemoglobin and changes in ALT

Of patients w/baseline Hb>LLN, % w/Hb<LLN at end of treatment			
M14-002 GT 1a w/o cirrhosis	3DAA+RBV 52/95 55%	3DAA 16/191 8%	p<0.001
Of patients w/baseline ALT>ULN, % w/ALT≤ ULN at end of treatment			
M11-646 GT 1 w/o cirrhosis	3DAA+RBV 352/360 98%	Placebo 18/114 16%	p<0.001
M13-098 GT 1 w/o cirrhosis	3DAA+RBV 217/224 97%	Placebo 10/77 13%	p<0.001

Overall, the results submitted by AbbVie demonstrate the efficacy of the Viekara pak (3DAA) with or without RBV for the treatment of genotype 1 patients with or without cirrhosis. The SVR rates for the new drug product are larger than pre-specified superiority thresholds which were based on rates for the historical control of telaprevir plus pegIFN and RBV. The choices of regimens based on genotype and the presence or absence of cirrhosis are supported by the results of comparisons of randomized arms in six Phase 3 studies.

2 INTRODUCTION

2.1 Overview

The applicant (AbbVie) has submitted an NDA for a combination product of three direct acting agents [2] (DAA) to be marketed as Viekira Pak (called 3DAA in this review). This product is composed of a fixed dose tablet of ombitasvir [ABT-267 12.5 mg], paritaprevir [ABT-450 75 mg] and ritonavir [ABT-538 50 mg] co-packaged with dasabuvir [ABT-333 250 mg]. Ritonavir alone has no antiviral activity against HCV and is not a DAA but does increase the bioavailability of ABT-450 (from applicant's Clinical Overview). An indication for the treatment of genotype 1 (GT1) chronic hepatitis C infection (HCV) in patients with or without cirrhosis is being sought by the applicant. Genotype 1 is the most common HCV genotype worldwide with genotype 1a most common in North America and Western Europe and genotype 1b most common in Southern and Eastern Europe, Latin America and Asia.

The applicant claims that each DAA (ombitasvir, paritaprevir or dasabuvir) has potent antiviral activity alone against HCV GT1; however, treatment as monotherapy generally results in selection of resistance. "The 3-DAA combination of ABT-450, ABT-267, and ABT-333 inhibits selection of resistance and has demonstrated elimination of the virus in vitro" (page 6 of section 2.5.4 of the Clinical Overview). The 3DAA is recommended by the applicant for all GT1 populations but the duration of treatment and whether to add ribavirin (RBV) to the regimen may be dependent on the patient population.

This 3DAA regimen was granted Breakthrough Therapy Designation in May 2013 and the NDA was accepted for rolling review. The final submission component arrived on April 21, 2014 but was followed up by additional submissions including updated efficacy and safety data.

AbbVie submitted the results of six Phase 3 trials to demonstrate the efficacy and safety of their 3DAA regimen (Table 2.1.1). In addition, two Phase 2 studies were considered key studies by the applicant. One study, M11-652 was designed with 14 arms to identify a regimen for use in the Phase 3 trials and is reviewed in full in Section 3.1.2.1. The second Phase 2 study, M14-103, is a small single-arm open label study in HCV genotype 1-infected non-cirrhotic adults who were on a stable opioid replacement therapy with methadone or buprenorphine ± naloxone. The results for Study M14-103 are not reviewed here because it is a small single arm study with positive results (SVR_{12} rate of 97% (37/38) with CI of 92% to 100%) but no comparative analyses (the applicant planned cross study comparisons to M11-652 but these were not done for efficacy). In addition, 8 Phase 2 studies and 49 Phase 1 studies were part of the clinical development program; these early, small (generally single-armed) clinical trials are not reviewed here. Two additional trials (M14-004 and M12-999) are part of the development program but were not included in the complete submission dated 4/21/14. If these trials are submitted and there is sufficient data to warrant a statistical review, these two trials will be reviewed in an addendum to this review.

All the Phase 3 trials were randomized, multicenter trials (Table 2.1.1); however, the primary efficacy comparison in all 6 trials is to the historical sustained virologic response (SVR) rate of telaprevir plus pegylated interferon (pegIFN) with RBV therapy (details regarding the historical controls are provided in Section 2.3 of this review). The applicant explained in their submission that enrolling an active comparator arm for demonstrating efficacy is infeasible because physicians generally recommend delaying treatment due to toxicity associated with the protease inhibitor/pegIFN/RBV regimens approved at the time of the design of these trials. In addition, the applicant states that an active comparator which contains pegIFN cannot be effectively blinded due to the high rate of adverse events. The randomized arms do allow one to address questions regarding the need for the addition of RBV in a non-cirrhotic

[2] DAA's are small molecules that target the hepatitis C virus.

population (Studies M13-389, M13-961 and M14-002,), regarding the safety of the 3DAA+RBV regimen (Studies M11-646 and M13-098) and regarding the duration of treatment for cirrhotic patients (Study M13-099). The emphasis of this review is primarily on the questions addressed by the randomized comparisons.

Table 2.1.1 Phase 2 and Phase 3 studies of 3DAA identified as key

Study	Design	Treatment Period	Follow-up Period	Randomized Arms (ITT N)	Study Population
Phase 2 Study Section 3.1.2.1 of review					
M11-652	OL, R, MC	8, 12 and 24 weeks	48 weeks	14 arms (see Table 3.1.1)	GT1 Non-cirrhotic TN & TE null responders
Phase 3 Studies by Design					
Studies comparing 3DAA+RBV versus 3DAA+Placebo Section 3.1.3.1 of review					
M13-389 PEARL II	OL, R, MC	12 weeks	48 weeks	3DAA+RBV (88) 3DAA (91)	GT 1b non-cirrhotic TE
M13-961 PEARL III	DB, R, MC	12 weeks	48 weeks	3DAA+RBV (210) 3DAA (209)	GT 1b non-cirrhotic TN
M14-002 PEARL IV	DB, R, MC	12 weeks	48 weeks	3DAA+RBV (100) 3DAA (205)	GT 1a non-cirrhotic TN
Studies comparing 3DAA+RBV versus Placebo Section 3.1.3.2 of review					
M11-646 SAPPHIRE I	DB, R, MC	12 weeks	48 weeks	3DAA+RBV (473) Placebo (158)	GT 1 non-cirrhotic TN
M13-098 SAPPHIRE II	DB, R, MC	12 weeks	48 weeks	3DAA+RBV (297) Placebo (97)	GT 1 non-cirrhotic TE
Study comparing duration of treatment for 3DAA+RBV in cirrhotic patients Section 3.1.3.3 of review					
M13-099 TURQUISE II	OL, R, MC	12 weeks and 24 weeks	48 weeks	3DAA+RBV 12 wks (208) 24 wks (172)	GT 1 cirrhotic TE & TN

OL=open label, DB=double-blind, R=randomized, MC=multicenter, Trt=treatment, FU=follow-up off treatment, TE=treatment experienced, TN=treatment naïve, ITT=intent-to-treat population of patients who were randomized and took at least one dose

The statistical evaluation section of this review (Section 3) is organized as shown in the table above with sections for the Phase 2 study and by each type of design for the Phase 3 trials.

2.2 Data Sources and Quality

The data was provided as tabulation files and as analysis files. The formats of the datasets and the accompanying define files were acceptable for statistical review and no changes were requested by this reviewer. The full application may be accessed in Global Submit at the following link: \\CDSESUB1\evsprod\NDA206619\206619.enx .

Because the applicant was granted a rolling application, submissions were accepted prior to the Final Drug Application which was received April 21, 2014. The reports and data for Studies M11-652, M11-646 and M13-098 were submitted under SN 2 SDN 3 dated March 18, 2014. The SVR_{12} data in four studies (M13-099, M13-961, M13-389 and M11-646) was updated with a submission on 7/1/2014; these data were included in the statistical results presented here.

All tables, figures and results shown in this review were created by the reviewer unless otherwise noted.

2.3 Historical Data for the Treatment of HCV

Primary efficacy is demonstrated in the Phase 3 trials of this submission based on comparisons of the test regimen (3DAA+RBV or 3DAA alone) to an historical control of telaprevir plus pegIFN and RBV. Based on historical data, thresholds for demonstrating non-inferiority (NI) and superiority (SUP) were named in the Phase 3 protocols. To demonstrate efficacy, a 95% confidence interval was constructed for the primary endpoint results (SVR_{12} rate) and the lower bound was compared against the appropriate threshold. Exceeding the thresholds is the basis for approval but does not allow for comparative claims.

Thresholds were specific for patient populations defined by HCV genotype (1, 1a, or 1b), treatment experience history (naïve [TN] or prior HCV treatment [TE]) and by presence or not of diagnosed cirrhosis. The thresholds defined in the protocols are summarized in the table below:

Thresholds	GT1 TE	GT1 TN	GT1a TE	GT1a TN	GT1b TE	GT1b TN
Non-cirrhotic						
NI	60%	70%	55%	65%	~66%	73%
SUP	70%	**80%**	**65%**	**75%**	**~76%**	**84%**
w/ cirrhosis						
NI	43%	56%	NC	NC	NC	NC
SUP	**54%**	**67%**				

Some approximate values are shown since the precise values named in the protocols with GT1b TE patients were based on estimates weighted by the expected balance in the specific trial population. NC=not computed

The thresholds were computed by the applicant from the results of trials REALIZE, ILLUMINATE and ADVANCE whose results are presented in the following three tables.

Table 2.3.1 Historical SVR rates for non-cirrhotic TE patients

REALIZE	GT1a TE	GT1b TE
Relapsers	119/142 (84%)	123/140 (88%)
Partial responders	26/55 (47%)	27/40 (68%)
Null responders	24/88 (27%)	22/59 (37%)
Weighted estimate (95% CI)	59% (53%, ~**65%**)	71% (64%, ~**77%**)

[1]Based on the labeling for INCIVEK and the publication: Zeuzem S et al. REALIZE trial final results: telaprevir-based regimen for GT1 hepatitis c virus infection in patients with prior null response, partial response or relapse to peginterferon/ribavirin. J Hepatol. 2011;54 Suppl:S3.

Table 2.3.2 Historical SVR rates for non-cirrhotic TN patients

	GT1 TN	GT1a TN	GT1b TN
ADVANCE[1]	270/342 (79%)	162/217 (75%)	119/142 (84%)
ILLUMINATE	367/479 (77%)	273/388 (70%)	112/149 (75%)
Meta-analysis estimate (95% CI)	78% (75%, **80%**)	72% (68%, **75%**)	80% (75%, **84%**)

[1]Based on the teleprevir labeling.

Table 2.3.3 Historical SVR rates for cirrhotic patients

	GT1
REALIZE TE pts	
Relapsers	48/57 (84%)
Partial responders	11/32 (34%)
Null responders	7/50 (14%)
Weighted estimate (95% CI)	47% (41%, **54%**)
ADVANCE TN pts	15/21 (71%)
ILLUMINATE	31/61 (51%)
Meta-analysis estimate (95% CI)	56% (45%, **67%**)

Superiority thresholds for the Phase 3 trials were based on the upper bound for the confidence interval about the combined historical estimate of the SVR rate for the specific population being studied; these

values are bolded in the above three tables. The estimates were weighted based on the projected population balance of sub-genotype or treatment experience history or based on a meta-analysis. For assessment of non-inferiority to the historical SVR rate, a margin of 10.5% was used and the threshold for non-inferiority was computed by subtracting 10.5 from the superiority threshold.

At the IND stage, the thresholds presented here were agreed upon by both the clinical and statistical FDA staff. This reviewer has summarized the thresholds but not researched the origins of these thresholds. It should be noted that the superiority thresholds are generally about 10% to 20% less than the lower bounds of the confidence intervals of the SVR_{12} estimates observed in the Phase 3 studies of this application so there are no borderline results that would be impacted by small changes in the thresholds based on the historical control of telaprevir plus pegIFN and RBV.

Since the onset of the trials in this application, a new drug (Sovaldi) has been approved for the treatment of HCV. For GT1 patients, Sovaldi is approved in combination with peg interferon alfa and ribavirin for 12 weeks of treatment. GT1 patients ineligible for interferon treatment may be treated with Sovaldi and ribavirin for 24 weeks. One single-arm study (NEUTRINO Study 110) was conducted in GT1 TN patients; the results are summarized below.

Sovaldi NEUTRINO Study 110 SVR_{12} rates[1]

	GT 1a TN SVR_{12} (95%CI)	GT 1b TN SVR_{12} (95%CI)
Non-Cirrhotics	168/180 93% (89%, 97%)	47/56 84% (72%, 92%)
Cirrhotics	36/43 84% (69%, 93%)	6/9 67% (30%, 93%)

[1]Rates were extracted from page 100 of the FDA statistical review; Clopper-Pearson exact confidence intervals were computed by this reviewer.

If we apply the same methodologies applied to the active control data used to compute the thresholds for the Phase 3 trials in this application to the Sovaldi data, we would obtain the following superiority thresholds based on the upper bounds of the confidence intervals:

Non-cirrhotics 1a TN 97% 1b TN 92%
Cirrhotics 1a TN 93% 1b TN 93%

Non-inferiority thresholds most likely would not be based on just subtracting 10.5% from these superiority thresholds since the 10.5% is based on prior telaprevir data not on Sovaldi data. These comparisons to Sovaldi are not appropriate when considering the approval of the product reviewed here. Although, given Sovaldi may presently (August 2014) be the best available treatment for HCV GT1 patients and the attention given to comparative efficacy across products, it may be important to consider the 3DAA+RBV results in the context of the Sovaldi results.

It should be noted in this discussion of historical controls that comparisons to historical data may be impaired by the biased selection of patients given the controls are not randomized or concurrent. Important differences in patient populations may not be accounted for in the analysis. Nevertheless as described earlier in the review, historical controls allow one to conduct a trial in populations difficult to study affording the possibility of improved treatment opportunities as is the case for this application.

3 STATISTICAL EVALUATION

3.1 Evaluation of Efficacy

3.1.1 Efficacy Endpoints and Statistical Methods Common to All Studies Reviewed

Efficacy Endpoints

For the Phase 2 trial and all Phase 3 trials reviewed here, the primary outcome is the sustained virologic response (SVR). SVR is based on the plasma HCV RNA levels measured at a central laboratory using the Roche COBAS TaqMan® real-time reverse transcriptase-PCR assay v2.0. The lower limit of detection (LLOD) is 15 IU/mL and lower limit of quantification (LLOQ) for the Roche assay is 25 IU/mL. Assays of HCV RNA were performed at baseline, at Weeks 1-4, 6, 8 and then monthly until the end of treatment and at post treatment weeks 2, 4, 8, 12, 24 and 48. Baseline HCV RNA was defined as the last non-missing measurement collected before the first dose of DAA study drug.

SVR is the sustained virologic response measured as the percentage of patients with HCV RNA<LLOQ at a specific timepoint after the last dose, without any confirmed quantifiable (≥ LLOQ) post-treatment value before or during that SVR window. In the Phase 2 trial, the primary outcome was SVR at post-treatment Week 24 while in all the Phase 3 trials the primary outcome was at post-treatment Week 12. SVR was also measured at other timepoints for several of the studies.

In addition to SVR, the applicant summarized results for other outcomes listed here:

- RVR: rapid virologic response measured as HCV RNA < LLOQ at Week 4
- EVR: early virologic response measured as HCV RNA < LLOQ at Week 8 for 8-week treatment or at Week 12 for 12-week treatment
- EOTR: end of treatment response HCV RNA< LLOQ at the end of 8, 12, or 24 weeks of treatment
- Relapse: two consecutive HCV RNA values ≥ LLOQ between the end of final treatment visit and end of the post-treatment period amongst subjects who took at least 11 weeks of treatment with a 12-week treatment regimen with HCV RNA < LLOQ at the Final Treatment Visit
- Rebound: confirmed increase from nadir in HCV RNA > 1 log10 IU/ml at any time point or confirmed HCV RNA > LLOQ (defined as two consecutive HCV RNA measurements ≥ LLOQ) at any point after HCV RNA > LLOQ during treatment.
- Virologic failure:
    - Failure to achieve an HCV RNA decrease of at least 2 log10 IU/mL at Week 1
    - Confirmed increase from nadir (defined as 2 consecutive HCV RNA measurements > 1 log10 IU/mL above nadir) in HCV at any time point (breakthrough)
    - Failure to achieve HCV RNA < 25 IU/mL (detected or not detected) at Week 6
    - Confirmed HCV RNA > LLOQ (defined as two consecutive HCV RNA measurements > LLOQ) at any point after HCV RNA < LLOQ

For the Phase 3 studies, this reviewer focused on SVR_{12} (SVR at post-treatment Week 12) as well as the reasons patients did not achieve an SVR (on treatment virologic failure (VF), relapse, and lost-to-follow-up). For the review, relapse is defined as a patient who is treated, responds with an HCV RNA value less than LLOQ and then subsequently, off treatment, has an HCV RNA value above LLOQ.

VF is defined above as for Study M11-652; for the Phase 3 studies, the first criterion of failure to achieve an HCV RNA decrease of at least 2 log10 IU/mL at Week 1 was not part of the VF definition.

Patients who met any of the above virologic failure criteria must be discontinued from treatment and could be offered pegIFN and RBV or other medications by the investigator.

Statistical Methods

The primary analysis population in all trials was an intent-to-treat population of all patients randomized who received at least one dose of DAA.

The applicant's imputation methods used for HCV RNA missing values and SVR missing values is summarized in the two tables below created by the reviewer. If more than one HCV RNA value is recorded within a window, the assessment closest to the nominal time was used. If HCV RNA was missing at the last post-treatment visit, the applicant carried forward a value from the previous visit (this only occurred for 4 patients in the entire Phase 3 database). If there are two observations equally distant to the nominal time, the latest one will be used in analyses. For SVR values, the last value in the window was always used.

Table 3.1.1.1 Flanking Imputation of missing HCV RNA values

Value before missing value	**Value after missing value**	**Imputed value for missing data**
Non-missing & Closest Value	Non-missing & Not Closest Value	Value before missing value
Non-missing & Not Closest Value	Non-missing & Closest Value	Value after missing value
Undetectable HCV RNA	Undetectable HCV RNA	Undetectable HCV RNA
Unquantifiable HCV RNA	Unquantifiable HCV RNA	Unquantifiable HCV RNA
Unquantifiable HCV RNA	Undetectable HCV RNA	Unquantifiable HCV RNA
Undetectable HCV RNA	Unquantifiable HCV RNA	Unquantifiable HCV RNA

Table 3.1.1.2 Flanking Imputation of missing SVR values

Status before missing value	**Status after missing value**	**Imputed value for missing data**
NA	HCV RNA value available	Use HCV RNA from window after missing value
NA	No value	Non-responder
Last available considered a non-responder	NA	All timepoints after the non-response will be considered non-responses (failures)
Patient discontinues treatment and starts rescue	NA	Non-responder

Similar statistical methodologies were used for all studies. Where more details are required for a specific study, those details are included with the efficacy results of that specific study.

For the primary efficacy analysis comparing a regimen (3DAA+RBV or 3DAA) to the pre-defined historical control, the applicant computed the percentage of patients with SVR_{12} and a 2-sided 95% CI using a simple proportion and variance along with the normal approximation to the binomial distribution.

When the rate was 0% or 100%, the applicant used Wilson's method to compute a confidence interval. The lower bound of the 95% CI was then compared to the predefined threshold for non-inferiority. If non-inferiority was shown, then a comparison to the superiority threshold was made. This reviewer used a more conservative approach by computing the confidence interval using the Clopper-Pearson exact method which usually results in a wider interval with small samples. The review contains CI's computed by the reviewer unless otherwise noted.

For the comparison of randomized arms, this reviewer performed a Mantel-Haenszel stratified analysis with the stratifiers used for randomization. The applicant used a stratum weighted estimate for the difference if a test of homogeneity showed a lack of homogeneity across strata.

To test for the impact of baseline characteristics on outcome, the applicant performed logistic regression analyses to determine which subgroups might be prognostic. This reviewer reviewed these results and considered them along with the protocol-defined subgroups when selecting which subgroup analyses to present. The results of the applicant's logistic regression analyses are not provided here.

To test for the homogeneity of the treatment difference of randomized arms across subgroups, the applicant used a Breslow-Day test while this reviewer used Zelen's exact test of homogeneity to test for an interaction.

[This space purposely left blank.]

3.1.2 Phase 2 Studies

3.1.2.1 Study M11-652

(conducted 10/2011 to 9/2013)

Design

Study M11-652 is an open label, randomized, 14-arm, Phase 2 study designed to examine multiple doses of ABT-450/r in combination with ABT-267 and /or ABT-333 and with or without RBV for treatment durations of 8, 12 or 24 weeks in both treatment-naïve subjects and previous null responders to pegIFN and RBV. All patients were non-cirrhotic. The design of the trial is illustrated in Table 3.1.2.1.1 below. Randomization was stratified by IL28B genotype (CC versus non-CC) and HCV subtype (1a versus non-1a). All study drugs were to be given with food. All patients receiving at least one dose were to be followed for an additional 48 weeks after their last dose.

Treatment-naïve subjects (Cohort 1 Arms A-I) should not have received previous anti-viral treatment for hepatitis C infection. Prior null responders (Cohort 2 Arms J-N) should have previously received pegIFN and RBV for at least 12 weeks and have failed to achieve a 2 log10 HCV RNA decrease at Week 12.

Table 3.1.2.1.1 Phase 2 Study M11-652 treatment arms[1]

Group Duration	ABT-450/r 150/100 QD	ABT-267 25 QD	ABT-333 250 BID	RBV Wt-based
Cohort 1 Treatment Naive				
A 8 wk	**X**	**X**	**X**	**X**
B 12 wk	**X**		**X**	**X**
C 12 wk	**100/100**	**X**		**X**
D 12 wk	**200/100**	**X**		**X**
E 12 wk	**X**	**X**	**X**	
F 12 wk	**100/100**	**X**	**X**	**X**
G 12 wk	**X**	**X**	**X**	**X**
H 24 wk	**100/100**	**X**	**X**	**X**
I 24 wk	**X**	**X**	**X**	**X**
Cohort 2 Treatment Experienced				
J 12 wk	**200/100**	**X**		**X**
K 12 wk	**100/100**	**X**	**X**	**X**
L 12 wk	**X**	**X**	**X**	**X**
M 24 wk	**100/100**	**X**	**X**	**X**
N 24 wk	**X**	**X**	**X**	**X**

[1]The large bolded X indicates that the drug and dose labeling the column was used in the randomized arm.

Entry criteria for this trial included evidence of chronic HCV genotype (GT) 1 (including interleukin 28B genotypes [CC, CT and TT] and HCV genotypes 1a and 1b), aged 18 to 70, using birth control (excluding oral contraceptives and ethinyl estradiol), 18>BMI>38 kg/m^2, no HepB or HIV, HbA1c<8%, no cirrhosis, and no abnormal labs.

This study was originally planned as an adaptive trial with statistical methods proposed to test for ineffective treatment arms and then dropping arms that met the statistical criteria. If the posterior probability that the SVR_{12} would be greater than 75% was less than 10%, an arm would be dropped. The latter did not occur in this trial so no arms were dropped.

Adherence to study drug was assessed by the Medication Event Monitoring systems (MEMS™, AARDEX Group Ltd., Switzerland) at each visit.

Plasma HCV RNA levels were measured for the primary endpoint at a central laboratory using a Roche assay with a lower limit of detection (LLOD) of 15 IU/mL and lower limit of quantification (LLOQ) of 25 IU/mL. Assays were performed at baseline, at Weeks 1-4, 6, 8 and then monthly until the end of treatment. HCV RNA was measured at post treatment weeks 2, 4, 8, 12, 24 and 48. The primary endpoint was the sustained virologic response at post-treatment Week 24 (SVR_{24}) defined as a value lower than the LLOQ (see Section 3.1.1 of this review for more details) with imputation for missing values and with no confirmed rebound/virologic failure or post-treatment relapse.

The primary efficacy objective of the study was to show a significantly higher percentage of patients with SVR_{24} for patients treated with 3 DAAs (ABT-450/r 150/100 mg+ABT-267+ABT-333) and RBV for 12 weeks than for patients treated with the same regimen for 8 weeks (**Arm G** [12 week regimen] versus **Arm A** [8 week regimen]). The primary comparison was powered at 80% assuming a rate of SVR_{24} of 66% in Arm A and 90% in Arm G with 80 subjects in Arm A and 40 subjects in Arm G based on Fisher's exact test with a two-sided significance level of 0.05. So a higher response rate was expected for the 12-week regimen than the 8-week regimen.

For secondary comparisons of the other arms in the study, the applicant planned (b) (4) This reviewer does not agree with the applicant's methods and will not present their secondary results. (b) (4)

This reviewer focused on comparisons between randomized arms where a specific hypothesis could be tested. The applicant also performed analyses comparing single arms and this reviewer checked those results against the results presented here. There were no notable differences in conclusions between the applicant and reviewer.

In addition to SVR_{24}, several secondary endpoints were analyzed by the applicant. This reviewer focused primarily on SVR_{24} as well as the reasons patients did not achieve an SVR (on treatment virologic failure, relapse, and lost-to-follow-up).

Patient Disposition

A total of 580 patients were randomized in 92 sites in the United States, Puerto Rico, Canada, France, Germany, Spain, United Kingdom, Australia and New Zealand. About 69% of the randomized patients were enrolled in the USA. Sites ranged in size from 1 to 20 patients with the largest site in the USA. The USA sites began enrollment about 3-4 months before any of the sites in other countries.

Patients were randomized to Arms A, F and G first in a 2:1:1 ratio. When 50% of patients were randomized into those three arms, randomization into the other Cohort 1 arms began. The randomization ratio, then, was 2:2:2:2:4:1:1:2:2 for Arms A to I with 80 patients in the 8 week DAA arm (Arm A) and in the 12 week DAA arm (Arm E), about 40 patients in Arms B to J and at least 20 in Arms K-N of Cohort 2 as planned. In 8 arms, all randomized patients had at least one dose and were counted as part of the ITT population (Table 3.1.2.1.1). In the other 6 arms, a total of 9 patients were never treated. In an open label study non-treatment of randomized patients can be a concern due to the possibility of introducing selection bias; however, this reviewer is not concerned given the small number of patients (only 1.5%) that were not treated coupled with the very high responder rates. These few patients will not impact the results.

In most arms, 90% or more of the patients completed the study. Among naïve patients, Arms H and I (24 week treatment arms) had more patients discontinuing treatment (13% and 18%, respectively) than what was seen in other arms; the primary reason was either non-compliance or patient refused to continue taking drug and 10 out of the 12 patients discontinued drug after 12 or more weeks on study. This pattern was not seen for the 24-week arms (M and N) of treatment experienced patients.

Table 3.1.2.1.2 Phase 2 Study M11-652 Patient Disposition[1]
Cohort 1 (Arms A-I, treatment naïve patients) rows have a clear background while
Cohort 2 (Arms J-N, treatment experienced null responders) rows have a grey background

Group	Randomized	Randomized & Treated (ITT)	Discontinued All Study Drugs[2]	Discontinued Study[3]	Completed Study
A	80 (100%)	80 (100%)	1 (1%)	3 (4%)	77 (96%)
B	43 (100%)	41 (95%)	3 (7%)	7 (16%)	36 (84%)
C	39 (100%)	39 (100%)	0	0	39 (100%)
D	40 (100%)	40 (100%)	2 (5%)	4 (10%)	36 (90%)
E	80 (100%)	79 (99%)	2 (3%)	8 (10%)	72 (90%)
F	39 (100%)	39 (100%)	2 (5%)	1 (3%)	38 (97%)
G	41 (100%)	40 (98%)	3 (7%)	4 (10%)	37 (90%)
H	40 (100%)	40 (100%)	5 (13%)	3 (7%)	37 (93%)
I	40 (100%)	40 (100%)	7 (18%)	3 (7%)	37 (93%)
J	47 (100%)	45 (96%)	1 (2%)	3 (6%)	44 (94%)
K	23 (100%)	23 (100%)	2 (9%)	2 (9%)	21 (91%)
L	23 (100%)	22 (96%)	0	2 (9%)	21 (91%)
M	23 (100%)	23 (100%)	2 (9%)	2 (9%)	21 (91%)
N	22 (100%)	20 (91%)	0	3 (14%)	19 (86%)

1 % computed by reviewer based on number randomized
2 Patients who were treated at least once and then discontinued all study drugs early
3 Includes patients never treated as well as patients treated and then discontinued

In Cohort 1, 6 patients (~1%) discontinued all study drugs due an adverse event (ADE) and no patients discontinued from the study due to an adverse event. In Cohort 2, 2 patients (~1%) discontinued all study drugs due an adverse event and 1 patient discontinued from the study due to an adverse event. No particular regimen had a notable ADE rate.

Baseline Demographics

The ITT population consisted of about 84% whites, 55% males, and 66% genotype 1a. Ages ranged from 19 to 70 years with an average of about 50. The groups were generally fairly well-balanced (Table 3.1.2.1.3). Cohort 2 had fewer patients enrolling from the US than Cohort 1. More than 85% of patients had a baseline HCV RNA greater than 800,000. Demographics for Arms A and G (the ones being compared for the primary comparison look) look similar. The most common previous HCV treatments for patients in arms J to N were peg-interferon alfa 2A (60%) and ribavirin (94%).

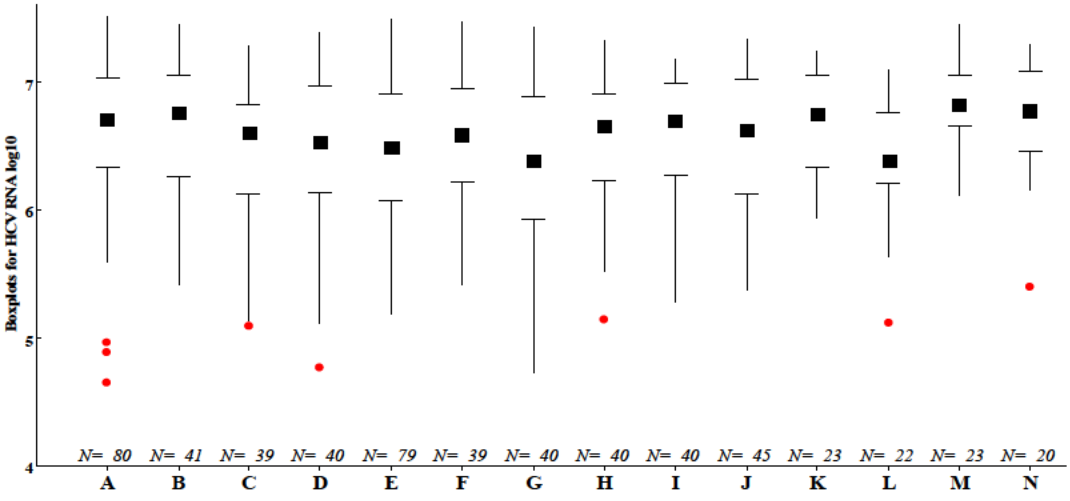
Table 3.1.2.1.3 Phase 2 Study M11-652 Key Patient Baseline Demographics ITT population
Cohort 1 (Arms A-I, treatment naïve patients) rows have a clear background while
Cohort 2 (Arms J-N, treatment experienced null responders) rows have a grey background

ARM	% USA	% Female	%White/ %Black	Mean age/ % ≥ 65	% Drinker	%GT1a/ %GT1b	IL28B GT %CC/CT/TT	HCV RNA (Log_{10}) Med
A (n=80)	75%	43%	85%/11%	50/4%	26%	70%/30%	28/51/21%	6.7
B (n=41)	90%	56%	85%/12%	51/2%	24%	71%/29%	34/44/22%	6.8
C (n=39)	72%	36%	80%/21%	51/0%	54%	67%/33%	26/56/18%	6.5
D (n=40)	65%	50%	75%/20%	49/3%	25%	65%/35%	28/57/15%	6.6
E (n=79)	70%	43%	81%/17%	48/3%	24%	66%/32%	29/49/22%	6.5
F (n=39)	85%	49%	77%/21%	49/5%	28%	69%/31%	28/49/23%	6.5
G (n=40)	95%	40%	83%/13%	51/3%	25%	68%/32%	28/50/22%	6.6
H (n=40)	68%	55%	95%/5%	51/10%	26%	68%/32%	28/62/10%	6.4
I (n=40)	68%	60%	93%/5%	52/10%	30%	68%/30%	28/52/20%	6.7
J (n=45)	49%	40%	78%/20%	51/9%	18%	58%/42%	2/62/36%	6.6
K (n=23)	43%	30%	83%/17%	49/9%	35%	65%/35%	4/48/48%	6.8
L (n=22)	36%	46%	86%/9%	51/9%	32%	59%/41%	5/77/18%	6.4
M (n=23)	43%	35%	91%/0%	51/9%	26%	61%/39%	4/61/35%	6.8
N (n=20)	60%	40%	85%/15%	55/20%	30%	65%/35%	0/55/45%	6.8

Source: Applicant's study report and reviewer results

The distributions of baseline values for HCV RNA (Figure 3.1.2.1.1) are similar across the 14 arms.

Figure 3.1.2.1.1 Boxplots of Baseline HCV RNA log10 by arm for the ITT population



Statistical Methods

See Section 3.1.1 of this review for a general description of the statistical methods used by the applicant and this reviewer. Methods specific to this study are described here.

SVR_{24} was the primary endpoint but additional timepoints (post treatment Weeks 4, 8 and 12) were also analyzed. SVR_{48} was originally in the protocol but then removed.

According to the original statistical analysis plan (SAP), the primary analysis to compare Arm A to Arm G will use a logistic regression model with factors for treatment group, baseline log10 HCV RNA level, HCV subgenotype (1a or 1b), geographic region and IL28B genotype (CC, non-CC). The SAP was amended to also include a stratum-adjusted Mantel-Haenszel (MH) method for efficacy analyses of RVR, EOTR, and all SVR endpoints with stratification by IL28B genotype (CC versus non-CC) and HCV subtype (1a versus non-1a). The latter method is the one used by this reviewer. An alpha level of 5% was planned with no adjustment for multiple comparisons. This reviewer thinks that an unadjusted alpha is acceptable because this is a Phase 2 trial designed to explore regimens that will be tested in a Phase 3 confirmatory trial so the results from this trial will not be used to establish the efficacy of the new drug regimen.

[This space purposely left blank.]

Efficacy Results

The primary endpoint (SVR_{24}) rates as well as the SVR_{12} rates are summarized in Table 3.1.2.1.4 for all 14 arms. For treatment naïve patients (Cohort 1), the lowest rate is seen for Arm B (arm excluding ABT-267) with an SVR_{24} of 83% and the highest rate is seen for Arm F (arm with 100 dose for ABT-450) at 97%. The proposed regimen for approval and the regimen used in the Phase 3 trial is Arm G where a rate of 95% was observed. For treatment-experienced null-responder patients, the lowest rate is seen for Arm J (arm excluding ABT-333 from the regimen) with an SVR_{24} of 89% and the highest rate is seen for Arm F (arm with full regimen for 24 weeks) at 100%. The full regimen for 12 weeks (Arm L) produced a rate of 95.5%. Overall the SVR_{24} rates are high for all regimens ranging from 83% to 100%.

Table 3.1.2.1.4 Study M11-652 SVR rates by arm; the regimens proposed for approval are bolded

Arm	ABT-450/r 150/100 QD	ABT-267 25 QD	ABT-333 250 BID	RBV	SVR_{12}	SVR_{24}
TRT NAIVE						
A 8 wk	X	X	X	X	71/80 (89%)	70/80 (87.5%)
B 12 wk	X		X	X	35/41 (85%)	34/41 (83%)
C 12 wk	100/100	X		X	35/39 (90%)	33/39 (85%)
D 12 wk	200/100	X		X	37/40 (92.5%)	37/40 (92.5%)
E 12 wk	X	X	X		72/79 (91%)	70/79 (89%)
F 12 wk	100/100	X	X	X	39/39 (100%)	38/39 (97%)
G 12 wk	**X**	**X**	**X**	**X**	**39/40 (97.5%)**	**38/40 (95%)**
H 24 wk	100/100	X	X	X	37/40 (92.5%)	37/40 (92.5%)
I 24 wk	X	X	X	X	37/40 (92.5%)	36/40 (90%)
TRT EXP						
J 12 wk	200/100	X		X	40/45 (89%)	40/45 (89%)
K 12 wk	100/100	X	X	X	21/23 (91%)	21/23 (91%)
L 12 wk	**X**	**X**	**X**	**X**	**21/22 (95.5%)**	**21/22 (95.5%)**
M 24 wk	100/100	X	X	X	22/23 (96%)	21/23 (91%)
N 24 wk	X	X	X	X	20/20 (100%)	20/20 (100%)

The primary comparison in this trial is the comparison of Arms A and G for naïve patients with the goal of determining whether more efficacy is observed with 12 weeks of treatment versus 8 weeks. The results for the primary comparison suggest that the longer treatment period is advantageous over the shorter one, based on the confidence interval for the difference in SVR_{24}, although the difference is not statistically significant (Table 3.1.2.1.5). The 12 week regimen looks better than the 8 week regimen for 1a patients (93% vs 84%) but minimally better for 1b patients (100% vs 96%) where both rates are high. For the IL28B genotypes, the confidence intervals are wide and the results are inconclusive for treatment of 12 weeks over 8 weeks. These results do not suggest any heterogeneity among subgroups and do suggest that 12 weeks of treatment is appropriate for all naïve patients.

Table 3.1.2.1.5 Study M11-652 SVR_{24} results for primary comparison of 12 weeks of treatment (Arm G) to 8 weeks of treatment (Arm A) in naïve patients

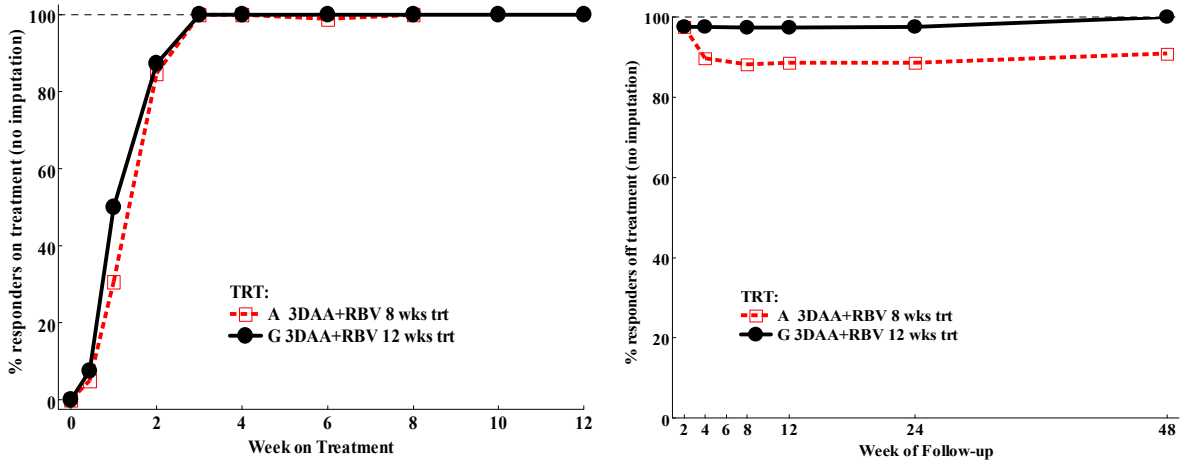
Pt Population	Trt diff G-A (95% CI)	Interpretation
All naive	+7.3% (-4%, +19%)[1]	NS, but favorable to 12 wks
1a	+9% (-6%, +24%)	NS, but longer treatment most favorable for 1a patients
1b	+4% (-13%, +20%)	
CC	+8% (-12%, +29%)	
CT	+5% (-11%, +21%)	
TT	+8% (-24%, +40%)	

NS=Not statistically significant with p>0.05; positive differences favor 12 weeks over 8 weeks of treatment

[1] Based on MH test stratifying on IL28B (CC, non-CC) and HCV genotype (1a, non-1a)

This reviewer plotted the observed percent of patients with Log10 HCV RNA ≤ 1.39794 (LLOQ Log10) by week on study. Because there is minimal missing data in either arm, the results for the observed data are very close to the results computed for the primary comparison with about a 7% treatment difference in favor of the 12 week regimen at the end of the study (graph to the right below). The graphs illustrate that all patients have responded by about Week 3 in both arms and that a difference between the arms is seen by post-treatment Week 4.

Figure 3.1.2.1.2 Percent of patients with a HCV RNA ≤ 1.39794 (LLOQ Log10) at each week for Arms A and G using observed data with no imputation for missing values. On treatment response are shown in the graph on the left and off treatment follow-up responses shown in the graph to the right



In addition to testing the primary hypothesis, the arms studied in naïve patients allow testing of several additional hypotheses and they are listed in Table 3.1.2.1.6 below. Although none of the differences are statistically significant (the trial is not adequately powered for this), the results suggest that the addition of ABT-267, ABT-333 and RBV, each individually, contribute to an increase in SVR rates. Also no advantage is seen for a longer treatment of 24 weeks or a dose of 150 for ABT-450 versus 100.

Table 3.1.2.1.6 Study M11-652 SVR_{24} results for primary comparison in all GT1 naïve patients

Test hypothesis	Arms	Treatment difference (95% CI)	Interpretation
Contribution of ABT-267	G-B	+12% (-3%, +26%)	NS, but favorable to adding ABT-267
Contribution of ABT-333	F-C	+13% (-1%, +28%)	Borderline significant, favorable to adding ABT-333
Contribution of RBV	G-E	+7% (-5%, +18%)	NS, slightly favorable to adding RBV
ABT-450/r dose 150/100 versus 100/100 (12 wk)	G-F	-2% (-14%, +6%)	NS
Trt for 24 wks vs. 12 wks	I-G	-5% (-18%, +8%)	NS

Based on MH test stratifying on IL28B (CC, non-CC) and HCV subgenotype (1a, non-1a)
NS=Not statistically significant

Results of subgroup analyses by GT1a/1b to test the contribution of ABT-267, ABT-333 and RBV (Table 3.1.2.1.7) show that each component appears to make a contribution to the efficacy of the 3DAA+RBV regimen for the 1a naïve population while, in the 1b naive population, rates are 100% regardless of regimen. Looking by IL28B genotype for the 1a population suggests the greatest advantage for the full regimen is seen for the CC genotype. However, there is no statistical heterogeneity across the subgroups so the results do not suggest more benefit in a specific subgroup over another.

Table 3.1.2.1.7 Study M11-652 SVR_{24} results by GT 1a and 1b and IL28B genotype for naïve patients

Test hypothesis	Arms	Treatment difference (95% CI)
Results for 1a naïve population		
Contribution of ABT-267	G-B	
All	25/27 (93%) - 22/29 (76%)	+16% (-3%, +36%)
IL28B CC	7/7 (100%) - 8/11 (73%)	+27% (+1%, +54%) **
IL28B CT	14/15 (93%) - 9/12 (75%)	+18% (-9%, +46%)
IL28B TT	4/5 (80%) - 5/6 (83%)	-3% (-49%, +42%)
Contribution of ABT-333	F-C	
All	26/27 (96%) – 20/26 (77%)	+19% (+0.3%, +39%)**
IL28B CC	8/8 (100%) – 7/8 (88%)	+12% (-10%, +35%)
IL28B CT	12/13 (92%) – 11/16 (69%)	+24% (-3%, +51%)
IL28B TT	6/6 (100%) – 2/2 (100%)	0% (-35%, +35%)
Contribution of RBV	G-E	
All	25/27 (93%) – 43/52 (83%)	+10% (-5%, +26%)
IL28B CC	7/7 (100%) – 14/16 (88%)	+12% (-4%, +29%)
IL28B CT	14/15 (93%) – 23/26 (88%)	+5% (-13%, +22%)
IL28B TT	4/5 (80%) – 6/10 (60%)	+20% (-26%, +66%)
Results for 1b naïve population		
Contribution of ABT-267	G-B 13/13 (100%) - 12/12 (100%)	0% (-20%, +20%)
Contribution of ABT-333	F-C 12/12 (100%) – 13/13 (100%)	0% (-19%, +19%)
Contribution of RBV	G-E 13/13 (100%) – 27/27 (100%)	0% (-15%, +15%)

**Statistically significant at $p<0.05$
Results for "all" are based on stratified MH and results by IL28B are based on an exact test

As shown in Table 3.1.2.1.4 on page 17, SVR rates in the five arms of treatment-experienced null responders ranged from 89% to 100%. Three hypotheses can be tested as shown in Table 3.1.2.1.8 but the interpretation is limited by small sample sizes and wide confidence intervals. Further analyses of these data by subgenotypes are not helpful given the small sample sizes.

Table 3.1.2.1.8 Study M11-652 SVR_{24} results for treatment experienced – null responders patients

Test hypothesis	Arm Difference	Treatment difference 95% CI	Interpretation
Trt for 24 wks vs. 12 wks	N-L	+5% (-10%, +20.5%)	NS
Contribution of ABT-333 (150/100v200/100)	L-J	+6% (-9%, +22%)	NS
ABT-450/r dose 150/100 versus 100/100 (12 wk)	L-K	+3% (-15%, +22%)	NS

NS=Not statistically significant

The table below summarizes the reasons patients did not achieve SVR_{24} by treatment arm.

Table 3.1.2.1.9 Phase 2 Study M11-652 Reasons patients did not achieve SVR_{24} ITT population
Cohort 1 (Arms A-I, treatment naïve patients) rows have a clear background while
Cohort 2 (Arms J-N, treatment experienced null responders) rows have a grey background
The proposed regimen for approval is bolded.

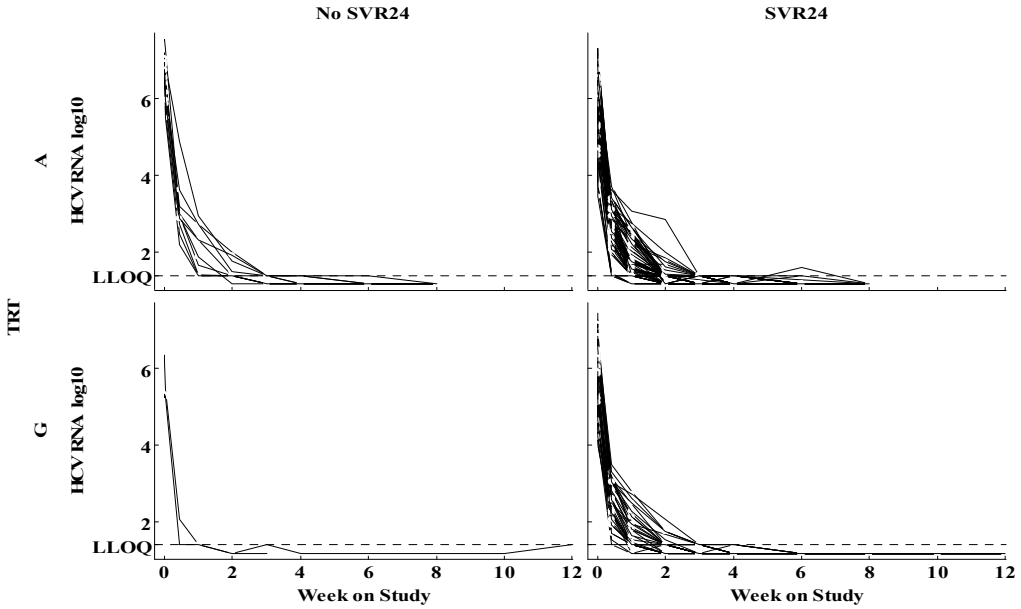
ARM (SVR_{24})	Regimen	Virologic failures Week of failure	Relapses	Lost-to-follow-up Week of loss and last SVR
A (70/80)	3DAA+RBV 8 wks	0	10	0
B (34/41)	450r+333+RBV 12 wks	1 Wk 8	4	1 discont wk6 no SVR data 1 discont ptwk12 SVR12 neg
C (33/39)	450r (100/100) +267+RBV 12 wks	0	6	0
D (37/40)	450r (200/100) +267+RBV 12 wks	1 Wk 4	2	0
E (70/79)	3DAA 12 wks	1 Wk 12	5	1 discont ptwk2, SVR2 neg 2 discont ptwk8, SVR12 neg
F (38/39)	3DAA (450r dose 100/100)+RBV 12 wks	0	0	1 discont ptwk8 SVR12 neg
G (38/40)	**3DAA+RBV 12 wks**	**0**	**1**	**1 discont ptwk8 SVR12 neg**
H (37/40)	3DAA (450r dose 100/100)+RBV 24 wks	0	1	1 discont wk10, no SVR data 1 discont wk20, no SVR data 1 discont ptwk8, SVR8 neg
I (36/40)	3DAA+RBV 24 wks	0	2	1 discont ptwk8, SVR8 neg 1 discont ptwk12 SVR12 neg
J (40/45)	450r (200/100) +267+RBV 12 wks	0	5	0
K (21/23)	3DAA (450r dose 100/100)+RBV 12 wks	2 Wks 6 & 8	0	0
L (21/22)	**3DAA+RBV 12 wks**	**1 Wk 12**	**0**	**0**
M (21/23)	3DAA (450r dose 100/100)+RBV 24 wks	1 Wk 16	0	1 discont ptwk12 SVR12 neg
N (20/20)	3DAA+RBV 24 wks	0	0	0

Source: Applicant's study report and reviewer results
All patients recorded as virologic failures were also recorded as rebounds in the database.

There were seven virologic failures observed in the 14 arms; one occurred in a 1b patient and all the rest were in 1a genotype patients. One virologic failure at Week 12 occurred for the proposed regimen (3DAA+RBV) in a null responder on Arm L. The highest relapse rates are seen for Arms C (15%) and J (11%) [arms without 333] and for Arm A (12.5%) [arm with 8 weeks of treatment]. A little more than half of the relapses occurred within 4 weeks and about 80% occurred within 8 weeks of stopping treatment. Most lost-to-follow-ups occurred during post-treatment weeks. Of the 13 patients lost to follow-up, 10 had an SVR at time of dropout (half at post-treatment week 12).

To see whether there is any suggestion that patients who do not attain an SVR_{24} have an early response to treatment that is distinguishable from the response of patients who do attain an $SVR_{24,}$ this reviewer plotted HCV RNA for each individual patient during treatment. The data in the graph below does not suggest a different pattern of response for the two possible outcomes that would be predictive of response. The earliest responses (decrease to LLOQ or less) are seen within a week and all patients have responded by about week 4 (these observations are true for all 14 treatment arms). This Phase 2 data does not suggest that the timing of response to the tested regimens is useful for identifying responders.

Figure 3.1.2.1.4 HCV RNA by treatment week by patient for main comparison A vs G and for patients with an SVR_{24} and without an SVR_{24}



The applicant performed a logistic regression analysis where they concluded that patients with lower baseline HCV RNA levels, with non-1a HCV subtype, or who received all 3 DAAs were more likely to achieve SVR compared with patients without these factors. The latter two findings are in agreement with the findings presented above. This reviewer also performed the primary analysis by subgroups defined by the median baseline HCV RNA (median=4,220,000 IU/mL) and also the analysis comparing the full regimen with RBV and without RBV. The latter comparison is an important comparison that was considered in three of the six Phase 3 trials. The treatment effects for both comparisons are higher for patients with HCV RNA greater than the baseline median; the interaction is significant (assuming an alpha of 0.10 for an interaction) when comparing the durations but not when comparing the addition of RBV. This data suggests that the 12 week duration is especially important for patients with high HCV RNA at the start of treatment.

Table 3.1.2.1.10 Results for two comparisons by the median baseline HCV RNA (4,220,000 IU/mL)

Baseline HCV RNA	3DAA+RBV 12 wks	3DAA+RBV 8 wks	Risk difference	Int. p-value[1]
< Median	24/26 (92%)	34/36 (94%)	-2% (-15%, +10%)	
≥ Median	14/14 (100%)	36/44 (82%)	+18% (+7%, +30%)	0.09
	3DAA+RBV 12 wks	3DAA 12 weeks	Risk difference	Int. p-value
< Median	24/26 (92%)	43/47 (91.5%)	+0.1% (-12%, +14%)	
≥ Median	14/14 (100%)	27/32 (84%)	+16% (+3%, +28%)	0.18

[1]P- value is based on test of homogeneity across subgroups. A significant p-value indicates that the treatment effects are statistically significantly different between the subgroups.

Summary

This reviewer has the following observations for Phase 2 study M11-652 which may also be relevant when considering the Phase 3 results:

1. The regimen appears to be well-tolerated with most arms completing treatment (>90%). Drug discontinuations were mostly seen in the treatment arms with a 24 week treatment period (9%-18%).
2. Because of patient discontinuations, SVR_{12} rates were higher (by 1-5%) than SVR_{24} rates in 8 out of the 14 arms.
3. Primary endpoint comparing 12 weeks to 8 weeks showed an advantage for 12 weeks of about 7% in SVR_{24}. There was 1 (2.5%) relapse in the 12 week arm versus 10 (12.5%) in the 8 week arm. The difference is driven by the 1a population where 9% more responders were seen for 12 weeks than 8 weeks).
4. ABT-267, ABT-333 and RBV were all shown to improve the SVR_{24} rates in the naïve 1a population so this study's results suggest that the full regimen of 3DAA plus RBV may be best for 1a naïve patients.
5. Response rates in the naïve 1b population were 100% in arms of the full regimen and when excluding any one of ABT-267, ABT-333 and RBV so there is no clear evidence in this trial that 1b patients need all the drugs in the regimen of 3DAA+RBV. However, given the paucity of data, one cannot rule out that the full regimen may be the most beneficial.
6. Subgroup results by IL28B in the 1a population are inconclusive but do not suggest differing effects across IL28B genotypes.
7. The data for null responders is minimal and offers minimal evidence regarding the contribution of the components, dosing or length of treatment.

3.1.3 Phase 3 Studies by Study Design

3.1.3.1 Studies comparing 3DAA+RBV versus 3DAA+Placebo

Studies M13-389, M13-961 and M14-002

Design

Studies M13-389, M13-961 and M14-002 were all Phase 3 trials designed to evaluate the safety and efficacy of the combination of ABT-450/r/ABT-267 and ABT-333 (3DAA) with and without RBV in non-cirrhotic patients with chronic HCV infection. The designs for these three studies are summarized in the table below.

Table 3.1.3.1.1 Summary of designs for three Phase 3 trials comparing 3DAA+RBV and 3DAA

	M13-389	M13-961	M14-002
Study name	PEARL-II	PEARL-III	PEARL-IV
Conducted	8/2012 - ongoing	12/2012 - ongoing	8/2012 - ongoing
Design	OL, R, MC	DB, R, MC	DB, R, MC
Arms	3DAA+RBV, 3DAA	3DAA+RBV, 3DAA	3DAA+RBV, 3DAA
Duration	12 wks Trt + 48 wks FU	12 wks Trt + 48 wks FU	12 wks Trt + 48 wks FU
Stratifiers	Null, nonresp/partial, relapser	CC, non-CC	CC, non-CC
Sites	43 US, Puerto Rico & EU	50 US, Russia, & EU	53 US, Canada & UK
Patient Population	GT 1b non-cirrhotic TE	GT 1b non-cirrhotic TN	GT 1a non-cirrhotic TN
Primary endpoint	SVR_{12}	SVR_{12}	SVR_{12}
Historical control	telaprevir plus pegIFN/RBV	telaprevir plus pegIFN/RBV	telaprevir plus pegIFN/RBV
Thresholds	NI=64% SUP=75%	NI=73% SUP=84%	NI=65% SUP=75%

OL=open label, DB=double-blind,R=randomized, MC=multicenter, Trt=treatment, FU=follow-up off treatment, TE=treatment experienced, TN=treatment naïve, NI=non-inferiority, SUP=superiority

All patients were to be treated for 12 weeks and then followed for 48 weeks (Figure 3.1.3.1.1). Patients who discontinued treatment early were also to be followed for 48 weeks. All study drugs were to be taken orally with food. Compliance was monitored using a MEMS cap. HCV RNA samples were collected at baseline, treatment weeks 1, 2, 4, 6, 8, 10 and 12 and at post-treatment weeks 2, 4, 8, 12, 24, 36 and 48 or last visit. The primary analysis was planned when all patients had completed at least 12 weeks post treatment or discontinued early.

Figure 3.1.3.1.1 Applicant's schematic of the trial design for Studies M13-389, M13-961 and M14-002

The protocols were amended in 2013, while the trials were ongoing, to not allow hormonal contraceptives during administration of study drugs for female patients. No women discontinued study drug due to being unwilling or unable to stop hormonal contraceptives.

As described in the protocol, the primary objective of all three studies was to compare the safety of 3DAA+RBV and 3DAA+placebo and secondly to show that each arm is non-inferior in SVR_{12} rates to the historical SVR rate of telaprevir plus pegIFN and RBV therapy based on the thresholds shown on the previous page in Table 3.1.3.1.1. Both the non-inferiority thresholds and the superiority thresholds were easily met so the primary focus of this review is to measure the contribution of the addition of RBV to 3DAA in improving the SVR_{12} rates. This question is also examined for several subgroups however the trials are not powered to show treatment differences by subgroups, so the goal is to determine if the effect seen overall is consistent across subgroups.

A secondary efficacy endpoint was the comparison of the percentage of subjects with a decrease in hemoglobin (Hb) to below the lower limit of normal (LLN) at the end of on treatment for patients with hemoglobins at or above the LLN at baseline. Decreases in Hb are an expected side-effect for RBV so investigators were blinded in Studies M13-961 and M14-002 (M13-389 is an open-label study) to Hb results unless changes in Hb suggested unacceptable toxicity as defined in the protocol. There were 22 patients in M13-961 and 6 patients in M14-002 whose assigned drug was unblinded due to toxicity; this reviewer is not concerned about this unblinding for an important safety issue and does not think it would bias the HCV RNA results.

Patient Disposition

Patients were screened at more than 140 sites in several countries. About 60-70% of patients screened were randomized to treatment with the primary reason for not being randomized was the patient did not meet entry criteria (Table 3.1.3.1.2)

Table 3.1.3.1.2 Patients screened and randomized by study

	M13-389	M13-961	M14-002
Screened	324	629	436
Randomized (% of screened)	187 (58%)	419 (67%)	305 (70%)
Primary reason for not being randomized (% of screened)	Did not meet entry criteria (94%)	Did not meet entry criteria (89%)	Did not meet entry criteria (90%)

Patient disposition was assessed when all patients had completed at least 24 weeks on study. Very few patients (2% or less) discontinued drug or study in Studies M13-389 and M13-961 (Table 3.1.3.1.3 on following page). For Study M14-002, the highest discontinuation rate was seen for the 3DAA arm with 11(5%) patients discontinuing study drug; six discontinued due to virologic failure (4 of the 6 discontinued from the study). Only one patient in all three studies was recorded as discontinuing the trial due to an adverse event. Less than 12% of patients in all studies had HCV RNA data out to 24 weeks post-treatment.

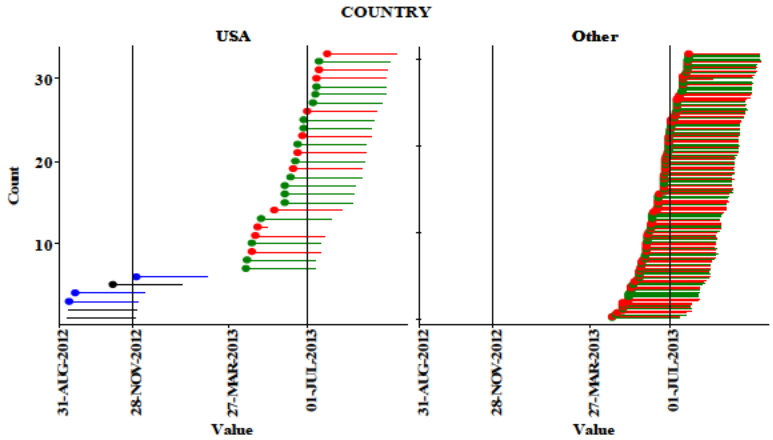
Table 3.1.3.1.3 Patient disposition by study

	M13-389		M13-961		M14-002	
	GT 1b TE		GT 1b TN		GT 1a TN	
	3DAA+RBV	3DAA	3DAA+RBV	3DAA	3DAA+RBV	3DAA
Randomized	92	95	210	209	100	205
Rand. +Trt (ITT)	91	95	210	209	100	205
Discont. drug	**2 (2%)**	**0**	**1 (0.5%)**	**1 (0.5%)**	0	**11 (5%)**
ADE	2 (2%)	0	0	0	0	2 (1%)
LTFU	0	0	0	0	0	3 (1.5%)
VF	0	0	0	0	0	6 (3%)
Withdrew consent	0	0	1 (0.5%)	1 (0.5%)	0	0
Discont. study	**1 (1%)**	**0**	**0**	**1 (0.5%)**	**0**	**7 (3%)**
ADE	1 (1%)	0	0	0	0	0
LTFU	0	0	0	1 (0.5%)	0	3 (1.5%)
VF/Entered ext	0	0	0	0	0	2 (1%)
VF/Other	0	0	0	0	0	2 (1%)
Weeks completed on study[1]						
8	91 (100%)	95 (100%)	210 (100%)	209 (100%)	100 (100%)	202 (98.5%)
12	91 (100%)	95 (100%)	210 (100%)	209 (100%)	100 (100%)	199 (97%)
20 (PTW8)	88 (97%)	95 (100%)	210 (100%)	207 (99%)	99 (99%)	197 (96%)
22 (PTW10)	88 (98%)	95 (100%)	208 (99%)	207 (99%)	98 (98%)	196 (96%)
24 (PTW12)	82 (90%)	88 (93%)	190 (90%)	192 (92%)	88 (88%)	172 (84%)
26 (PTW14)	12 (13%)	14 (15%)	20 (10%)	21 (10%)	3 (3%)	8 (4%)
30 (PTW18)	11 (12%)	11 (11%)	20 (10%)	19 (9%)	1 (1%)	2 (1%)
36 (PTW24)	11 (12%)	10 (11%)	19 (9%)	16 (8%)	1 (1%)	0
48 (PTW36)	3 (3%)	3 (3%)	1 (0.5%)	1 (0.5%)	0	0

[1] PTW=post-treatment week. These counts are based on the last available HCV RNA data for each patient. Note that a patient may have SVR data in the PTW12 window but not be counted as completing a full 24 weeks on study.

In the study report for Study M13-389, the applicant stated that they excluded 6 patients from the ITT population because they were not given the co-formulated product (as was the case for M11-652). This reviewer included these 6 patients as part of the ITT population because they were all treated with their randomized drugs and followed. The graph below illustrates that these 6 subjects were the first enrolled at USA sites.

Figure 3.1.3.1.2 Time of enrollment and duration on study by patient and country (USA or other)



Baseline Demographics

Within each of the three studies (M13-389, M13-961 and M14-002), the treatment arms are comparable based on baseline demographics and disease characteristics. In all studies, the majority of patients are white (~90%) with an average age of about 50 years. All studies enrolled patients at US sites with Study M14-002 having the largest enrollment at 81%. The median baseline HCV RNA $\log_{10}$ ranged from 6.4 to 6.7.

Table 3.1.3.1.4 Baseline demographics by study

	M13-389		M13-961		M14-002	
	GT 1b TE		GT 1b TN		GT 1a TN	
	3DAA+RBV N=91	3DAA N=95	3DAA+RBV N=210	3DAA N=209	3DAA+RBV N=100	3DAA N=205
% USA	15%	20%	23%	22%	81%	81%
% Canada	NA	NA	NA	NA	11%	10%
% Europe	85%	80%	66%	61%	8%	9%
% Other	NA	NA	11%	17%	NA	NA
% Female	50%	40%	50%	59%	30%	37%
% White	92%	91%	94%	94%	86%	83%
Age						
Mean (SD)	54 (11)	54 (11)	48 (12)	49 (12)	52 (11)	51 (11)
Min-Max	26-68	28-70	19-70	22-70	19-69	21-70
% ≥ 65 years	17%	17%	7%	9%	10%	6%
HCV genotype						
1a	2%	1%	0%	0%	100%	99.5%
1b	98%	98%	100%	100%	0%	0.5%
Other	0%	1%	0%	0%	0%	0%
IL28B						
CC	11%	7%	21%	21%	31%	31%
CT	65%	71%	61%	63%	58%	51%
TT	24%	22%	19%	16%	11%	18%
Treatment experience						
Null responder	35%	35%	NA	NA	NA	NA
Partial responder	29%	28%				
Relapser	36%	37%				
Baseline HCV RNA Log_{10} IU/mL						
Mean (SD)	6.6 (0.6)	6.5 (0.5)	6.3 (0.8)	6.3 (0.7)	6.6 (0.5)	6.5 (0.7)
Median	6.6	6.5	6.4	6.5	6.7	6.7

NA=not applicable

Two patients in each arm of Study M13-389 (a study of 1b patients) who were recorded as 1a sub-genotypes are included in the ITT population. For Study M14-002, a study of 1a patients, one patient was infected with HCV subgenotype 1b and was counted as part of the ITT population. This reviewer agrees with including these randomized patients.

Statistical Methods

See Section 3.1.1 of this review for a general description of the statistical methods used by the applicant and this reviewer. Methods specific to this study are described here.

The applicant updated the data for Studies M13-389 and M13-961 with a submission dated 7/1/2014. This updated data was used in the analyses performed by this reviewer.

The same statistical methods were used for all three studies described here. A fixed sequence testing procedure was planned where the test of non-inferiority of each arm to the historical control was performed first and then secondary analyses were only to be performed if non-inferiority was demonstrated. Secondary analyses included test of superiority compared to the historical control and head-to-head comparisons of the two arms.

The difference in SVR_{12} rates between treatment arms was analyzed using the stratum-adjusted Mantel-Haenszel (MH) test with a continuity correction.

A Fisher's exact test was used to compare the percentage of patients with hemoglobin (Hb) below the lower limit of normal (LLN) at the end of treatment.

Results

For all six arms in the three studies, each arm is shown to be superior to telaprevir+pegIFN+RBV with the lower bound of the confidence interval (all 90% or larger) around the observed SVR_{12} rate being larger than the predefined threshold for superiority (84% for M13-961 and 75% for M13-389 and M14-002).

Table 3.1.3.1.5 SVR_{12} rates by study and by arm

Study	3DAA+RBV SVR_{12} rate (95% CI)[1]	3DAA SVR_{12} rate (95% CI)	3DAA+RBV - 3DAA Difference (95% CI)[2]	p-value
M13-389 GT 1b TE	89/91 98% (92%, 100%)	95/95 100% (96%, 100%)	-2% (-6%, +2%)	0.16
M13-961 GT 1b TN	209/210 99.5% (97%, 99.9%)	209/209 100% (98%, 100%)	+0.5% (-1%, +2%)	>0.30
M14-002 GT 1a TN	97/100 97% (94%, 100%)	185/205 90% (86%, 94%)	+6.8% (+1.3%, +12%)	0.035

[1]Clopper Pearson exact confidence interval [2] Results of MH stratified analysis

A comparison of randomized arms shows no statistically significant difference between the arms in Studies M13-389 and M13-961 indicating that RBV does not significantly improve the SVR_{12} rate over 3DAA alone in patients with HCV subgenotype 1b. A statistically significant treatment difference of about 7% (p=0.035) is seen in Study M14-002 indicating that the addition of RBV to the 3DAA regimen is beneficial for GT 1a treatment naïve patients.

For Studies M13-389 and M13-961, focusing on only the 3DAA arm, there was one failure in both the GT 1b TE population and the GT 1b TN population (Table 3.1.3.1.6 on the following page).

Study M14-002 results show a clear treatment difference in relapses with only 1% on the full regimen and 5.4% on the reduced regimen without RBV (Table 3.1.3.1.6 on the following page); the treatment difference is statistically significant [-4% (95% CI of -8% to -1%) Fisher's exact test p=0.047].

Table 3.1.3.1.6 Reasons for not achieving SVR12 by study

	3DAA+RBV			3DAA		
	On trt Virologic failures	Relapses	Lost-to-follow-up	On trt Virologic failures	Relapses	Lost-to-follow-up
M13-389 GT 1b TE	0/91 0%	1/91 1% PTW4 Only 4 wks of trt	1/91 1% TW1 RNA 7 PTW12	0/95 0%	0/95 0%	0/95 0%
M13-961 GT 1b TN	1/210 0.5% TW10	0/210 0%	0/210 0%	0/209 0%	0/209 0%	0/209 0%
M14-002 GT 1a TN	1/100 1% TW8	1/100 1%** PTW2	1/100 1% TW12 w/ RNA 1.17	6/205 2.9% TWs 2-8	11/205 5.4%** 6 at PTW2 5 at PTW4 1 w/4 wks of trt	3/205 1.5%

**Relapse rates different p=0.047 Fisher's exact test

Subgroup analyses for Study M14-002 further support the benefit of adding RBV to the 3DAA regimen for treatment naïve GT 1a patients with subgroups consistently showing an increase in SVR_{12} rates for the full regimen over the 3DAA. Only for subgroups defined by median age is there any evidence of heterogeneity across subgroups with older patients showing more benefit adding RBV (p=0.09 for a test for interaction).

Table 3.1.3.1.7 Study 14-002 Subgroup results for the 1a TN population

	3DAA+RBV	3DAA	Trt diff (95%CI)	p-value for Int.[1]
Overall	97/100 97%	185/205 90%	+6.8% (+1.3%, +12%)	
Subgenotype				
CC	31/31 100%	61/63 97%	+3% (-1%, +8%)	
Non-CC	66/69 96%	124/142 87%	+8% (+1%, +16%)	>0.3
Baseline HCV RNA				
By median				
< 5,080,000	45/47 96%	94/104 89%	+6% (-2%, +15%)	
≥ 5,080,000	52/53 98%	92/101 91%	+7% (+0.4%, +14%)	>0.3
By tertiles				
<2,730,000	29/30 97%	65/71 92%	+5% (-4%, +14%)	
2,730,000-7,960,000	35/37 95%	56/66 85%	+10% (-2%, +21%)	>0.3
>7,960,000	33/33 100%	64/68 94%	+6% (+0.3%, +11%)	
Sex				
Male	67/70 (96%)	113/129 (88%)	+8% (+0.7%, +16%)	
Female	30/30 (100%)	72/76 (95%)	+5% (+0.2%, +10%)	>0.3
Median Age years				
<54	45/48 (94%)	90/98 (92%)	+2% (-7%, +11%)	
≥ 54	52/52 (100%)	95/107 (89%)	+11% (+5%, +17%)	**0.09**
Median BMI kg/m^2				
<26	39/39 (100%)	98/106 (92%)	+8% (+2%%, +13%)	
≥ 26	58/61 (95%)	87/99 (88%)	+4% (-1%, +16%)	>0.3
Geographic area				
US	78/81 (96%)	147/166 (89%)	+8% (+1%, +14%)	
Other	19/19 (100%)	38/39 (97%)	+3% -2%, +8%)	>0.3

[1] Zelen's exact test of homogeneity was used to test for an interaction; a large p-value indicates homogeneity across the subgroup effects.

Because of the expected effect of RBV to decrease hemoglobin (Hb), the percentage of patients with a decrease in Hb below LLN at end of treatment was named as a secondary endpoint in all three protocols. This reviewer focused on the Hb results for Study 14-002 where the full regimen was seen to be beneficial for the GT 1a treatment naïve population. Additionally, it is unlikely that treatment with 3DAA and RBV will be recommended for the HCV 1b population.

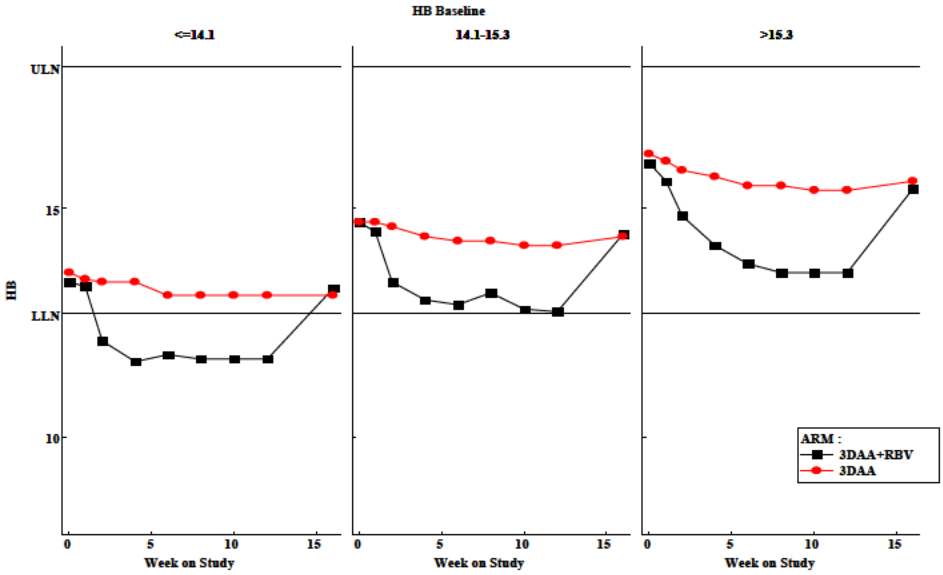
The applicant reported that the percentage of subjects with a decrease in Hb below the LLN at the end of treatment was statistically significantly higher in the 3-DAA + RBV treatment group (42.0%) than in the 3-DAA treatment group (5.5%) among subjects with Hb≥ LLN (12.7 g/dL) at baseline. This reviewer obtained slightly different numbers than the applicant but the results are nevertheless consistent (Table 3.1.3.1.8). The changes in Hb are baseline related and on average the decreases in Hb occur in the first 5 weeks of treatment (see Figure 3.1.3.1.3 below).

Table 3.1.3.1.8 Study M14-002 Hb results (g/dL)

	3DAA+RBV	3DAA	p-value
Patients w/Bsl Hb>LLN			
% w/Hb<LLN EOT	52/95 55%	16/191 8%	<0.001
% w/Hb<LLN EOT by baseline Hb tertiles			
Hb<14.1	24/29 83%	9/49 18%	
14.1≤Hb≤15.3	16/32 50%	5/73 7%	
Hb>15.3	12/34 35%	1/68 1%	
All Patients Mean (SD)			
Baseline Hb	14.7 (1.3)	14.8 (1.4)	
End of trt Hb	12.6 (1.4)	14.2 (1.3)	
End of study Hb	14.3 (1.3)	14.3 (1.4)	
Anemia on study	9/100 9%	1/205 0.5%	<0.0003

EOT=end of treatment EOS=end of study Analysis of ADLBHEMA dataset by reviewer

Figure 3.1.3.1.3 Study M14-002 All patients Median Hb overtime by baseline Hb tertiles and treatment arm



To observe whether the addition of RBV is associated with increased adverse events in treatment naïve genotype 1a patients in Study M14-002, this reviewer summarized the results for adverse events (selected based on input from the clinical reviewer) where a nominal p-value below 0.05 was observed and for adverse events that may be associated with RBV treatment. The most significant difference is seen for bilirubin increases (7% vs. 0%). The importance of these differences is a clinical issue.

Table 3.1.3.1.9 Study 14-002 Adverse event rates based on reported Meddra preferred terms during the 12 week treatment period

	3DAA+RBV N=100	3DAA N=205	Treatment Difference (95%CI)	*p-value*
Blood bilirubin inc	7 (7%)	0	+7% (2%, 12%)	*<0.001*
Anemia	4 (4%)	0	+4% (0.2%, 8%)	*0.011*
Insomnia	17 (17%)	16 (8%)	+9% (1%, 17%)	*0.019*
Pruritus	10 (10%)	12 (5%)	+5% (-2%, 11%)	*0.15*
Fatigue	44 (44%)	72 (35%)	+9% (-3%, 21%)	*0.17*
Asthenia	3 (3%)	1 (1%)	+2% (-1.6%, 6%)	≥0.3
Dizziness	7 (7%)	13 (6%)	+1% (-5%, +7%)	≥0.3
Dyspnea	4 (4%)	6 (3%)	+1% (-3%, +6%)	≥0.3
Hyperbilirubinaemia	1 (1%)	0	+1% (-1%, +3%)	≥0.3
Nausea	21 (21%)	28 (14%)	+7% (-2%, 17%)	≥0.3
Pyrexia	3 (3%)	4 (2%)	+1% (-3%, +5%)	≥0.3
Rash	5 (5%)	9 (4%)	+1% (-4%, +6%)	≥0.3
Upper Abdom. Pain	2 (2%)	3 (1.5%)	+0.5% (-3%, +4%)	≥0.3
Vomiting	4 (4%)	4 (2%)	+2% (-2%, +6%)	≥0.3
Any severe AE	2 (2%)	3 (1.5%)		

Summary

Studies M13-389 (TE 1b patients) and M13-961 (TN 1b patients) both showed that treatment with 3DAA without RBV is as effective as the regimen with RBV for patients with subgenotype 1b. No patients in either study relapsed or had an on-treatment virologic failure on 3DAA. The SVR_{12} rate with both studies combined is 100% with a tight confidence interval of 98.7% to 100%. These two studies provide strong statistical evidence to recommend 3DAA treatment for patients with subgenotype 1b regardless of treatment experience history.

Study M14-002 was conducted in treatment naïve genotype 1a patients. A statistically significant increase of 7% (95% CI +1.3%, +12%) for the regimen with RBV compared to the regimen without demonstrates that RBV should be recommended in combination with 3DAA for HCV 1a TN patients. Also subgroup analyses consistently showed increased SVR_{12} rates with the addition of RBV.

One of the safety issues with RBV administration is the effect on red blood cells with hemoglobin decreasing and the incidence of anemia increasing. In Study M14-002, the difference in anemia rates was about 8.5% (9% on 3DAA+RBV and 0.5% on 3DAA, $p<0.0003$) and significant decreases in hemoglobin were observed within about 5 weeks of treatment. Hemoglobin generally returned to baseline a few weeks after stopping RBV treatment. It does not appear that the benefit of the added RBV is outweighed by this potential safety issue.

3.1.3.2 Studies comparing 3DAA+RBV versus Placebo

Studies M11-646 and M13-098

Design

Studies M11-646 and M13-098 were two double-blind, Phase 3 trials designed to evaluate the safety and efficacy of the combination of ABT-450/r/ABT-267 and ABT-333 (3DAA) with RBV in non-cirrhotic genotype 1 patients with chronic HCV infection. The designs for these two studies are summarized in the table below.

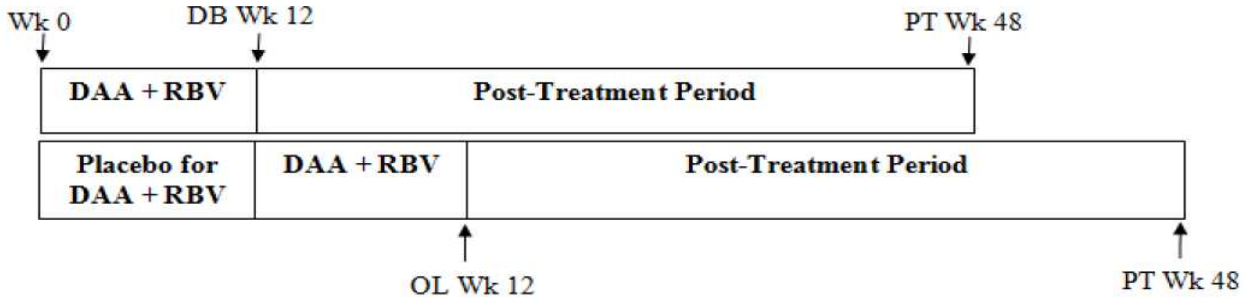
Table 3.1.3.2.1 Summary of designs for two Phase 3 trials comparing 3DAA+RBV and Placebo

	M11-646	**M13-098**
Study name	SAPPHIRE-I	SAPPHIRE-II
Conducted	12/2012 - ongoing	11/2012 - ongoing
Design	DB, R, MC	DB, R, MC
Arms & Duration	3DAA+RBV 12 wks Trt + 48 wks FU Placebo 12 wks; 3DAA+RBV12wks+48wks FU	3DAA+RBV 12 wks Trt + 48 wks FU Placebo 12 wks; 3DAA+RBV12wks+48wks FU
Stratifiers	1a or non-1a and CC or non-CC	1a or non-1a and null, nonresp/partial or relapser
Sites	79 US, Canada, & EU	76 US, Australia, Canada, & EU
Patient Population	GT 1 non-cirrhotic TN	GT 1 non-cirrhotic TE
Primary endpoint	SVR_{12}	SVR_{12}
Historical control	telaprevir plus pegIFN/RBV	telaprevir plus pegIFN/RBV
Margins 1 1a 1b	 NI=70% SUP=80% SUP=75% SUP=84%	 NI=60% SUP=70% SUP=65% SUP=77%

DB=double-blind, R=randomized, MC=multicenter, Trt=treatment, FU=follow-up off treatment, TE=treatment experienced, TN=treatment naïve, NI=non-inferiority, SUP=superiority

These studies were unique in that they included a placebo arm. Patients were randomized to 3DAA+RBV or placebo. Patients randomized to 3DAA+RBV were treated for 12 weeks and then followed for 48 weeks post treatment (Figure 3.1.3.2.1). Patients randomized to placebo were followed for 12 weeks and then treated open label with 3DAA+RBV for 12 weeks and then followed off treatment for 48 weeks.

Figure 3.1.3.2.1 Applicant's trial schematic



The applicant stated that the inclusion of a blinded placebo arm allows for "*characterization of adverse events in an untreated, chronically HCV infected population.*" A primary purpose of the trial then appears to be to compare safety outcomes during the initial 12 weeks.

With regards to efficacy, the 3DAA+RBV arm was compared against the historical rate for telaprevir plus pegIFN and RBV therapy. Based on predefined boundaries, non-inferiority was tested first and if demonstrated then superiority was tested.

The first secondary endpoint was a comparison of the percentage of subjects with ALT normalization at the final treatment visit in the DB treatment period between the placebo and 3DAA+RBV. ALT normalization was defined as having a final ALT ≤ ULN in the DB Treatment Period in patients with ALT above the upper limit of normal (ULN) at baseline.

Patient Disposition

Patients were screened at more than 150 sites in several countries. About 72% of patients screened were randomized to treatment with the primary reason for not being randomized was the patient did not meet entry criteria (Table 3.1.3.2.2)

Table 3.1.3.2.2 Patients screened and randomized by study

	M11-646	M13-098
Screened	855	562
Randomized (% of screened)	636 (74%)	395 (70%)
Primary reason for not being randomized (% of screened)	Did not meet entry criteria (93%)	Did not meet entry criteria (92%)

Patients were randomized in a 3:1 ratio to 3DAA+RBV and placebo. A total of 631 patients were randomized and treated in the study of treatment naïve patients (M11-646) and 394 treatment-experienced patients were treated in Study 13-098. Six patients in both studies did not receive randomized drug for a variety of reasons, such as change in health status, withdrew consent and did not satisfy entry criteria; there does not appear to be any bias introduced by choosing not to treat these randomized patients.

Table 3.1.3.2.3 Patient disposition by study

	M11-646		M13-098	
	GT 1 TN		GT 1 TE	
	3DAA+RBV	Placebo	3DAA+RBV	Placebo
Randomized	477	159	297	98
Rand. +Trt (ITT)	473	158	297	97
Discont. drug[1]	7 (1.5%)	1 (1%)	5 (1.6%)	2 (2%)
ADE	2 (0.4%)	1 (1%)	3 (1%)	0
LTFU	1 (0.2%)	0	0	0
Withdrew consent	2 (0.4%)	0	1 (0.3%)	2 (2%)
Noncompliant	2 (0.4%)	0	1 (0.3%)	0
Discont. study [2]	7 (1.5%)	6 (4%)	6 (2%)	2 (2%)
ADE	2 (0.4%)	0	1 (0.3%)	0
LTFU	2 (0.4%)	0	0	0
Withdrew consent	3 (0.6%)	4 (2.5%)	4 (1%)	2 (2%)
Entered ext	0	1 (0.6%)	1 (0.3%)	0
VF/Other	0	1 (0.6%)	0	0
Completed DB TW12	464 (98%)	157 (99%)	292 (98%)	96 (99%)

[1] Based on ADEFFOUT dataset

[2]These counts are based on the TW 12 HCV RNA data for each patient. Note that a patient may have data in the TW12 window but not be counted as completing a full 12 weeks on study.

About 98% of patients completed the double-blind period. Very few patients discontinued treatment or the study due to an adverse event (≤1%).

Baseline Demographics

The treatment groups in each study were reasonably balanced with regard to baseline characteristics (Table 3.1.3.2.4). About half the patients in each study were enrolled in North America (~40% in USA). The rest were mostly enrolled in Europe. About half the patients were male and about 91% were white. The average age was about 52 and less than 10% were 65 or older; in Study M13-098, the placebo group had statistically significantly more patients over 65. A majority of patients were genotype 1a in each study.

Table 3.1.3.2.4 Baseline demographics by study

	M11-646		M13-098	
	GT 1 TN		GT 1 TE	
	3DAA+RBV N=473	Placebo N=158	3DAA+RBV N=297	Placebo N=97
% USA	41%	42%	40%	32%
% Canada	7%	4%	6%	2%
% Europe	45%	45%	50%	63%
% Aus/NZ	7%	9%	4%	3%
% Female	43%	54%	44%	38%
% White	91%	91%	91%	89%
Age				
Mean (SD)	49 (11)	51 (10)	52 (10)	55 (8)
Min-Max	18-70	21-70	19-71	30-69
% ≥ 65 years	4%	6%	7%	13%
HCV genotype				
1a	68%	67%	58%	59%
1b	32%	33%	41%	41%
Other	0%	0%	0.3%	0%
IL28B				
CC	30%	32%	11%	7%
CT	54%	52%	67%	72%
TT	16%	16%	21%	21%
Treatment experience				
Null responder	NA	NA	49%	49%
Partial responder			22%	21%
Relapser			29%	29%
Baseline HCV RNA Log_{10} IU/mL				
Mean (SD)	6.4 (0.6)	6.5 (0.6)	6.6 (05)	6.5 (0.5)
Median	6.5	6.6	6.7	6.6

NA=not applicable

Statistical Methods

See Section 3.1.1 of this review for more details regarding the statistical methods used by the applicant and this reviewer. The applicant updated the data for Study M11-646 with a submission dated 7/1/2014. This updated data was used in the analyses performed by this reviewer.

A fixed-sequence testing procedure was used to proceed through the primary and secondary efficacy endpoints with no adjustments made for multiple comparisons.

The percentage of patients in the 3DAA+RBV arm with ALT normalization was compared to the percentage of patients in the placebo arm with ALT normalization using Fisher's exact test.

Results

The SVR_{12} results for the 3DAA+RBV arm are summarized in the table below. Looking first at the treatment naïve population in Study M11-646, it is evident that the test regimen is superior to the historical control with lower limits of the 95% confidence (93% and 94%) well about the predefined superiority threshold for both genotype 1a and 1b patients. Subgroups defined by sub-genotypes of CC or non-CC also had high event rates; no threshold for superiority was named but it is clear that the confidence intervals suggest that the regimen was effective across subgroups.

Table 3.1.3.2.5 SVR_{12} rates (95% CI) by study and by arm 3DAA+RBV

Study	SVR_{12} rate	95% CI[1]	Superiority Threshold
M11-646 TN			
All	456/473 96%	94%, 98%	80%
1a	308/322 96%	93%, 98%	75%
1b	148/151 98%	94%, 99.5%	84%
M11-646 TN			
1a CC	103/106 97%	92%, 99%	
1a Non-cc	205/216 95%	91%, 97%	
1b CC	36/38 95%	82%, 99%	
1b Non-cc	112/113 99%	95%, 100%	
M13-098 TE			
All	286/297 96%	93%, 98%	70%
1a	166/173 96%	92%, 98%	65%
1b	120/124 97%	92%, 99%	77%
M13-098 TE			
1a Null	83/87 95%	89%, 99%	
1a Partial	36/36 100%	90%, 100%	
1a Relapser	47/50 94%	83%, 99%	
1b Null	56/59 95%	86%, 99%	
1b Partial	29/29 100%	88%, 100%	
1b Relapser	35/36 97%	85%, 100%	

[1]Clopper Pearson exact confidence interval

The results in the treatment experienced population of Study M13-098 also showed highly significant results for 3DAA+RBV compared to the historical control with lower boundaries of the 95% CI much larger than the thresholds for superiority (93% vs. 70% for the full population). Consistent results are seen regardless of HCV genotype or type of treatment experience with SVR_{12} rates ranging from 94% to 100% across the different subgroups.

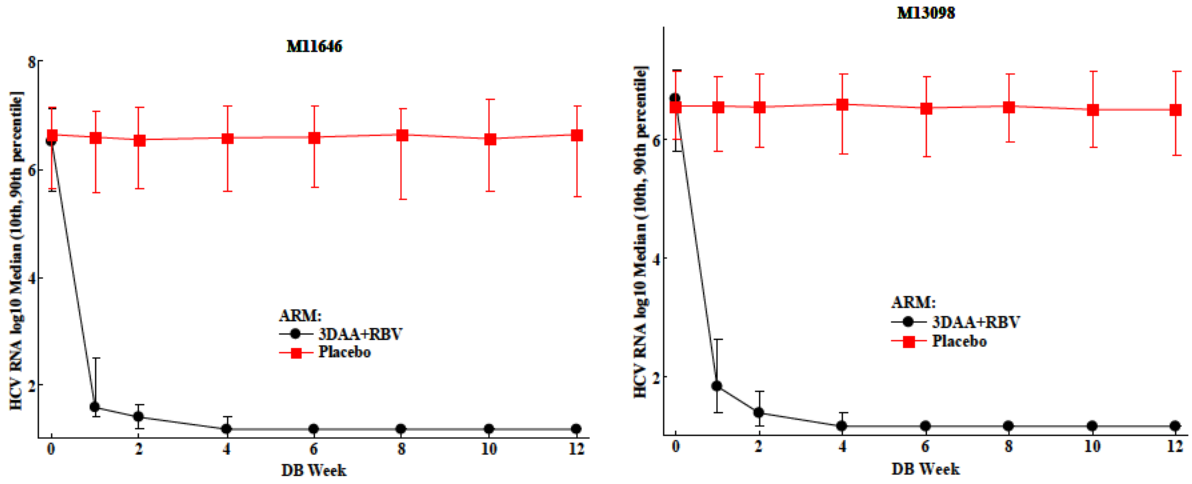
Only one patient (Study M11-646) had a virologic failure (Table 3.1.3.2.6). Eight patients in each study relapsed; 6 of the 16 relapses occurred at post-treatment week 8 or later.

Table 3.1.3.2.6 Reasons for not achieving SVR12 by study for all GT1patients

	3DAA+RBV		
	On Trt Virologic Failures	Relapses	Lost-to-follow-up
M11-646 GT 1 TN	1/473 0.2% TW12 1a	8/473 1.7% 4:PTW2 3:PTW8 1:PTW12 (6:1a)	8/473 1.7% 3:TW1 2:TW2 1:PTW8 2:PTW8w/SVR$_8$
M13-098 GT 1 TE	0	8/297 3% 2:PTW2 4:PTW4 2:PTW8 (6:1a) One	3/297 1% 1:TW1 1:TW2 1:TW8

A plot of the HCV RNA (log10) data shows that most patients respond within a week of treatment. As expected, there is no change in HCV RNA over time in the placebo group.

Figure 3.1.3.2.2 Median HCV RNA data in the double blind period by study and treatment group



ALT normalization was named as the most important secondary efficacy endpoint. The percentage of patients who had a baseline ALT above the upper limit of normal (ULN) at baseline and reached normalization at the end of double blind treatment was computed. This reviewer also computed and summarized ALT results for all randomized patients (Table 3.1.3.2.7).

More than three quarters of the randomized patients in both studies had baseline ALT above ULN (Table 3.1.3.2.7). For these patients, in both studies, the treatment difference in normalization rates was highly significant ($p<0.0001$) with treatment differences over 80% as shown here:

Study M11-646 +82% (95% CI +75%, +89%)
Study M13-098 +84% (95% CI +76%, +92%)

In all patients treated with 3DAA+RBV, ALT levels dropped by about 50 U/L at the end of the double blind period. For patients with ALT above ULN at baseline, the mean drop was slightly great by about 10 U/L. The data clearly illustrates the dramatic effect on ALT levels due to 3DAA+RBV treatment.

Table 3.1.3.2.7 ALT results for all patients and for the subset of patients with baseline ALT>ULN computed by the reviewer using dataset ADUSALT

	M11-646		M13-098	
	GT 1 TN		GT 1 TE	
	3DAA+RBV	Placebo	3DAA+RBV	Placebo
For all pts	N=469[1]	N=158	N=296[2]	N=96
Baseline ALT				
Mean (SD)	68.5 (43)	63 (39)	64.5 (37)	65 (38)
Median	57	52	54	55.5
Change at EOT				
Mean (SD)	-51 (43)	-4 (19)	-46 (44)	-0.9 (25)
Median	-51	-3	-38	0
% w/Bsl ALT>ULN	360/469 77%	114/158 72%	224/296 76%	77/96 80%
% w/EOT ALT>ULN	9/469 2%	103/158 65%	7/296 2%	74/96 77%
For pts w/baseline ALT>ULN	N=360	N=114	N=224	N=77
Baseline ALT				
Mean (SD)	80 (42)	76 (39)	75 (36)	73 (39)
Median	66.5	66.5	63	59
Change at EOT				
Mean (SD)	-62 (41)	-7 (22)	-54 (47)	-3.5 (25)
Median	-48	-6	-45	-3
%w/ALT≤ ULN EOT	**352/360 98%**	**18/114 16%**	**217/224 97%**	**10/77 13%**

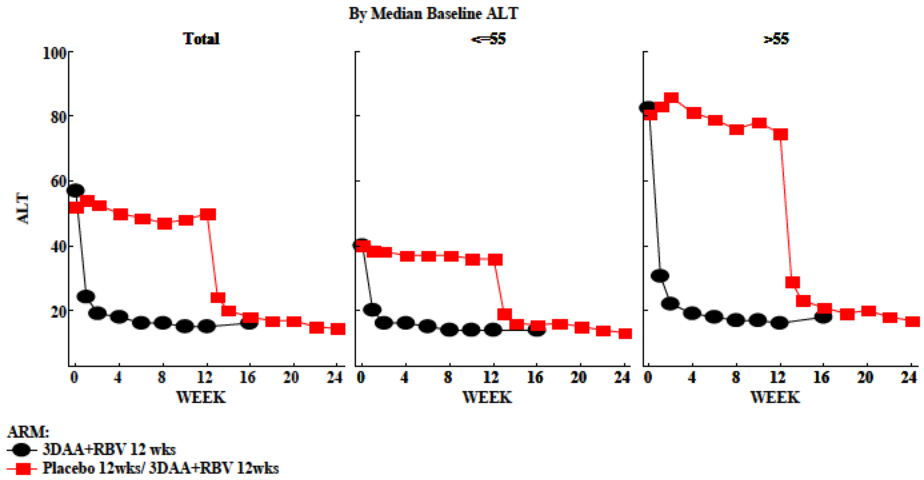
EOT=end of treatment

[1] 4 3DAA+RBV pts had no baseline ALT data; all discontinued study early

[2] One 3DAA+RBV and one placebo had no baseline ALT data

A graph of median ALT by week for Study M11-646 (results for M13-098 look essentially the same) illustrates that the median ALT drops below the ULN within a week due to 3DAA+RBV treatment. This drop is also shown for placebo patients switched to 3 DAA+RBV. The magnitude of the decrease is baseline related as illustrated.

Figure 3.1.3.2.3 Median ALT (U/L) by week on study and treatment for Study M11-646 GT1 TN



One of the main objectives of the two placebo-controlled trials, Studies M11-646 and M13-098, was to compare adverse event rates during the double blind treatment period. This reviewer has computed the incidence of selected adverse events ordered by the nominal p-value for comparing the event rates. The studies were combined after checking results for the individual studies and finding the results to be similar. Also the designs for the two studies are the same including using the same randomization ratio. For more details and a discussion of the safety outcomes, see the review of the FDA clinical reviewer, Russell Fleischer.

Table 3.1.3.2.8 Studies M11-646 and M13-098 adverse event rates for the 12-week double blind period based on Meddra terms reported in the tabulation dataset AE

	3DAA+RBV N=770	Placebo N=255	Treatment Difference (95%CI)	*p-value*
Pruritus	119 (15%)	11 (4%)	+11% (8%, 15%)	*<0.0001*
Anemia	39 (5%)	0	+5% (4%, 6%)	*<0.0001*
Asthenia	102 (13%)	16 (6%)	+7% (3%, 11%)	*0.002*
Dry skin	49 (6%)	4 (2%)	+5% (2%, 7%)	0.002
Decreased appetite	56 (7%)	7 (3%)	+4.5% (2%, 7%)	0.007
Hemoglobin decreased	17 (2%)	0	+2% (1%, 3%)	0.01
Insomnia	101 (13%)	19 (7%)	+6% (2%, 10%)	0.013
Nausea	169 (22%)	38 (15%)	+7% (2%, 12%)	0.015
Fatigue	259 (34%)	66 (26%)	+8% (+1%, +14%)	0.024
Vomiting	43 (6%)	6 (2%)	+3% (+1%, +6%)	0.041
Dizziness	62 (8%)	11 (4%)	+4% (+1%, +7%)	0.049
Dyspnea	73 (9%)	14 (5%)	+4% (+0.5%, +7%)	0.052
Rash	72 (9%)	15 (6%)	+3% (-0.1%, +7%)	0.09
Upper Abdom. Pain	41 (5%)	11 (4%)	+1% (-2%, +4%)	≥0.3
Pyrexia	18 (2%)	5 (2%)	+0.4% (-1.6%, +2.4%)	≥0.3
Hyperbilirubinaemia	3 (0.4%)	0	NC	NC
Blood bilirubin inc	1 (<1%)	0	NC	NC
Any Severe AE	24 (3%)	1 (0.3%)		

NC=not computed

Summary

With two placebo-controlled studies (one in treatment naïve GT1 patients and one in treatment-experienced GT1 patients), the effectiveness of 3DAA+RBV in decreasing ALT to normal levels was demonstrated. About 98% of patients with a baseline ALT greater than ULN had an ALT below ULN at the end of the study. The drop in ALT occurred quickly within the first week of treatment and the magnitude was baseline related with larger drops seen with larger baselines.

The SVR_{12} findings are consistent with the results for the 3DAA+RBV regimen seen in other studies in this application with lower bounds of the confidence intervals about the SVR_{12} estimate 10% or more greater than the superiority threshold.

One of the main objectives of the trial was to demonstrate the safety of 3DAA+RBV against placebo during the 12-week treatment period. There were several adverse events for which a nominally significant treatment difference was seen; most of the adverse events were ones generally associated with RBV (e.g. anemia, asthenia, hemoglobin decreases). This safety data is fully discussed by the FDA clinical reviewer.

3.1.3.3 Study comparing duration of treatment for 3DAA+RBV in cirrhotic patients

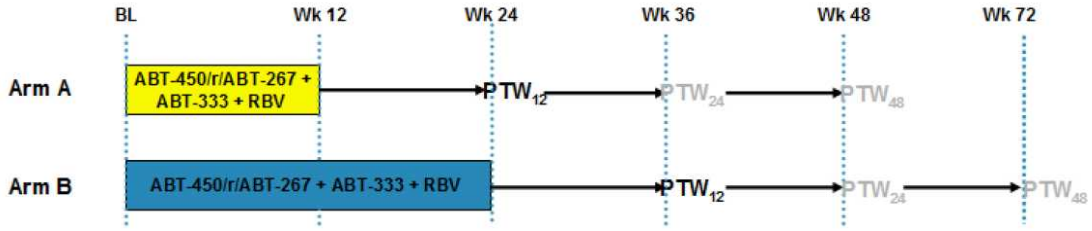
Study M13-099

(conducted 3/14/2013 to 12/12/2013)

Design

Study M13-099 was a Phase 3, randomized, open-label, multicenter study of HCV GT 1, treatment naïve or treatment experienced patients with compensated cirrhosis[3]. Patients were randomized stratified by treatment-experience type and HCV genotype or treatment-naïve and HCV subgenotype (1a versus non-1a) and IL28B genotype (CC versus non-CC) to either 12 weeks of treatment with 3DAA+RBV or to 24 weeks of 3DAA+RBV. At the start of the study, the first 200 subjects were randomized in a 3:5 ratio to the 12- and 24-week arms. After the first 200 subjects were enrolled, the remaining subjects were randomized in a 3:1 ratio to the 12- and 24-week arms. A total of 380 patients were planned to achieve a power greater than 90% assuming an SVR_{12} rate of 68% in each arm. All patients receiving at least one dose were to be followed for an additional 48 weeks after their last dose (Figure 3.1.3.3.1). Patients who had an on-treatment virologic failure or a relapse could be offered alternative treatment.

Figure 3.1.3.3.1 Applicant's schematic of the study design[4]



All study drugs were to be given with food. Compliance was monitored using a MEMS cap.

HCV RNA samples were collected at Weeks 1, 2, 4, 6, 8 and 12 and at additional weeks of 16, 20 and 24 for the 24-week arm. Also samples were collected at post-treatment weeks 2, 4, 8, 12, 24, 36 and 48 or last visit.

The primary objective of the study was to compare the 12-week regimen SVR_{12} rate to the historical SVR_{12} rate of telaprevir plus pegIFN and RBV using a 2-sided 97.5% confidence interval and a non-inferiority margin of 43% and a superiority margin of 54%. A secondary objective was to compare the SVR_{12} rates between the12- and 24-week treatment arms. The focus of this review will be the secondary objective because the primary objective was easily achieved with SVR_{12} rates over 90% seen in each arm. Therefore the primary question addressed in this review is whether 24 weeks of treatment is better than 12 weeks of treatment for all cirrhotic patients. The trial is not powered to show treatment differences by subgroups, so several subgroup analyses are performed to determine whether the treatment difference for the two arms is consistent across subgroups. The primary analysis was planned when all patients had completed at least 12 weeks post treatment or discontinued early.

[3] *Compensated cirrhosis* means that the body still functions fairly well despite damage to the liver.

[4] In error, the applicant's schematic labels Wk 48 as PTW_{48} whereas it should be labeled as PTW_{36}

Patient Disposition

A total of 671 subjects were screened and 381 were randomized (209 on the 12 week regimen and 172 on the 24 week regimen). The primary reason patients were not randomized was they did not satisfy the entry criteria. The first 200 patients were randomized over about a 4 month period in a 3:5 ratio (Wk 12 arm to Wk 24 arm); the ratio was changed to 3:1 and the remaining 181 patients were randomized quickly within less than a month. The goal was to have approximately equal numbers in each group but to allow sufficient time for patients in the 24-week arm to complete treatment and the 12 week follow-up.

As would be expected considering the trial is ongoing, more post-treatment data is available for the 12-week arm compared to the 24-week arm with 94% of 12-week arm patients with data at post-treatment Week 24 data compared to only 12% for the 24-week arm. Therefore SVR_{24} cannot be compared for the two arms.

In each arm, only 5 patients discontinued study early. About twice as many patients discontinued treatment on the 24-week arm (5%) than the 12-week arm (2%). Five of the Week-24 arm patients discontinued after treatment week 12; of those 5, 3 achieved an SVR_{12}.

Of the 4 Wk-12 arm patients who discontinued study drug, 2 discontinued the study early; of the 9 Wk-24 arm patients who discontinued study drug, 4 discontinued the study early. One patient being treated with the 12 week regimen died on Day 93 for reasons unlikely to be drug related.

Table 3.1.3.3.1 Study M13-099 Patient disposition[1]

	3DAA+RBV Trt 12 weeks	3DAA+RBV Trt 24 weeks
Randomized	209	172
Randomized and treated (ITT)	208 (100%)	172 (100%)
Discontinued study drug early	4 (2%)	9 (5%)
Virologic failure	0	3 (2%)
Subject non-compliant	0	1 (<1%)
ADE	4 (2%)	4 (2%)
Incarcerated	0	1 (<1%)
Day discontinued study drug		
Range	Day 10-32	Day 13-156
# (%) discontinued after Wk 12	NA	5 (3%)
ADE days	10, 11,18, 32	13, 141, 154, 156
Discontinued study	5 (2%)	5 (3%)
ADE	2 (1%)	0 (0%)
Lost-to-follow-up	0	3 (2%)
Withdrew consent/non-compliant	2 (1%)	1 (<1%)
Virologic failure	1 (<1%)	1 (<1%)
Weeks completed on study		
8	206 (99%)	172 (100%)
12	206 (99%)	170 (99%)
18	206 (99%)	169 (98%)
22	206 (99%)	169 (98%)
24 * (PTW12 for 12 week arm)	206 (99%)	169 (98%)
26	197 (95%)	169 (98%)
30	195 (94%)	169 (98%)
36 ** (PTW24 12-wk arm; PTW12 24-wk arm)	195 (94%)	168 (98%)
48 **(PTW24 for 24 week arm)	10 (5%)	20 (12%)

[1] Numbers are based on the ADEFFOUT dataset provided by the applicant

Baseline Demographics

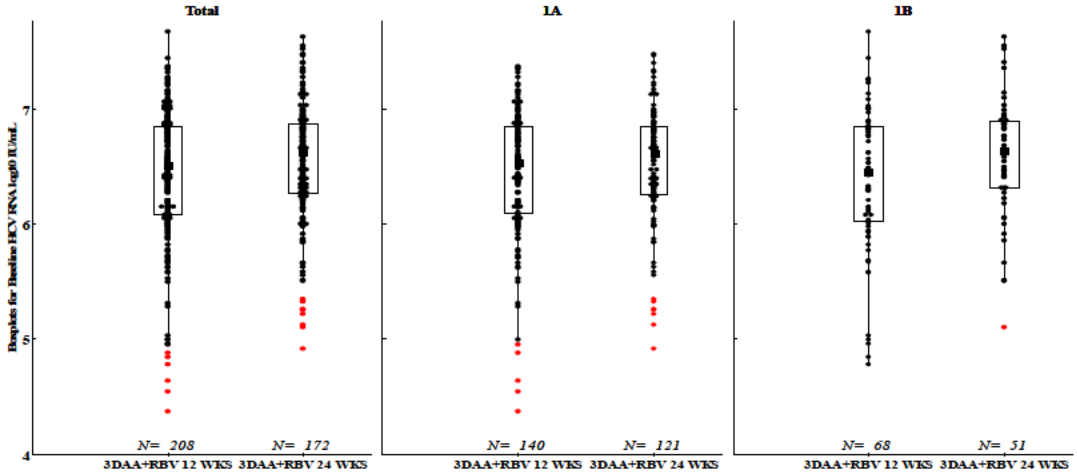
The treatment groups were comparable at baseline as shown in the table below. About half the patients were enrolled in North America and half in Europe. The majority of patients were male (70%) and white (95%). The mean age was 57 with 13% of patients 65 or older. About 2/3 were GT 1a. About 42% of the population was naïve to treatment and about 36% were treatment-experienced null responders.

Table 3.1.3.3.2 Study M13-099 Baseline demographics by treatment arm

	3DAA+RBV Trt 12 weeks (n=208)	3DAA+RBV Trt 24 weeks (n=172)
% USA	41%	45%
% Canada	9%	9%
% Europe	50%	46%
% Female	30%	30%
% White	96%	94%
Age		
Mean (SD)	57 (7)	57 (8)
Min-Max	25-71	21-71
% ≥ 65 years	13%	13%
HCV genotype		
1a	67%	70%
1b	33%	30%
IL28B		
CC	17%	20%
CT	64%	61%
TT	20%	19%
Treatment experience		
Naïve	41%	43%
Experienced	59%	57%
Null responder	36%	36%
Partial responder	9%	8%
Relapser	14%	13%
Baseline HCV RNA Log_{10} IU/mL		
Mean (SD)	6.4 (0.6)	6.5 (0.5)
Median	6.5	6.6
Baseline Child Pugh score		
5	82%	81%
6	18%	16%
>6	0%	3%

The applicant's study report suggested a difference in baseline HCV RNA by reporting a p-value for a comparison of the arms of 0.04 (page 163 in applicant's report) but this reviewer found no difference at baseline for the two arms. The distributions for baseline HCV RNA are shown on the following page (Figure 3.1.3.3.2). There is a slight difference in the lower tails of the distributions but this difference is not notable and does not suggest any imbalance between the arms at baseline. Also the treatment arms are balanced for baseline HCV RNA for both 1a and 1b populations.

Figure 3.1.3.3.2 Boxplots of baseline HCV RNA by treatment arm and genotype 1a/1b



Statistical Methods

See Section 3.1.1 of this review for a general description of the statistical methods used by the applicant and this reviewer. Methods specific to this study are described here.

According to the protocol, patients randomized to the Week 12 arm could have their treatment extended to 24 weeks and patients discontinuing treatment early could have been offered add-on pegIFN/RBV treatment; neither of these situations that could potentially bias the interpretation of the outcomes occurred.

The applicant performed analyses to demonstrate the efficacy of each treatment arm against the historical control. A 2-sided 97.5% confidence interval was computed for the proportion of patients achieving an SVR_{12} in each arm and the lower bound must exceed 43% to demonstrate non-inferiority and must exceed 54% to demonstrate superiority to the active control of telaprevir plus pegIFN and RBV therapy. The confidence intervals were computed stratifying on the randomization stratifiers. The difference in SVR_{12} rates between treatment arms was analyzed using a logistic regression model and using the stratum-adjusted Mantel-Haenszel (MH) test with a continuity correction by the applicant. The latter approach was used by this reviewer.

Efficacy Results

Confidence intervals for SVR_{12} by arm and HCV genotype all excluded 54% (the superiority threshold) demonstrating superiority to the historical control (Table 3.1.3.3.3) .

Table 3.1.3.3.3 Study M13-099 SVR12 rates and 97.5% CI

	3DAA+RBV 12 weeks	3DAA+RBV 24 weeks
1	191/208 92% (87%, 96%)	166/172 96.5% (92%, 99%)
1a	124/140 89% (81%, 94%)	115/121 95% (89%, 98%)
1b	67/68 98.5% (91%, 100%)	51/51 100% (92%, 100%)

To address whether 24 weeks of treatment is more effective than 12 weeks of treatment with 3DAA plus RBV, the two SVR_{12} proportions are subtracted such that a positive value for the difference indicates a higher SVR rate for the 24-week arm over the 12-week arm. Looking at all the patients in this trial, a treatment difference of about 5% was observed (Table 3.1.3.3.4) and this difference was borderline

significant ($p<0.055$) and there was no statistically significant measure of heterogeneity across the two genotypes of 1a and 1b. However the results for 1a patients (+6% with CI -1% to 13%) suggest slightly more favorable results for 24 weeks over 12 weeks of treatment than for 1b patients (+1% with CI -6% to +8%). It seems that 12 weeks of treatment is sufficient for 1b cirrhotic patients with only one patient (a treatment experienced partial responder in the 12-week arm) not achieving an SVR_{12} (relapsed at PTW4).

Table 3.1.3.3.4 Study M13-099 SVR_{12} treatment differences by GT and treatment experience.

	3DAA+RBV 12 weeks (N=208)	3DAA+RBV 24 weeks (N=172)	Treatment Difference (95%CI) 24-wk arm minus 12-wk arm
Overall (1a & 1b)	191/208 92%	166/172 96%	**+5% (+0.07%, +10%)[1]**
1a	124/140 89%	115/121 95%	**+6 % (-0.6%, +13%)**
TN	59/64 92%	53/56 95%	**+2.5% (-7%, +12%)**
CC	19/19 100%	15/16 94%	-6% (-31%, +12%)
Non-CC	40/45 89%	38/40 95%	+6% (-5%, +17%)
TE	65/76 85.5%	62/65 95%	**+9.7% (-0.4%, +20%)**
Null	40/50 80%	39/42 93%	+13% (-1.8%, +28%)
Partial	11/11 100%	10/10 100%	0% (-29%, +28%)
Relapser	14/15 93%	13/13 100%	+7% (-19%, +31%)
1b	67/68 98.5%	51/51 100%	**+1% (-6%, +8%)**

[1]MH for the risk difference stratified by genotypes and treatment experience. P=0.055

Looking at 1a patients by treatment experience, it appears that treatment experienced patients (~+10%) benefit from longer treatment more than treatment naïve patients (~+2.5%); however the confidence intervals for these subgroups are overlapping and a test of homogeneity does not suggest a difference in treatment effects in these two subgroups ($p>0.2$). Furthermore, results for subgroups within these subgroups (CC and non-CC for TN patients and previous outcome type for TE patients) do not suggest heterogeneity among the results with wide overlapping confidence intervals for the treatment differences. In general, it appears that the 1a population may benefit from longer treatment if no safety issues outweigh the small benefit gained with an additional 12 weeks of therapy.

The reasons for not achieving SVR_{12} (Table 3.1.3.3.5) support the advantage of 24 weeks of treatment over 12 weeks of treatment in the 1a population with the difference primarily driven by patients who relapse on the 12-week treatment arm (4%). The 4 virologic failures occur early so length of treatment is not a factor in their outcomes.

Table 3.1.3.3.5 Reasons for not achieving SVR12 in the 1a population

	3DAA+RBV 12 weeks (n=140)			3 DAA+RBV 24 weeks (n=121)		
	On trt Virologic failures	Relapses	Lost-to-follow-up	On trt Virologic failures	Relapses	Lost-to-follow-up
TN n Comments		5 Non-CC at PTW 2-4 One pt w/~4wks of trt			2 All at PTW4 One pt w/~1wk of trt	1 Non-CC at TW 8w/RNA<LLOQ
TE n Comments	1 Previous relapser at TW10	8 Null responders at PTW 2-8 One pt w/~4wks of trt	2 TW<4 with no response	3 Null responders at TW 4-10	0	0

Subgroups

This reviewer looked at subgroups based on baseline HCV RNA, sex, age and geographic region for the 1a population. Analysis by race was not done because there were insufficient numbers of non-white randomized patients (~5%). Only the 1a population was analyzed because most subgroups in the 1b populations had a 100% response in both arms with only one 1b patient not achieving an SVR_{12}. Results for subgroup analyses for the TN 1a and TE 1a populations are shown in Appendix 6.1.

Table 3.1.3.3.6 Study 13-099 Subgroup results for the 1a population (stratified on treatment naïve and exp)

	3DAA+RBV 12 weeks	3DAA+RBV 24 weeks	24 wks-12 wks Trt diff (95%CI)	p-value for Int.[1]
Overall 1a	124/140 89%	115/121 95%	**+6 % (-0.6%, +13%)**	
Baseline HCV RNA				
By median				
< 3,630,000	66/74 89%	52/56 93%	+4% (-6%, +13%)	
≥ 3,630,000	58/66 88%	63/65 97%	+10% (-0.7%, +20%)	>0.3
By tertiles				
<2,120,000	45/48 90%	36/38 95%	+5% (-6%, +16%)	
2,120,000-5,690,000	43/47 91%	39/41 95%	+4% (-7%, +14%)	>0.3
>5,690,000	38/45 84%	40/42 95%	+11% (-2%, +23%)	
Sex				
Male	89/101 (88%)	84/89 (94%)	+6% (-2%, +14%)	
Female	35/39 (90%)	31/32 (97%)	+7% (-4%, +18%)	>0.3
Median Age years				
<57	67/74 (91%)	53/56 (95%)	+4% (-4%, +13%)	
≥ 57	57/66 (86%)	62/65 (95%)	+9% (-1%, +19%)	>0.3
Median BMI kg/m^2				
<28	58/67 (87%)	65/67 (97%)	+10% (+1%, +20%)	
≥ 28	66/73 (90%)	50/54 (93%)	+2% (-8%, +12%)	>0.3
Geographic area				
US & Canada	76/86 (88%)	70/76 (92%)	+4% (-5%, +13%)	
Europe	48/54 (89%)	45/45 (100%)	+11% (+3%, +19%)	**0.13**
Site 44318	5/10 (50%)	7/7 (100%)	+50% (+3%, +82%)	
w/o Site 44318	119/130 (92%)	108/114 (95%)	+3% (-3%, +9.5%)	0.21
Randomization date				
Before or on 4/8/13[2]	47/56 (84%)	88/92 (96%)	+12% (+1%, +22%)	
After 4/8/13	77/84 (92%)	27/29 (93%)	+1% (-10%, +12%)	0.23

[1] Zelen's exact test of homogeneity was used to test for an interaction; a large p-value indicates homogeneity across the subgroup effects.

[2] 4/8/2013 is the date the 200th patient was randomized. The randomization scheme (Wk 12 arm to Wk 24 arm) was to change from 3:5 to 3:1 after 200 patients enrolled.

In most 1a subgroups, more responders were seen on the Week 24 arm compared to the Week 12 arm. The treatment by subgroup interactions were generally non-significant. The exceptions were subgroups defined by geographic region and subgroups defined by time of randomization (before and after the time the randomization ratio was changed). In the US and Canada, a small difference is seen between the arms while in Europe the treatment difference of 11% in favor of longer treatment is significant based on a nominal p-value of 0.03. It is worth noting that the mean BMI for US & Canada is 29 kg/m^2 while the BMI for Europe is 27 kg/m^2. So as was seen when looking at subgroups by BMI, results by region also

show that a lower BMI is associated with a larger difference between the two regimens in favor of the longer treatment.

Looking at the analysis by date of randomization, a +12% treatment effect is seen for the first 200 patients and +1% for the patients randomized after 4/8/2013. This reviewer checked the demographics for the latter two groups of patients and saw no differences that would suggest that the type of patients enrolled changed or that there were any imbalances by arm (see Appendix 6.2). A look at the data by both region and randomization date shows that the difference in treatment effect seen for the by-date analysis is due wholly to the differing US/Canada results before and after 4/8/2013. This reviewer was unable to discern any reason for this difference.

US & Canada: Before or on 4/8/13 +9% (-4%, +22%) After 4/8/13 -9% (-29%, +11%)
Europe Before or on 4/8/13 +18% (-0.1%, +36%) After 4/8/13 +8% (-0.1%, +17%)

The results of the largest site in the study, French site 44318, were very positive (+50% with CI +3% to +82%) such that excluding that site resulted in a treatment difference of 3% (p>0.3) favoring the 12-week arm. The interaction p-value of 0.21 suggests this could be a chance finding; however, it does point to the weakness of the data in favor of the Week 24 arm. This site was inspected by FDA and no major problems were identified.

Adverse Events

To examine further whether the 24 weeks of treatment is preferable for the 1a population, this reviewer looked at the adverse event data for both treatment groups focusing on adverse events examined by the FDA clinical reviewer and also listing those where a significant increase in risk was observed.

Fatigue was the most common adverse event where a significant difference was seen between the groups (Table 3.1.3.3.6). There appears to be no important increases in events that have been associated with RBV or the 3DAA. This safety data suggests no increased risk to 1a patients receiving 24 weeks of therapy versus 12 weeks of therapy in this trial.

Table 3.1.3.3.6 Selected adverse event rates for 1 a population only based on Meddra preferred terms as reported

	3DAA+RBV 12 weeks N=140	3DAA+RBV 24 weeks N=121	Treatment Difference (95%CI) 24-wk arm minus 12-wk arm	*p-value*
Fatigue	47 (34%)	61 (50%)	+17% (+5%, +29%)	0.01
Upper Abdom. Pain	3 (2%)	11 (9%)	+7% (+1%, +13%)	0.02
Dyspnea	9 (6%)	13 (11%)	+4% (-3%, +11%)	≥0.3
Irritability	12 (9%)	15 (12%)	+4% (-4%, 11%)	≥0.3
Vomiting	6 (4%)	7 (6%)	+2% (-4%, +7%)	≥0.3
Anemia	8 (6%)	10 (8%)	+2.5% (-4%, +9%)	≥0.3
Rash	14 (10%)	16 (13%)	+3% (-5%, +11%)	≥0.3
Hyperbilirubinaemia	5 (4%)	2 (2%)	-2% (-6%, +2%)	≥0.3
Blood bilirubin inc	11 (8%)	5 (4%)	-4% (-9%, +2%)	≥0.3

Summary

The main question that may be addressed with this study is whether 24 weeks of treatment with 3DAA+RBV for patients with compensated cirrhosis is more effective than 12 weeks of treatment. For patients with GT 1b, there is no advantage to longer treatment with all but one patient achieving SVR_{12} with 12 weeks of treatment.

For the 1a population, the answer is not completely clear. Overall the 24 weeks of treatment seems preferable although the difference between the arms is not statistically significant (+6% with 95% CI -0.6% to +13%). The observations that support treating the 1a cirrhotic population with 24 weeks of treatment are:

- Generally positive results in favor of 24 weeks were seen for patients who might be considered more difficult to treat, such as patients with failed treatment experience, males, older patients, and patients with high HCV RNA at presentation.
- With the exception of fatigue, there does not appear to be a higher risk of side effects with treatment for 24 weeks compared to 12 weeks.
- Relapse was the primary reason for not achieving an SVR_{12} in the 12-week arm.

The data from the European sites was more convincingly in favor of 24 weeks of treatment than the US/Canada data. One French site (the largest site in the study) showed results overwhelmingly in favor of 24 weeks of treatment (+50% treatment difference) and lower BMI was seen for the European sites with low BMI seemingly associated with larger treatment effects for the Week 24 regimen.

(b) (4) The null responders were a group that showed the largest treatment effect in favor of 24 weeks of treatment (+13% [-2%, +28%]) but the effect was not statistically significant. Also an interaction test comparing the treatment effect for the GT1a null responders to the rest of the 1a population showed no evidence of heterogeneity with p=0.65. (b) (4)

Overall with the data in this study alone, it is not possible to definitively select, based on statistical evidence (i.e. significant interactions and significant differences within subgroups), which types of 1a patients should receive 12 weeks of treatment and which should receive 24 weeks of treatment. Therefore, treating all 1a patients with 24 weeks of treatment may be most beneficial to patients.

3.2 Evaluation of Safety

For a full review of the safety data, see the FDA clinical review by Russell Fleischer.

Safety results were provided in earlier sections of this review to support the recommendation of a treatment regimen for a specific patient population. The lack of important safety signals between randomized arms suggested that the choice of a regimen could depend primarily on demonstrating more significant efficacy.

In this section adverse events that are proposed for labeling (Section 6.1 of the proposed labeling) are summarized by the treatment regimens proposed for four groups of patients:

Genotype 1a without cirrhosis 3DAA+RBV for 12 weeks
Genotype 1a with cirrhosis 3DAA+RBV for 24 weeks
Genotype 1b without cirrhosis 3DAA for 12 weeks
Genotype 1b with cirrhosis 3DAA+RBV for 12 weeks

This reviewer combined the results of studies of the same design including the same randomization ratio and studies having essentially the same safety results by treatment arm (Table 3.2.1). The safety of the full regimen 3DAA+RBV for the 1a non-cirrhotic population is best demonstrated with the placebo-controlled studies (Studies M11-646 and M13-098). The safety of the 3DAA regimen for the 1b population cannot be assessed directly against placebo within a randomized trial. The 3DAA regimen was tested against the 3DAA plus RBV in Studies M13-389 and M13-961 so that comparison measures the impact of RBV on safety; however, the full regimen of 3DAA+RBV was tested against placebo in 1b non-cirrhotics in Studies M11-646 and M13-098. Assuming the populations across these 4 studies are similar and the results for the two 3DAA+RBV are similar, one could determine whether there may be any notable differences in adverse event rates between 3DAA and placebo. For example, in the 1b non-cirrhotic population, there is significantly more insomnia due to treatment with 3DAA+RBV versus placebo (11% vs. 2%) or versus 3DAA (11% vs. 3%) suggesting that insomnia for the 3DAA arm is similar to placebo.

For the 1a non-cirrhotic population, Table 3.2.1 shows that adverse event rates are higher for the full regimen (3DAA+RBV) for the six selected adverse events with the difference being nominally significantly different for pruritus (itching), asthenia (weak, lacking energy) and anemia. As further support for these results, the results for the 3DAA+RBV arm in 1a patients in Study M14-002 were generally similar (see Table 3.1.3.1.9 on page 30).

Table 3.2.1 Adverse event rates[1] for 1a & 1b non-cirrhotic ITT population during the 12-week treatment period

	1a Non-cirrhotics		1b Non-cirrhotics			
	Studies M11-646 and M13-098		Studies M11-646 and M13-098		Studies M13-389 and M13-961	
Adverse Event	**3DAA+RBV[2] N=495**	Placebo N=162	3DAA+RBV N=275	Placebo N=93	3DAA+RBV N=301	**3DAA[2] N=304**
Fatigue	**183 (37%)**	49 (30%)	76 (28%)	17 (18%)	73 (24%)	**62 (20%)**
Nausea	**115 (23%)**	26 (16%)	54 (20%)	12 (13%)	38 (13%)*	**14 (5%)**
Pruritus	**77 (16%)***	7 (4%)	42 (15%)*	4 (4%)	37 (12%)*	**19 (6%)**
Insomnia	**72 (15%)**	17 (10%)	29 (11%)*	2 (2%)	32 (11%)*	**9 (3%)**
Asthenia	**50 (10%)***	5 (3%)	52 (19%)	11 (12%)	31 (10%)*	**17 (6%)**
Anemia	**24 (5%) ***	0	15 (5%) *	0	23 (7%) *	**1 (<1%)**

[1]Adverse events are defined by Meddra preferred terms (b) (4)
[2]FDA recommended regimen
* Fisher's exact test nominal unadjusted p-value <0.05 for comparisons of randomized arms within study

For the 1b non-cirrhotic population where the 3DAA regimen is recommended, the comparison of the 3DAA to 3DAA+RBV illustrates that the difference in rates is due to RBV (Table 3.2.1). Noting first that the rates for the 3DAA+RBV arms are similar for several adverse events, the 3DAA results generally look similar to the results seen for placebo.

Only one study (M13-099) included patients with cirrhosis. In this study, two durations of treatment were compared; 12 weeks versus 24 weeks. The 24-week regimen is recommended for 1a cirrhotics and the 12-week regimen for 1b cirrhotics. Comparing adverse event rates for these two durations, the only notable difference among the six events listed is for fatigue with more fatigue observed with 24 weeks of treatment. Note that the event rates for cirrhotics are very similar to the ones seen for non-cirrhotics.

Table 3.2.2 Adverse event rates[1] for the 1a and 1b cirrhotic ITT populations during the treatment period

	1a Cirrhotics Study M13-099		1b Cirrhotics Study M13-099	
Adverse Event	3DAA+RBV 12wks N=140	**3DAA+RBV[2] 24 wks N=121**	**3DAA+RBV[2] 12wks N=68**	3DAA+RBV 24 wks N=51
Fatigue	47 (34%)*	**61 (50%)**	**19 (28%)**	18 (35%)
Nausea	27 (19%)	**20 (17%)**	**10 (15%)**	12 (24%)
Pruritus	23 (16%)	**22 (18%)**	**14 (21%)**	11 (22%)
Insomnia	24 (17%)	**23 (19%)**	**5 (7%)**	6 (12%)
Asthenia	14 (10%)	**12 (10%)**	**14 (21%)**	10 (20%)
Anemia	8 (6%)	**10 (8%)**	**7 (10%)**	7 (14%)

[1]Adverse events are defined by Meddra preferred terms (b) (4)
[2]Recommended regimen
* Fisher's exact test p-value <0.05

(b) (4)

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

The results in this section are organized by genotype and by presence of cirrhosis to be consistent with the FDA labeling recommendations for treatment of these groups. The labeling as proposed by FDA recommends the following:

Genotype 1a without cirrhosis 3DAA+RBV for 12 weeks
Genotype 1a with cirrhosis 3DAA+RBV for 24 weeks
Genotype 1b without cirrhosis 3DAA for 12 weeks
Genotype 1b with cirrhosis 3DAA+RBV for 12 weeks

The SVR_{12} rates in the 1b population are very high with very few patients not achieving an SVR as summarized in Table 4.1. With only a total of 5 patients not achieving an SVR, the results by subgroup will be consistent showing no important treatment differences as is seen for the overall 1b population and interactions will not be measurable. Therefore subgroup results are only shown for the 1a population, not the 1b population. For subgroup analyses of the 1a population, only two studies have randomized arms to test efficacy; Study M14-002 in non-cirrhotics and Study M13-099 in cirrhotics.

Table 4.1 SVR_{12} rates in the 1b population

Studies M13-389 & M13-961 Genotype 1b w/o cirrhosis	3DAA+RBV 298/301 (99%)	3DAA 304/304 (100%)	Trt diff (95%CI) -1% (-2%, +1%)
Study M13-099 Genotype 1b w/ cirrhosis	3DAA+RBV/12 weeks 67/68 (98.5%)	3DAA+RBV/24 weeks 51/51 (100%)	Trt diff (95%CI) +1% (-6%, +8%)

4.1 Gender, Race, Age, and Geographic Region

More than 90% of patients were Caucasian in all 6 studies so there are insufficient numbers to perform analyses by race.

Subgroup results by gender and geographic region show consistent treatment differences (p-value to test interaction greater than 0.3) for both non-cirrhotic and cirrhotic GT1a patients (Table 4.1.1). For non-cirrhotics, older patients (treatment difference of 11%) appear to receive more benefit from the addition of RBV than younger patients (treatment difference of 2%) when using a cutpoint of the median age; when using 60 or 65 years as a cutpoint, the interaction results are not significant with p>0.3.

Table 4.1.1 Subgroup results for genotype 1a patients with and without cirrhosis

Genotype 1a without cirrhosis Study M14-002

	3DAA+RBV	3DAA	Trt diff (95%CI)	p-value for Int.[1]
Overall 1a w/o cirrhosis	97/100 97%	185/205 90%	+6.8% (+1.3%, +12%)	
Gender				
Male	67/70 (96%)	113/129 (88%)	+8% (+0.7%, +16%)	
Female	30/30 (100%)	72/76 (95%)	+5% (+0.2%, +10%)	>0.3
Age (years)				
< Median 54	45/48 (94%)	90/98 (92%)	+2% (-7%, +11%)	
≥ Median 54	52/52 (100%)	95/107 (89%)	+11% (+5%, +17%)	**0.09**
< 65	87/90 (97%)	172/192 (90%)	+7% (+1%, +13%)	
≥ 65	10/10 (100%)	13/13 (100%)	0% (-25%, +33%)	>0.3
Geographic area				
US	78/81 (96%)	147/166 (89%)	+8% (+1%, +14%)	
Other	19/19 (100%)	38/39 (97%)	+3% -2%, +8%)	>0.3

Genotype 1a with cirrhosis Study M13-099

	3DAA+RBV 12 weeks	3DAA+RBV 24 weeks	Trt diff (95%CI) 24 wks-12 wks	p-value for Int.[1]
Overall 1a w/cirrhosis	124/140 89%	115/121 95%	+6 % (-0.6%, +13%)	
Gender				
Male	89/101 (88%)	84/89 (94%)	+6% (-2%, +14%)	
Female	35/39 (90%)	31/32 (97%)	+7% (-4%, +18%)	>0.3
Age (years)				
< Median 57	67/74 (91%)	53/56 (95%)	+4% (-4%, +13%)	
≥ Median 57	57/66 (86%)	62/65 (95%)	+9% (-1%, +19%)	>0.3
< 65	114/129 (88%)	99/105 (94%)	+6% (-1%, +13%)	
≥ 65	10/11 (91%)	16/16 (100%)	+9% (-8%, +26%)	>0.3
Geographic area				
US	63/73 (86%)	60/65 (92%)	+6% (-4%, +16%)	
Other	61/67 (91%)	55/56 (98%)	+7% (-0.5%, +15%)	>0.3

[1]Zelen's exact test for interaction was used to compute the p-values.

4.2 Other Special/Subgroup Populations

Subgroups were defined by baseline HCV RNA, ALT and BMI and by genotypes. There are no significant interactions with the results being consistently favorable to the 3DAA+RBV for 12 weeks for non-cirrhotic patients with GT1a and generally favorable to the 3DAA+RBV for 24 weeks for cirrhotic patients with GT1a.

Table 4.2.1 Subgroup results for genotype 1a patients with and without cirrhosis

Genotype 1a without cirrhosis Study M14-002 (all TN)

	3DAA+RBV	3DAA	Trt diff (95%CI) 3DAA+RBV – 3DAA	p-value for Int.[1]
Overall 1a w/o cirrhosis	97/100 97%	185/205 90%	+6.8% (+1.3%, +12%)	
Bsl Median HCV RNA				
< 5,080,000	45/47 96%	94/104 89%	+6% (-2%, +15%)	
≥ 5,080,000	52/53 98%	92/101 91%	+7% (+0.4%, +14%)	>0.3
Subgenotype				
CC	31/31 100%	61/63 97%	+3% (-1%, +8%)	
Non-CC	66/69 96%	124/142 87%	+8% (+1%, +16%)	>0.3
Bsl Median ALT				
< 57 U/L	48/48 (100%)	95/104 (91%)	+9% (+3%, +14%)	
≥ 57 U/L	49/52 (94%)	90/101 (89%)	+5% (-4%, +14%)	0.27
Bsl Median BMI kg/m^2				
<26	39/39 (100%)	98/106 (92%)	+8% (+2%%, +13%)	
≥ 26	58/61 (95%)	87/99 (88%)	+4% (-1%, +16%)	>0.3

Genotype 1a with cirrhosis Study M13-099

	3DAA+RBV 12 weeks	3DAA+RBV 24 weeks	Trt diff (95%CI) 24 wks - 12 wks	p-value for Int.[1]
Overall 1a w/cirrhosis	124/140 89%	115/121 95%	+6 % (-0.6%, +13%)	
Bsl Median HCV RNA				
< 3,630,000	66/74 89%	52/56 93%	+4% (-6%, +13%)	
≥ 3,630,000	58/66 88%	63/65 97%	+10% (-0.7%, +20%)	>0.3
Treatment experience				
Naïve	59/64 92%	53/56 95%	+3% (-7%, +12%)	
Experienced	65/76 85.5%	62/65 95%	+10% (-0.4%, +20%)	>0.3
Subgenotype				
CC	24/25 (96%)	24/25 (96%)	0% (-11%, +11%)	
Non-CC	100/115 (87%)	91/96 (95%)	+8% (+0.2%, +15%)	>0.3
Trt exp. patients				
Null	40/50 80%	39/42 93%	+13% (-1.8%, +28%)	
Partial	11/11 100%	10/10 100%	0% (-29%, +28%)	
Relapser	14/15 93%	13/13 100%	+7% (-19%, +31%)	>0.3
Bsl Median ALT				
< 88 U/L	62/71 (87%)	55/57 (96%)	+9% (+0.1%, +18%)	
≥ 88 U/L	62/69 (90%)	60/64 (94%)	+4% (-5%, +13%)	>0.3
Bsl Median BMI kg/m^2				
<28	58/67 (87%)	65/67 (97%)	+10% (+1%, +20%)	
≥ 28	66/73 (90%)	50/54 (93%)	+2% (-8%, +12%)	>0.3

Bsl=Baseline

[1]Zelen's exact test for interaction was used to compute the p-values.

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

During the review of this application, there were a few statistical issues that arose and they are listed here:

1. All primary efficacy comparisons were defined in the protocol as being a comparison to an historical control. The lack of a randomized or concurrent control may bias the comparison primarily due to not controlling for population characteristics (both those that can be measured and those that cannot). Nevertheless, historical controls allow one to conduct a trial in populations difficult to study affording the possibility of improved treatment opportunities.
2. Only a Phase 2 trial provided data regarding the contribution of the two DAA's (ABT-267 and ABT-333) added to ABT-450. The applicant argued that the full regimen (3DAA) was needed based on in vitro evidence of resistance without all three. Whether the resistance argument is sufficient or whether the combination rule should have been applied more rigorously to this 3DAA is a clinical issue.
3. This reviewer and the applicant used different statistical methods for some analyses (e.g. an exact confidence interval versus an asymptotic confidence interval). These differences had no impact on interpretation of the outcomes.
4. The applicant excluded 6 patients randomized in Study M13-389 because they were not treated with co-formulated product. This reviewer included these patients because they were part of the randomized population and they were all treated and followed as were all other patients in the study. The inclusion of these patients has no notable effect on the outcomes.
5. The applicant (b) (4) This reviewer did not agree that this was appropriate (b) (4) Though the power is affected (b) (4) the results were sufficient from which to draw some conclusions.
6. The randomization ratio was changed in Study M13-099 when about half the patients were enrolled. Given this was an open label study, this type of change could lead to biased patient selection. This reviewer checked patient characteristics before and after the change and saw no measurable difference in the patient populations.
7. For four patients in three Phase 3 trials, SVR_{12} was assessed by imputing Week 12 HCV RNA with post-treatment Week 8 HCV RNA (i.e. carrying forward a value); this imputation approach was planned and described in the protocol. The assumption must be made with these patients that they would not have relapsed after post-treatment Week 8. This reviewer presented results with them counted as having an SVR_{12} as planned in the protocol and also checked to see if counting them as not converted would impact the outcomes in any appreciable way and noticed no appreciable effect.
8. The applicant defined relapse as patients who had or nearly had completed treatment, achieved an HCV RNA<LLOQ and then had an HCV RNA above the LLOQ off treatment. Additionally, this reviewer counted patients who had not completed treatment, achieved an HCV RNA<LLOQ while on treatment and then during follow-up post-treatment had an HCV RNA above the LLOQ as relapsers also. In all cases, only one additional patient was included in the relapse count and not counted as a lost-to-follow-up so the impact on the interpretation of the relapse data is small. The length of treatment for these patients is noted in each table listing the reasons for not achieving SVR_{12}. Clearly the differing definitions for relapse do not impact the estimate of SVR_{12}

9. One site in Study M13-099 had a large impact on the results so this reviewer asked the FDA inspectors to examine the data for specific patients during their planned site visit. Their site inspection revealed no major problems with the site. It is reasonable to assume for this site that the extreme results seen at this site were due to chance.

None of these issues had an impact on the statistical conclusions that were drawn from the data.

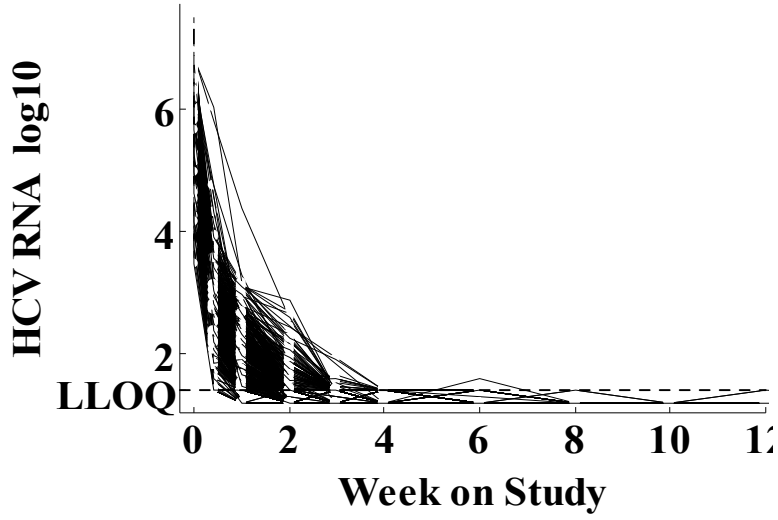
5.2 Collective Evidence

The applicant, AbbVie, has submitted the results of one 14-arm Phase 2 study and 6 Phase 3 studies to support the treatment of HCV in genotype 1 patients with or without cirrhosis (see Table 2.1.1 for a brief description of the design of each study).

A regimen of three direct acting antivirals with or without ribavirin was studied in the Phase 3 trials. This regimen (referred to as 3DAA+RBV in this review) is composed of a fixed dose tablet of ombitasvir [ABT-267 12.5 mg], paritaprevir [ABT-450 75 mg] and ritonavir [ABT-538 50 mg] co-packaged with dasabuvir [ABT-333 250 mg].

The Phase 2 study (M11-652) was designed to study different durations of treatment and also regimens excluding ABT-267, ABT-333 or RBV in naïve patients and in patients who were previous null responders. Figure 5.2.1 illustrates the usual pattern of response seen by patients in both the Phase 2 study and Phase 3 studies with most patients' HCV RNA values falling below the LLOQ within 4 weeks of treatment. In the naïve 1a population of Study M11-652, ABT-267, ABT-333 and RBV were all shown to improve the SVR_{24} rates so this study's results suggested that the full regimen of 3DAA plus RBV is best for 1a naïve patients. Response rates in the naïve 1b population were 100% in arms of the full regimen and when excluding any one of ABT-267, ABT-333 or RBV so there is no evidence in this Phase 2 study that 1b patients need all the drugs in the regimen of 3DAA+RBV. The data for null responders was limited and offered minimal evidence regarding the contribution of the components, dosing or length of treatment.

Figure 5.2.1 HCV RNA plotted over time by patient for 3DAA+RBV in Study M11-652



All Phase 3 trials were ongoing at the time of the submission to FDA with the analyses submitted planned when all patients had completed at least12 weeks post-treatment or discontinued early to assess the primary endpoint SVR_{12}. There was very limited SVR_{24} (less than 10% of patients) with the exception of the 12-week treatment arm in Study M13-099, the study in cirrhotics where about 94% of patients had data to post-treatment Week 24. Both the full regimen of 3DAA+RBV and the regimen without RBV were well tolerated in the Phase 3 trials with only about 2% of patients over all trials discontinuing treatment early.

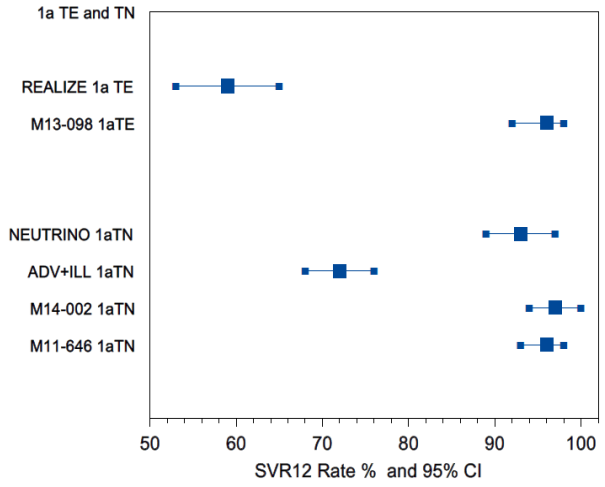
The patient populations across the 6 Phase 3 trials were quite homogenous with the majority being Caucasian (>90%), male (~65%) and under 65 years (>90%). About half the patients were enrolled in US sites.

The primary efficacy comparison in all trials was a comparison of SVR_{12} rates of each randomized arm (3DAA+RBV or 3DAA alone) to an historical control of telaprevir plus pegIFN and RBV. Thresholds were computed by the applicant from the telaprevir plus pegIFN and RBV results of trials REALIZE, ILLUMINATE and ADVANCE (see Section 2.3 of this review for more details). REALIZE was a trial in GT1 treatment-experienced patients and ILLUMINATE and ADVANCE were conducted in GT1 treatment-naïve patients. The results for non-cirrhotics from those historical trials are shown in the two graphs below with the regimens proposed for approval (3DAA+RBV for 1a patients and 3DAA for 1b patients in Studies M13-098, M14-002, M11-646, M13-389 and M13-961). In addition, included on the graphs are the results for NEUTRINO, a study in GT1 treatment-naïve patients treated with Sovaldi, the most recently approved HCV drug as of August 2014.

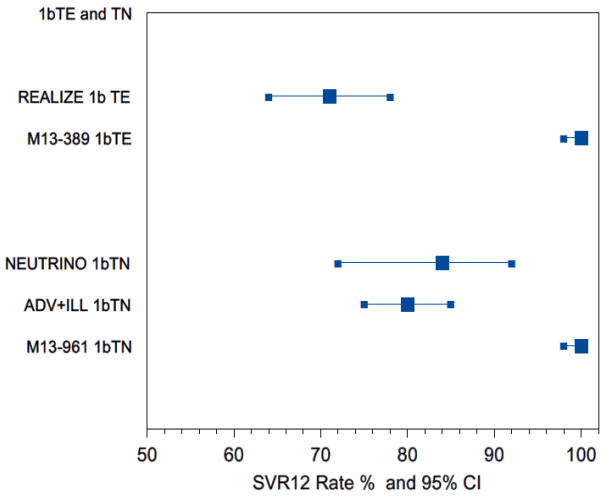
It is clear that the rates observed for the recommended 3DAA+RBV regimen (3 estimates to the far right in the graph on the left) in GT1a patients are superior to those observed for the pre-specified historical control (REALIZE, ILLUMINATE and ADVANCE) as illustrated by the lack of overlap of the confidence intervals for the historical control arms and the 3DAA+RBV arms. The confidence intervals for the rates of the 3DAA+RBV arms overlap with the confidence interval from NEUTRINO suggesting the results for Sovaldi and AbbVie's 3DAA+RBV in a treatment-naive 1a population may be comparable.

Figure 5.2.2 SVR results for non-cirrhotics from AbbVie trials and historical control trials

3DAA+RBV & historical controls SVR In GT1a patients



3DAA & historical controls SVR In GT1b patients

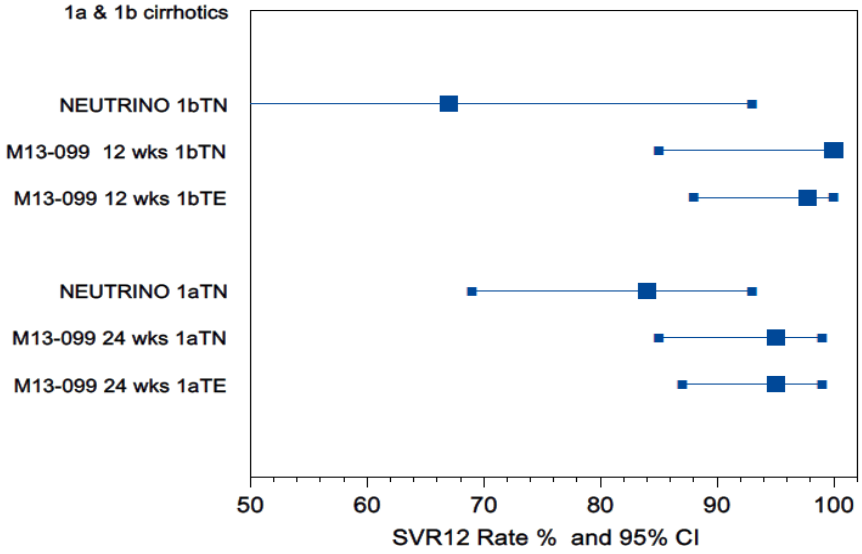


The 3DAA results for the GT1b (TE and TN) population (the graph on the right above) are superior to the results for the historical controls with no overlap of the confidence intervals.

For the cirrhotic population, the lower bounds of the 95% confidence intervals for the SVR rates from Study M13-099 were significantly higher than the superiority thresholds of 54% and 67% based on GT1 cirrhotic patients in REALIZE, ILLUMINATE and ADVANCE (see Table 3.1.3.3.3 for details); these comparisons are not shown in the graph on the following table because historical results by 1a and 1b for cirrhotics were not provided in the application.

Figure 5.2.3 illustrates the results for cirrhotics from Study M13-099 and results from the Sovaldi trial NEUTRINO. The magnitudes of the SVR rates for 1a patients treated with 24 weeks of 3DAA+RBV and for 1b patients treated with 12 weeks of 3DAA+RBV are larger than the rates seen for Sovaldi but the confidence intervals are overlapping suggesting that the AbbVie product may not be more effective than Sovaldi.

Figure 5.2.3 SVR results for GT1a and GT1b cirrhotics from AbbVie trial M13-099 and Sovaldi NEUTRINO trial



The randomized arms in the Phase 3 trials were defined to address questions regarding the choice of a regimen for specific populations. In Studies M13-389, M13-961 and M14-002, non-cirrhotic patients were randomized to either 3DAA+RBV or 3DAA; these designs allowed one to test whether RBV contributed to the efficacy of the 3DAA alone. In Study M13-099, cirrhotic patients were randomized to either 12 weeks of treatment with 3DAA+RBV or 24 weeks of treatment with 3DAA+RBV. These studies suggested different treatment regimens by genotype.

For non-cirrhotic patients with GT1a, a regimen including RBV produced significantly more patients achieving an SVR (increase in SVR_{12} of about 7%, Table 5.2.1). For cirrhotic patients with GT1a, a 24-week regimen was more effective than a 12-week regimen (increase in SVR_{12} of about 6%, Table 5.2.1). From Table 5.2.2 on the following page, it can be seen that primary reason for the difference between arms was a difference in relapse rates.

Table 5.2.1 SVR_{12} rates for the 1a population

M14-002	**3DAA+RBV[1]**	3DAA	Trt diff (95%CI)
GT 1a w/o cirrhosis	**97/100 97% (94%, 100%)**	185/205 90% (86%, 94%)	+6.8% (+1%, +12%)
M13-099	**3DAA+RBV/24 weeks**	3DAA+RBV/12 weeks	Trt diff (95%CI)
GT 1a w/ cirrhosis	**115/121 95% (97%, 100%)**	124/140 89% (81%, 94%)	+6% (-1%, +13%)

[1]The regimens in the labeling proposed by FDA are bolded.

Table 5.2.2 On-treatment virologic failure and relapse rates for the 1a population

M14-002 GT 1a w/o cirrhosis	**3DAA+RBV**	3DAA
On-trt VF	**1/100 (1%)**	6/205 (3%)
Relapses	**1/100 (1%)**	11/205 (5%)
M13-099 GT 1a w/ cirrhosis	**3DAA+RBV/24 weeks**	3DAA+RBV/12 weeks
On-trt VF	**3/121 (2%)**	1/140 (1%)
Relapses	**0/121 (0%)**	13/140 (9%)

For non-cirrhotic patients with GT1b, 3DAA alone was as effective as the regimen including RBV with a tight confidence interval on the treatment difference (difference of -1% with 95% CI of -2% to +1%, Table 5.2.3). For cirrhotic patients with GT1b, a 12-week regimen was as effective as the12-week regimen (difference of +1% with 95% CI of -6% to +8%, Table 5.2.3). In all three studies reported below, there was only one on-treatment virologic failure and one relapse.

Table 5.2.3 SVR_{12} rates for the 1b population

M13-389 & M13-961 GT 1b w/o cirrhosis	3DAA+RBV 298/301 99% (97%, 100%)	**3DAA[1]** **304/304 100% (99%, 100%)**	Trt diff (95%CI) -1% (-2%, +1%)
M13-099 GT 1b w/ cirrhosis	3DAA+RBV/24 weeks 51/51 100% (92%, 100%)	**3DAA+RBV/12 weeks** **67/68 98.5% (91%, 100%)**	Trt diff (95%CI) +1% (-6%, +8%)

[1]The regimens in the labeling proposed by FDA are bolded.

Table 5.2.4 On-treatment virologic failure and relapse rates for the 1a population for the 1b population

M13-389 & M13-961 GT 1b w/o cirrhosis	3DAA+RBV	**3DAA**
On-trt VF	1/301 (<1%)	**0**
Relapses	1/301 (<1%)	**0**
M13-099 GT 1b w/ cirrhosis	3DAA+RBV/24 weeks	**3DAA+RBV/12 weeks**
On-trt VF	0	**0**
Relapses	0	**1/68 (1.5%)**

The data in the Phase 3 trials clearly support the treatment regimens in the labeling proposed by the FDA as follows:

Genotype 1a without cirrhosis 3DAA+RBV for 12 weeks
Genotype 1a with cirrhosis 3DAA+RBV for 24 weeks
Genotype 1b without cirrhosis 3DAA for 12 weeks
Genotype 1b with cirrhosis 3DAA+RBV for 12 weeks

The evidence in favor of the above regimens was consistent across a number of subgroups.

5.3 Conclusions and Recommendations

Overall, the results submitted by AbbVie demonstrate the efficacy of the Viekara pak (3DAA) with or without RBV for the treatment of genotype 1 patients with or without cirrhosis. The SVR rates for the new drug product are larger than pre-specified superiority thresholds which were based on rates for the historical control of telaprevir plus pegIFN and RBV. The choices of regimens based on genotype and the presence or absence of cirrhosis are supported by the results of comparisons of randomized arms in six Phase 3 studies.

6 APPENDICES

6.1 Study 13-099 Baseline Demographics Before and After 4/8/2013

	before 4/8/2013		after 4/8/2013	
	3DAA+RBV Trt 12 weeks N=56	3DAA+RBV Trt 24 weeks N=92	3DAA+RBV Trt 12 weeks N=84	3DAA+RBV Trt 24 weeks N=29
% USA/Canada	70%	64%	56%	59%
% Europe	30%	36%	44%	41%
% Female	25%	24%	30%	34%
Age Mean (SD)	56 (7)	57.5 (7)	56 (6)	53 (9)
IL28B				
CC	10%	20%	23%	24%
CT	59%	64%	64%	55%
TT	16%	16%	13%	21%
Prior treatment experience				
Naïve	46%	46%	45%	48%
Experienced	54%	54%	55%	52%
Null responder	34%	35%	36%	34%
Partial responder	9%	9%	7%	7%
Relapser	11%	11%	11%	10%
Baseline HCV RNA Log_{10} IU/mL Mean (SD)	6.48 (0.6)	6.50 (0.5)	6.37 (0.6)	6.51 (0.6)
Baseline Child Pugh score				
5	80%	74%	83%	90%
6	20%	21%	17%	10%
>6	0%	5%	0%	0%

6.2 Study 13-099 Subgroup SVR_{12} results for the 1a population by TE and TN

Study 13-099 Subgroup results for the 1a TE population

	3DAA+RBV 12 weeks	3DAA+RBV 24 weeks	24 wks-12 wks Trt diff (95%CI)	p-value for Int.[1]
Overall 1a TE	65/76 85.5%	62/65 95%	**+9.7% (-0.4%, +20%)**	Trt diff p=0.09
TE Category				
Null	40/50 80%	39/42 93%	+13% (-2%, +28%)	
Partial resp/relapse	25/26 96%	23/23 100%	+4% (-4%, +11%)	>0.3
Null	40/50 80%	39/42 93%	+13% (-2%, +28%)	
Partial responder	11/11 100%	10/10 100%	0% (-29%, +28%)	
Relapser	14/15 93%	13/13 100%	+7% (-19%, +31%)	
Baseline HCV RNA By median				
< 3,630,000	35/41 85%	28/29 97%	+11% (-2%, +24%)	
≥ 3,630,000	30/35 86%	34/36 94%	+9% (-5%, +23%)	>0.3
Sex				
Male	53/61 (87%)	46/49 (94%)	+7% (-4%, +18%)	
Female	12/15 (80%)	16/16 (100%)	+20% (-0.2%, +40%)	>0.3
Median Age years				
<57	38/43 (88%)	26/27 (96%)	+8% (-4%, +20%)	
≥ 57	27/33 (82%)	36/38 (95%)	+13% (-2%, +28%)	>0.3
Median BMI kg/m^2				
<28	29/34 (85%)	36/37 (97%)	+12% (-1%, +25%)	
≥ 28	36/42 (86%)	26/28 (93%)	+7% (-7%, +21%)	>0.3
Geographic area				
US & Canada	29/34 (85%)	33/36 (92%)	+6% (-8%, +21%)	
Europe	36/42 (86%)	29/29 (100%)	+14% (+4%, +25%)	>0.3
Site 44318	5/10 50%	6/6 100%	+50% (+2%, +81%)	
w/o Site 44318	60/66 91%	56/59 95%	+5% (-6%, +14%)	0.24
Randomization date				
Before or on 4/8/13[2]	25/30 (83%)	47/50 (94%)	+11% (-4%, +26%)	
After 4/8/13	40/46 (87%)	15/15 (100%)	+13% (+3%, +23%)	>0.3

[1] Zelen's exact test of homogeneity was used to test for an interaction; a large p-value indicates homogeneity across the subgroup effects.

[2] 4/8/2013 is the date the 200th patient was randomized. The randomization scheme (Wk 12 arm to Wk 24 arm) was to change from 3:5 to 3:1 after 200 patients enrolled.

Study 13-099 Subgroup results for the 1a TN population

	3DAA+RBV 12 weeks	3DAA+RBV 24 weeks	24 wks-12 wks Trt diff (95%CI)	p-value for Int.[1]
Overall 1a TN	59/64 92%	53/56 95%	**+2.5 % (-7%, +12%)**	
TN Category				
CC	19/19 100%	15/16 94%	-6% (-31%, +12%)	
Non-CC	40/45 89%	38/40 95%	+6% (-5%, +17%)	>0.3
Baseline HCV RNA By median				
< 3,630,000	30/32 94%	24/27 89%	-4% (-19%, +10%)	
≥ 3,630,000	29/32 91%	29/29 100%	+9% (-0.7%, +19%)	0.17
Sex				
Male	36/40 (90%)	38/40 (95%)	+5% (-6%, +16%)	
Female	23/24 (96%)	15/16 (94%)	-2% (-16%, +12%)	>0.3
Median Age years				
<57	29/31 (93.5%)	27/29 (93%)	-0.4% (-13%, +12%)	
≥ 57	30/33 (91%)	26/27 (96%)	+5% (-7%, +18%)	>0.3
Median BMI kg/m^2				
<28	29/33 (88%)	29/30 (97%)	+9% (-4%, +22%)	
≥ 28	30/31 (97%)	24/26 (92%)	-4% (-16%, +8%)	>0.3
Geographic area[2]				
US & Canada	47/52 (90%)	37/40 (92.5%)	+2% (-9%, +14%)	
Europe	12/12 (100%)	16/16 (100%)	0% (-21%, +27%)	>0.3
Randomization date				
Before or on 4/8/13[3]	22/26 (85%)	41/42 (98%)	+13% (-2%, +28%)	
After 4/8/13	37/38 (97%)	12/14 (86%)	-12% (-31%, +7%)	0.04

[1] Zelen's exact test of homogeneity was used to test for an interaction; a large p-value indicates homogeneity across the subgroup effects.

[2] Results for the French site are not shown because there is only one 1aTN patient in that site

[3] 4/8/2013 is the date the 200th patient was randomized. The randomization scheme (Wk 12 arm to Wk 24 arm) was to change from 3:5 to 3:1 after 200 patients enrolled.

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

/s/

JOY D MELE
09/19/2014

GUOXING SOON
09/19/2014

DIONNE L PRICE
09/19/2014
concur

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

STATISTICAL REVIEW AND EVALUATION

CARCINOGENICITY STUDY

NDA:	206619
Drug Name:	Ombitasvir/ABT-450/ritonavir copackaged with dasabuvir
Indication:	Treatment of chronic hepatitis C virus infection, including patients with cirrhosis
Applicant:	AbbVie, Chicago, Illinois, 60064 CRO: (b) (4)
Date:	Assigned to Reviewer 3 February 2014 Data submitted 21 April 2014
Review Priority:	NME NDA
Biometrics Division:	Division 6
Statistical Reviewer:	Steve Thomson
Concurring Reviewer:	Team Leader: Karl Lin, Ph. D.
Medical Division:	Antiviral Products
Toxicologist:	Reviewer: CAPT Mark Seaton, Ph.D, DABT
Project Manager:	Katherine Schumann, MS.
Keywords:	Carcinogenicity, Cox regression, Kaplan-Meier product limit, Survival analysis, Trend test

Table of Contents

1. EXECUTIVE SUMMARY

According to the report for rats provided by Contract Research Organization, this submission was "conducted for Abbe Vie Inc. to evaluate the carcinogenic potential of A-1043422 (Lot Number 88310PP12) when coformulated with A-84538 [i.e., Ritonavir] (Lot Numbers 82554TL and 04301TL) and administered once daily for 104 weeks to Sprague-Dawley rats.". (page 13 of rat report, Study 126-501). Note that the CRO used these codes as labels for the drug combination under review, apparently with some variation in composition. The results of three additional studies in Tg.rasH2 mice, with similar objectives, were also submitted. All four studies were conducted [redacted] (b) (4) The descriptions of the studies are taken from the CRO final reports.

1.1. Conclusions and Recommendations

This submission summarizes the results of a rat study and three different Tg.RASH2 studies with daily dosing. The designs of the four studies are summarized below with details for each study in Sections 3.2.1 – 3.2.4.

Table 1. Design of Study 126-501: Sprague-Dawley Rat Study (dose volume 2 mL/kg)

Treatment Group	# Main study animals (# TK[1] animals)/gender	Nominal Dose (mg/kg/day)	Nominal Dosing Concentation (mg/mL)
0. H2O	80 (7)	0	0
1.Vehicle	80 (7)	0	0
2.Vehicle	80 (7)	0	0
3.Low	80 (7)	6/30	3/15
4.Medium	80 (7)	60/30	30/15
5. High	80 (7)	300/30	150/15

[1]Toxicokinetic phase animals, including 2 replacements.
In Study 126-501, in rats, an additional 20 animals per dose group were used as sentinel animals.

In the rat study, 126-501, and Tg.rasH2 study 126- 641 the compound being tested is identified as A-84530 (ritonavir) with A-1043422 (described as free acid) and vehicle Cremophor EL™: PEG 400:oleic acid at (10:10:80). In Tg.rasH2 study 126-701 the test compound is identified A-998821, sodium salt, with vehicle, 0.2% hydroxypropyl methylcellulose (HPMC) in distilled water. In Tg.rasH2 study 126-712 the test compound is identified A-1233617 in vehicle, 40% Phosal 53 MCT: 20% Polyethylene Glycol 400: 20% Poloxamer 124: 20% Cremophor RH40.

In each of the three following Tg.rasH2 studies, 15 animals were used a positive control, but were not included in data sets provided for the FDA statistical analysis.

Table 2. Design of Study 126-641: Tg,RASH2 Mice Study (dose volume 2 mL/kg)

Treatment Group	# Main study animals/gender	Nominal Dose (mg/kg/day)	Nominal Dosing Concentation (mg/mL)
0.H2O	35	0	0
1.Vehicle	35	0	0
2.Low	35	6/30	3/15
3..Medium	35	60/30	30/15
4. High	35	300/30	150/15

.**Table 3. Design of Study 126-701: Tg.RASH2 Mice Study** (dose volume 10 mL/kg)

Treatment Group	# Main study animals	Nominal Dose (mg/kg/day)	Nominal Dosing Concentation (mg/mL)
0.Vehicle	25	0	0
1. Low	25	200	20
2..Medium	25	600	60
3. High	25	2000	200

Table 4. Design of Study 126-712: Tg.RASH2 Mice Study (dose volume 2 mL/kg)

Treatment Group	# Main study animals	Nominal Dose (mg/kg/day)		Nominal Dosing Concentation (mg/mL)	
		Males	Females	Males	Females
0. H2O	28	0		0	
1.Vehicle	28	0		0	
2. Low	28	2.5	5	1.25	2.5
3..Medium	28	10	20	5	10
4. High	28	150		75	

Kaplan-Meier survival curves for each gender in study 126-501 in rats are given in Appendix 1 and results of statistical tests of differences in survival are summarized in Table 5, below:

Table 5. Study 126-501: Statistical Significances of Tests of Homogeneity and Trend in Survival in Rats

Hypotheses	Males		Females	
	Logrank	Wilcoxon	Logrank	Wilcoxon
Homogeneity over groups 1-4.	0.9938	0.6931	0.2190	0.1623
No Trend over four groups 1-4.	0.8793	0.8229	0.1675	0.2169
No difference between groups 1 & 4	0.9350	0.9825	0.3760	0.5254
No difference between groups 0 & 1.	0.2491	0.0996	0.4537	0.6368

From Figure A.1.1 in Appendix 1, it seems that in male rats the medium group seems to have roughly the highest survival, the vehicle group generally the lowest, and the other groups largely intertwined. Although lack of proof does not imply proof of lack, this is not sufficient to result in any statistically significant tests of overall homogeneity, trend, or differences between the high dose and controls (i.e., all six logrank or Wilcoxon $p \geq 0.6931$). Again, these tests are based on the pooled vehicle controls. The test of differences in survival between the pooled vehicle groups and the water group is the closest to the usual 0.05 statistical significance (logrank $p = 0.2491$, Wilcoxon $p= 0.0996$).

Among females, from Figure A.1.2 in Appendix 1, at the end of the study the low dose group seems to eventually tend toward somewhat higher mortality than the other dose groups, which in turn are largely intertwined. However, none of the eight statistical tests are particularly close to being statistically significant at Sir R.A. Fisher's suggested 0.05 level (i.e., all 8 $p \geq 0.1623$).

Summary survival tables are provided for each of the three Tg.rasH2 studies in sections 3.2.2 through 3.2.4 and in Tables A.1.2 through A.1.7 in Appendix 1. Note that none of these three Tg.rasH2 studies seem to show any particularly strong evidence of dose related differences in survival.

A large number of tumors are typically identified in the analysis of neoplasms, implying a large number of statistical tests. The problem of adjusting for this multiplicity of statistical tests is discussed in Section 1.3.1.4, below. To adjust for the multiplicity of statistical tests in tumors in the two year study, we currently recommend using the Haseman-Lin-Rahman (HLR) rules to adjust the interpretation of observed significance levels. That is, when testing for trend over dose and the difference between the highest dose group with vehicle group, to control the overall Type I error rate to roughly 10% one compares the unadjusted significance level of the trend test to 0.005 for common tumors and 0.025 for rare tumors (incidence < 1%), and the pairwise test to 0.01 for common tumors and 0.05 for rare tumors. Using these adjustments for other tests, like testing the comparisons between the Low and Medium dose groups versus vehicle can be expected to increase the overall type I error rate to some value above the nominal rough 10% level, possibly considerably higher than the nominal 10% rate. As discussed in Section 1.3.1.4, the effect of the excessive multiplicity due to three Tg.rasH2 studies is not well understood. When available, the sterile water groups are used to determine whether a tumor should be classified as rare or common, determination is based on incidence in the vehicle group. Thus in these analyses, if there were none of that specific tumor in the water group the tumor would be classified as rare, otherwise common.

Tables 6 and 7, below, shows those tumors in male Rats that had at least one non-multiplicity adjusted test that was statistically significant at a 0.10 level. The column headings annotations are as described in Appendix 2.

Table 6. Study 126-501: Tumor Incidence and Results in Male Rats

Organ/Tumor	H2O	Veh1	Veh2	Low	Med	High	ptrend	phigh vsVeh	pmed vsVeh	plow vsVeh
liver										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.0	50.7	56.8	49.1				
CARCINOMA, HEPATOCELLULAR	1	0	1	1	3	3	.0591	.1080	.1369	.5631
Adj. # at Risk	57.1	45.9	53.0	50.7	57.3	49.1				
Adenoma/Carcinoma, Hepato.	1	2	2	2	7	4	.2020	.2538	.0580	.6623
multicentric neoplasm										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.0	51.3	56.1	49.8				
Leukemia, any	0	0	0	1	1	2	.0541	.1096	.3636	.3423
pituitary gland										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.7	45.7	53.0	50.7	56.1	49.1				
CARCINOMA, PARS DISTALIS	2	1	0	0	0	2	.0969	.2577	1	1
skin										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.7	46.7	53.0	50.7	56.6	49.8				
Kerato./Sq.Cell Papilloma/ Carcinoma	4	3	1	0	2	4	.0668	.2492	.7065	1
Adj. # at Risk	57.1	45.7	53.0	50.7	56.4	49.1				
PAPILLOMA, SQUAMOUS CELL	0	0	0	0	1	2	.0316	.1096	.3636	.
Adj. # at Risk	57.4	45.7	53.0	50.7	56.4	49.1				
Papilloma/Carc.Squamous Cell	2	0	0	0	1	2	.0316	.1096	.3636	.
skin, subcutis										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.6	50.7	56.3	49.5				
SCHWANNOMA	0	0	1	0	1	2	.0922	.2547	.5935	1

None of the tests of trend in male rats are statistically significant (i.e., whether one considers the tumor to be common, so $p > 0.005$, or rare, $p > 0.025$). Thus no joint tests would be statistically significant. The only pairwise test that even achieves a significance level below 0.10 is the comparison between the medium dose group in pooled hepatocellular adenoma and carcinoma, and even allowing the increase in error due to applying the Haseman-Lin-Rahman rules to this comparison, it would not be considered as statistically significant ($p = 0.0580 > 0.01$, since it would be classified as a common tumor).

Table 7. Study 126-501: Tumor Incidence and Results in Female Rats

Organ/Tumor	H2O	Veh1	Veh2	Low	Med	High	ptrend	phigh vsVeh	pmed vsVeh	plow vsVeh
Systemic										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.2	55.0	53.2	47.5	50.4	55.4				
HEMANGIOSARCOMA	0	0	0	0	2	0	.4154	.	.0988	.
thyroid gland										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.2	55.0	53.2	48.1	50.7	55.4				
ADENOMA, FOLLICULAR CELL	0	0	0	2	2	2	.1268	.1125	.0988	.0933
Adj. # at Risk	53.8	55.0	53.2	49.2	50.7	55.4				
Adenoma/Carc. Foll. Cell	1	0	2	4	2	2	.4507	.4147	.3771	.0765
vagina										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.2	55.0	53.3	47.5	49.7	56.4				
GRANULAR CELL TUMOR	0	0	2	0	0	4	.0208	.1042	1	1

In female rats the test of trend in granular cell tumor in the vagina would be classified as barely statistically significant ($p = 0.0208 < 0.025$, since tumor would be classified as rare). No other tests achieved statistical significance.

Complete results of statistical poly-k tests of tumor trend and differences between dose groups in rats are given in Tables A.2.3 and A.2.4 in Appendix 2. As can also be seen in Appendix 2, in the Tg.rasH2 studies 126-641, 126-701, and 126-712 no tests of dose related trend or differences from vehicle were statistically significant. Complete tumor incidence tables for each gender within each Tg.rasH2 study are presented in Tables A.2.5-A.2.10.

1.2. Brief Overview of the Studies

One Standard Lab Rat study and three Tg.rasH2 studies, all conducted (b) (4) were submitted:

Study 126-501: 104-Week Oral Dose Carcinogenicity Study with A-1043422 and A-84538 in Rats

The results in the rat study are mostly negative, and it is this reviewer's opinion that there is little evidence of a particularly strong dose effect on survival or carcinogenicity. Even more clearly, results of the three Tg.rasH2 studies, 126-641, 126-701, and 126-712, suggest no strong dose related differences in survival or carcinogenity in any of the the three studies.

1.3. Statistical Issues and Findings

1.3.1. Statistical Issues

In this section, several issues, typical of statistical analyses of these studies, are considered. These issues include comments on the dual vehicles, details of the survival analyses, tests on tumorigenicity, multiplicity of tests on neoplasms, and the validity of the designs.

1.3.1.1. Water and Dual Vehicle Controls:

The rat study has three possible control groups, a water vehicle and two nominally identical vehicle control groups. These are numbered as groups 0, 1, 2, respectively. Most tables and plots in this report distinguish between these three groups. However the actual statistical tests are based on the pooled vehicle groups, ignoring the water group.

Note that some researchers would propose analyses that distinguish between the two vehicle control groups. The primary issue with such a procedure is that unless there are systemic problems with the conduct of the study, any observed differences should be due to random fluctuations between the treatment groups. That is, pre-study randomization to two identical controls should be equivalent to post-study randomization into two control groups. In the latter circumstances it would seem that few analysts would place any weight on any observed differences between the control groups (since a simple rerandomization would almost surely

eliminate any differences). But then logically no weight should be placed on any observed differences between vehicle controls in the current studies, and on possibly differing results when control groups are tested against other treatment groups.

1.3.1.2. Survival Analysis:

Appendix 1 reviews the specific FDA animal survival analyses in some detail.

The survival analyses in Rats presented in this report are based on both the log rank test and the Wilcoxon test comparing survival curves. The log rank tests tend to put higher weight on later events, while the Wilcoxon test tends to weight events more equally, and thus is more sensitive to earlier differences in survival. The logrank test is most powerful when the survival curves track each other, and thus the hazards, i.e., the conditional probability of the event in the next infinitesimal interval, would be roughly proportional. Note the logrank test seems to be the test usually recommended by statisticians, and seems to be the test used by the Sponsor. Both tests are used in the FDA analysis of mortality. The corresponding results of the Sponsor's analysis are summarized in Section 3.2.1.1.

Survival in each of the three Tg.rasH2 studies is high, with no obvious trends over increasing dose. With such high survival it was felt that simple incidence tables would be the most appropriate way to analyze results. If actual statistical tests were needed, due to the low sample size and low incidence it was felt that exact, permutation based, logrank tests would be more appropriate than the the asymptotic tests used for rats. These results are presented with the survival summaries.

1.3.1.3. Multiplicity of Tests on Survival:

Using both the logrank and Wilcoxon tests, for each gender in Rats there are 8 tests of survival differences. Assuming tests were performed at the usual 0.05 level, and the tests were stochastically independent, but there were actually absolutely no differences in survival across groups (so one would hope no tests would be statistically significant), the probability of at least one statistically significant result in each gender was about 0.336. These bounds assume the tests are independent, which they clearly are not, but these values can give some idea of the possible price paid for the multiplicity of hypothesis tests in the statistical frequentist paradigm.

1.3.1.4. Tests on Neoplasms:

The data sets requested for the analysis of rodent carcinogenicity studies are supposed to include a record for each animal organ combination that was not evaluated. It is possible that some such records are missing from the provided data sets. If a number of the animals are not examined, but the proportions of animals showing the tumor under study in each treatment group is roughly the same as in the subset of animals actually reported the calculated p-values will generally be too large, i.e., results will be less statistically significant than they should be, possibly much less. If we can assume the process that determines whether or not a tumor is analyzed in each specific tumor is random, it is perhaps appropriate to consider such endpoints to be both analyzed AND have the tumor.

Ignoring these possible problems, the Sponsor's analyses of tumorigenicity are Peto tests, where logrank tests are used for fatal and mortality independent tumors and Mantel-Haenzel test where incidental tumors are grouped by time of detection. These tests are criticized since they require accurate determination of whether a tumor is fatal or incidental. In this submission the results of these Peto tests seem to agre with the FDA analysis.

This FDA analysis is based on a modification of the Cochran-Armitage test of trend in mortality (please see Bailer & Portier, 1988, Bieler & Williams, 1993). Inspecting a large number of studies, Bailer and Portier noted that survival time seemed to fit a Weibull distribution, generally with a shape parameter of between 1 and 5, with 3 a typical value. With t_{max} denoting the maximal time to terminal sacrifice and t_{obs} the time to detection of the tumor in the animal, they proposed weighting the animal by $(t_{obs}/t_{max})^k$, so that an animal that survives for say 52 weeks in 104 week study without the tumor being analyzed is counted as $(1/2)^k$ of an animal in the risk set for that tumor. For k = 3, that means that particular animal would count as 1/8 of an animal. Further, the k = 3 specification seems to represent tumor incidence where some animals are perhaps more sensitive and respond earlier to the insult than the remining animals. Under this structure time to incidence would tend to follow a cubic expression. Thus an animal with the specific tumor being studied or who survives to terminal sacrifice without the tumor will be given a weight of 1 when counting the number of animals at risk. However, animals that die early without the tumor are down weighted when counting the number of animals in the risk set for that specific tumor. With differential mortality, this can mean a substantial reduction in the size of that risk set. Note this seems to be an appropriate adjustment for dose groups that are terminated early. The report of the Society of Toxicological Pathology "town hall" meeting in June 2001 recommeded the use of this poly-k modification of the so-called Cochran-Armitage tests of trend over the corresponding Peto tests used by the Sponsor.

The computed significance levels are based on small sample exact permutation tests of tumor incidence. In the tumor incidence tables the effective size of the risk set for each tumor is listed in the row labeled "Adjusted # at risk ", and seems to be a more appropriate denominator when comparing incidence rates than the simple unadjusted number evaluated.

1.3.1.6. Multiplicity of Tests on Neoplasms:

There were a large number of gender by organ by tumor combinations to be tested, with both tests of trend, overall homogeneity, and pairwise tests between the high dose and vehicle. To control the probability of false posititives,it seems appropriate to use a correction for multiplicity in the statistical tests. Current FDA practice is based on the Haseman-Lin-Rahman adjustments.

The Haseman-Lin-Rahman rules are based on the original multiplicity adjustment of Haseman (1983) and extended by Lin and Rahman with various simulations. Based on his extensive experience with such analyses, for pairwise tests in a two species study comparing control to the High dose group, Haseman (1983) claimed that for a roughly 0.10 (10%) overall

false positive error rate, rare tumors should be tested at a 0.05 (5%) level, and common tumors (with a historical control incidence greater than 1%) at a 0.01 level. Lin & Rahman (1998) proposed a further p-value adjustment for tests of trend. That is, for a roughly 0.10 (10%) overall false positive error rate in tests of trend, rare tumors should be tested at a 0.025 (2.5%) level and common tumors at a 0.005 (0.5%) level. Other specifications are presented in the Table 5 below. This approach is intended to balance both Type I error and Type II error (i.e., the error of concluding there is no evidence of a relation to tumorgenicity when there actually is such a relation).

The proposed Haseman-Lin-Rahman bounds are taken from the draft *Guidance for Industry Statistical Aspects of the Design, Analysis, and Interpretation of Chronic Rodent Carcinogenicity Studies of Pharmaceuticals*, (HHS, 2013). The bounds on the right in table 8, below, are grouped so that the first two columns (columns 1-2) correspond to testing either trend or pairwise between the high dose and control, but not both. The remaining two columns (columns 3-4), correspond to testing both overall trend and pairwise tests between the high dose and control. The bounds (on levels of significance) cited below are for a submission with either two chronic, two-year study in rats or one two-yesr study in rats and another short term or medium term study in mice. In this analysis we will use the observed incidence in the no treatment group to decide if a tumor is rare or common.

Table 8. Recommended Multiplicity Adjusted Bounds on Significance Levels in Rat Study

	Separate Tests		Joint test	
	Trend	Pairwise	Trend	Pairwise
Common Tumor	0.005	0.01	0.005	0.05
Rare Tumor	0.025	0.05	0.025	0.10

For a single Tg.rasH2 mouse study, it is recommended that all tests be assessed at a 0.05 level. Note that with three Tg.rasH2 studies as in this submission, applying this procedure to each study increases the number of tries and thus such is clearly going to be anti-conservative. However, the effect of the multiple Tg.rasH2 studies has not been investigated. For this submission, since no tests of trend or pairwise comparisons in the Tg.rash2 studies were statistically significant, this point is moot.

1.3.1.6. Validity of the Designs:

When determining the validity of designs there are two key points:

1) adequate drug exposure

2) tumor challenge to the tested animals.

1) is related to whether or not sufficient animals survived long enough to be at risk of forming late-developing tumors and 2) is related to the Maximum Tolerated Dose (MTD), designed to achieve the greatest likelihood of tumorigenicity.

Lin and Ali (2006), quoting work by Haseman, have suggested that in standard laboratory rodent species, a survival rate of about 25 animals, out of 50 or more animals (i.e.

50%), between weeks 80-90 of a two-year study may be considered a sufficient number of survivors as well as one measure of adequate exposure. From tables 15 and 16 on page 18, as a percentage of the High dose group animals that survived to week 91, this criterion seems to slightly fail in both genders male rats (43.8%), and slightly in female rats (48.8%).

The mean weight values used to derive differences and ratios in the following tables were taken directly from the Sponsor's report (Table 5, Body Weight group mean values (g)), pages 123-132). The change from baseline in the tables below is the simple difference between the means at the specified dates, and thus animals that die early are only counted at the study initiation, not at the end of the study.

Table 9. Mean Weights and Changes (in g) in Male Rats

Dose Group	Dose mg/kg/day	Week		Change from Baseline	% change relative to pooled vehicle	% change relative to H2O
		-1	104			
0. H2O	0	207.1	797.7	590.6	111.3%	
1+2. Vehicle	0	203.8	734.3	530.5		89.8%
3. Low	6/30	204.2	722.9	518.7	97.8%	87.8%
4. Medium	60/30	204.0	727.1	523.1	98.6%	88.6%
5. High	300/30	203.9	754.1	550.2	103.7%	93.2%

Table 10. Mean Weights and Changes (in g) in Female Rats

Dose Group	Dose mg/kg/day	Week		Change from Baseline	% change relative to pooled vehicle	% change relative to H2O
		-1	98			
0. H2O	0	147.6	496.3	348.7	113.0%	
1+2. Vehicle	0	146.8	455.5	308.7		88.6%
3. Low	6/30	149.9	453.3	303.6	98.3%	87.1%
4. Medium	60/30	147.0	459.9	312.9	101.4%	89.7%
5. High	300/30	147.1	451.7	304.6	98.7%	87.4%

Chu, Ceuto, and Ward (1981), citing earlier work by Sontag et al. (1976) recommend that the MTD "is taken as 'the highest dose that causes no more than a 10% weight decrement as compared to the appropriate control groups, and does not produce mortality, clinical signs of toxicity, or pathologic lesions (other than those that may be related to a neoplastic response) that would be predicted to shorten the animal's natural life span' " From Tables 9 and 10 above, if the water only group is considered as the appropriate control all the other study groups have about a 10% weight decrement. If the vehicle group is considered to be the appropriate control there seems to be no particular weight decrement, possibly suggesting the MTD was not achieved, since decrement in weight is uaually expected when the dose is close to the MTD. Of course, the actual interpretation of these observations is in the toxicologists's bailiwick.

` However, the Sponsor summarizes weight changes as follows: "Vehicle-related decreases in mean body weight gain were noted for both sexes in combined vehicle controls when

compared to water controls, as were test article related (and likely vehicle exacerbated) decreases in mean body weight gain in males at all A-1043422/A-84538 doses (generally peaking within the timeframe of Weeks 26-58) and in females at the two highest A-1043422/A-84538 doses (generally peaking within the timeframe of Weeks 42-58) when compared to combined vehicle controls. However, these decreases did not exhibit a clear dose response pattern and were not considered adverse due to their generally low magnitudes, as all A-1043422/A-84538 groups still exhibited substantial weight gain over the course of the study." (pages 29-30 of report)

In a related issue the Sponsor notes that in the rat study there was a general decrease in food consumption in the actual treatment groups and especially the vehicle controls relative to water controls.

Again from 2) above, excess mortality not associated with any tumor or sacrifice in the higher dose groups might suggest that the MTD was exceeded. This suggests that a useful way to assess whether or not the MTD was achieved is to measure early mortality not associated with any identified tumor. If this is high in the higher dose groups it suggests that animals tend to die before having time to develop tumors. Table 11, below, displays the number of animals in each dose group in rats that died of a natural death or moribund sacrifice, but did not show any tumors (i.e., the "Event"):

Table 11. Study 126-501: Natural Death with No Identified Tumor in Rats (Male/Female)

	0.H2O.	1.Vehicle	2.Vehicle	3 Low	4.Medium	5.High
Males Event	9	19	16	19	10	18
No event	71	61	64	61	70	62
Females Event	1	4	5	8	7	2
No event	79	76	75	72	73	78

Unlike most of tests conducted in this review, these tests include the water group. Note that including this group makes the study groups more disparate and thus increases the significance of any test of homogeneity (i.e. reduces the p-value). However, even including this significant (Males p = 0.1372, Females p = 0.1160).

The role of such an assessment of the MTD in Tg.rasH2 studies is not clear to this reviewer, however similar results in these three studies are presented below:

Table 12. Study 126-641: Natural Death with No Identified Tumor

	0.H2O.	1.Vehicle	2.Low	3.Medium	4.High
Males Event	1	2	2	7	2
No event	34	33	33	28	33
Females Event	1	2	1	1	1
No event	34	33	34	34	34

Although arguably not needed, due to the low incidence in the Tg.rasH2 studies, Fisher exact, permutation based tests were used to assess differences in this "event" among study groups. In Study 126-641 the Fisher test of homogeneity were not statistically significant (Males p = 0.1211, Females p = 1.0).

Table 13. Study 126-701: Natural Death with No Identified Tumor

	0.Vehicle	1 Low	2.Medium	3.High
Males Event	3	0	2	1
No event	22	25	23	24
Females Event	0	0	3	0
No event	25	25	22	25

Table 14. Study 126-712: Natural Death with No Identified Tumor

	0.H2O.	1.Vehicle	2 Low	3.Medium	4.High
Males Event	1	0	1	0	1
No event	24	25	24	25	24
Females Event	0	1	1	0	1
No event	25	24	24	25	24

. In Studies 126-701 and 126-712 the Fisher exact tests of homogeneity were not statistically significant (Study 126-701: Males p = 0.5057, Females p = 0.0569, Study 126-712: both Males and Females p = 1.0). Thus, with the possible exception of females in Study 126-701, there is no evidence of differences in early deaths without tumor. Like the other observations above, this require the expertise of the toxicologist, but these tests may provide evidence that the MTD was not exceeded in any of the studies.

1.3.2. Statistical Findings

Please see Section 1.1 above.

2. INTRODUCTION

2.1. Overview

Results from a single study in Sprague-Dawley rats and three studies in Tg.rasH2 mice and supporting data sets were submitted to assess the carcinogenic potential of the Ritonavir mixture.

2.2. Data Sources

For each of the four studies the Sponsor provided a transport file containing data sets tumor.sas7bdat. It should be noted that the results of the positive control group were not included in any of the three Tg.rasH2 provided data sets.

3. STATISTICAL EVALUATION

3.1. Evaluation of Efficacy

NA

3.2. Evaluation of Safety

More detailed results on the study are presented below.

3.2.1 A 104-Week Oral Dose Carcinogenicity Study with A-1043422 and A-84538 in Rats

STUDY DURATION: Male Rats 105 weeks, Females 99 weeks.
DOSING STARTING DATE: 9 November 2010.
NECROPSY COMPLETED: 8 November 2012
STUDY ENDING DATE (Draft Report Mailed): 22 November 2013.
RAT STRAIN: CD® [Crl:CD®(SD)].
ROUTE: Daily Oral Gavage

In this rat study, 126-501 tested is identified as A-84530 (ritonavir) with A-1043422 (described as free acid) and vehicle Cremophor EL™: PEG 400:oleic acid at (10:10:80). The following table (a repeat of Table 1 above) summarizes the study design:

Table 15. Design of Study 126-501: Spargue-Dawley Rat Study (dose volume 2 mL/kg)

Treatment Group	# Main study animals (# TK[1] animals)/gender	Nominal Dose (mg/kg/day)	Nominal Dosing Concentation (mg/mL)
0. H2O	80 (7)	0	0
1.Vehicle	80 (7)	0	0
2.Vehicle	80 (7)	0	0
3.Low	80 (7)	6/30	3/15
4..Medium	80 (7)	60/30	30/15
5. High	80 (7)	300/30	150/15

[1]Toxicokinetic phase animals, including 2 replacement.

In Study 126-501 an additional 20 animals per dose group were used as sentinel animals.

The Sponsor summarizes the study conduct as follows : " Three treatment groups of 80 male and 80 female CD® [Crl:CD®(SD)] rats were orally administered the test articles (A-1043422/A-84538) at respective dose levels of 6/30, 60/30, and 300/30 mg/kg/day at a dose volume of 2 mL/kg. An additional 20 animals/sex were designated as sentinel animals for health screening purposes only. One additional group of 80 animals/sex served as the water control group and received deionized water, at a dose volume of 2 mL/kg. Two additional groups of 80

animals/sex/group served as the vehicle control groups and received the vehicle control article, Cremophor EL™: PEG 400:oleic acid (10:10:80, w/w/w), at a dose volume of 2 mL/kg. The water control, vehicle controls, or test articles were administered to all groups via oral gavage, once a day for up to 731 consecutive days. Males were necropsied on Days 729-731 (Week 105), while females were necropsied on Days 693-694 (99 weeks of dosing) after water control females reached a survival level of 20 animals on Week 99. Additionally, six groups of seven animals/sex/group served as toxicokinetic (TK) animals and received the water control, vehicle control, or test articles in the same manner and the same dose level as the main study groups." (page 13 of rat report)

Dose levels were justified as follows: "Selection of dose levels for this study were based in part upon results of a 3-month toxicology study with A-1043422/A-84538, with subsequent presentation and approval of an integrated dose selection justification to the FDA Carcinogenicity Advisory Committee.." (page 23 of rat report) "The water control, vehicle control, or test articles will be administered orally by gavage. The control animals will be given the water control article or vehicle control article at the same volume as the treated animals. Individual doses will be based on the most recent body weights. Upon arrival in the animal room, all prepared formulations will be stirred prior to and during the administration period." (page 65 of rat report)

Animals were individually housed in a wire mesh cage, unless there was evidence of foot damage. Food and water were available ad libitum. The Sponsor states that "Observations for morbidity, mortality, injury, and the availability of food and water were conducted twice daily for all animals." (page 22 of report) Detailed observations were made "weekly."

3.2.1.1 Sponsor's Results and Conclusions

This section will present a summary of the Sponsor's analysis on survivability and tumorigencity in rats.

Survival analysis:

The Sponsor summarized mortality results as follows: "No dose-response trend in mortality was detected in males or females, nor were A-1043422/A-84538 dose groups different in mortality from the vehicle control group for either sex. No mortality difference was detected between the water control group and the vehicle control group for either sex." (page 11 of statistical report, page 5441 of rat report).

Tumorigenicity analysis:

"Tumor data were organized by tumor type within each organ/tissue system. The tumor incidence data were analyzed using Peto's method. For palpable tumors, the onset-rate method was used. For all other tumors, the time of detection was the time of death for an animal. If the tumor was the cause of the death, it was termed a fatal tumor; otherwise, it was termed an incidental tumor. For fatal tumors, the death-rate method of analysis was used. For incidental tumors, the prevalence method of analysis was employed, using a set of time intervals (Weeks 0-

50, 51-80, 81-end and terminal necropsy) suggested by Center for Drug Evaluation and Research (CDER) of Food and Drug Administration. Results from these analyses were combined to form a final assessment of carcinogenicity of A-1043422/A-84538

"Three statistical tests were performed for each tumor type by organ/tissue: a test for a positive dose-response trend (Groups 2 and 3 combined, Group 4, Group 5 and Group 6), a comparison between the water control group (Group 1) and the vehicle control group (Groups 2 and 3 combined), and a comparison between the high-dose group (Group 6) and the vehicle control group (Groups 2 and 3 combined). All tests were one-sided. The alternative hypothesis for the comparison between the water control group and the vehicle control group was that the vehicle group had more animals with tumors; and the alternative hypothesis for the comparison between the high-dose group and the vehicle control group was that the high-dose group had more animals with tumors. The dose response trend test was conducted using a scale of scores corresponding to the actual dose levels of 0, 6, 60, and 300 mg/kg/day of A-1043422. For both trend and pairwise comparison tests, the p-values were computed using exact permutation" (page 10 of statistical report, page 5440 of rat report).

To "control the overall Type I error rate when testing many different tumor types in different organs, we adopted the decision rule recommended by the CDER for declaring significance in these statistical tests (multiplicity adjustment). For rare tumors (having a historical control incidence of 1% or less), the recommended significance level is 0.025 for the trend test and 0.050 for the pairwise comparison between the high-dose group and the control; for common tumors, the recommended significance level is 0.005 for the trend test and 0.010 for the pairwise comparison between the high-dose group and the control." (pages 10-11 of statistical report, page 5440-5441 of rat report).

The Sponsor summarized results in rats as follows: "There are four statistical tests in these tables which have unadjusted p-values less than or equal to 0.05:

(1) The test of a positive dose response for carcinoma, hepatocellular in liver in the males (exact test p=0.050). There were three tumors found in the high dose group, three tumors found in the middle dose group, one tumor found in low dose group and one tumor found in vehicle control group (Groups 2 and 3 combined). The pairwise comparison test of high dose versus control group comparison was not significant (exact test p=0.097). The trend test was not significant when the CDER multiplicity adjustment[5] level was used.
(2) The test of a positive dose response for papilloma, squamous cell in skin in the males (exact test p=0.03). There were two tumors found in the high dose group, one tumor found in the middle dose group and no tumor found in the other groups. The high dose versus control group comparison was not significant (exact test p=0.103). The trend test was not significant when the CDER multiplicity adjustment[5] level was used.
(3) The test of a positive dose response for adenoma in parathyroid glands in the females (exact test p=0.026). There were four tumors found in the high dose group, two tumors found in vehicle control group (Groups 2 and 3 combined), and no tumor found in the low and middle dose

groups. The pairwise comparison test of high dose versus control group comparison was not significant (exact test p=0.123). The trend test was not significant when the CDER multiplicity adjustment[5] level was used.
(4) The test of a positive dose response for granular cell tumor in vagina in the females (exact test p=0.016). There were four tumors found in the high dose group, two tumors found in the vehicle control group (Groups 2 and 3 combined), and no tumor found in the low and middle dose groups. The pairwise comparison test of high dose versus control group comparison was not significant (exact test p=0.083). The granular cell tumor in vagina is considered a common tumor based on the historical control data. The historical control data consists of 54 control groups from the studies conducted by Charles River Laboratories in years 2001-2011. The overall tumor rate across these 54 studies is 2.107%. The group size of these studies ranges from 50 to 75 animals and the tumor rate ranges from 0% to 11.7%. The trend test result from the current study was not considered statistically significant when the CDER multiplicity adjustment level 0.005 for common tumors was used. Furthermore, the observed tumor rate in the high dose group 5% is well within the historical control range (see below the histogram of the historical control tumor rate distribution for granular cell tumor in vagina in female rats). More than 16% (9/54) of these control groups have the tumor rate greater than the observed tumor rate in the high dose female group in this study." (pages 12-13 of statistical report, pages 5442-5443 of rat report).

3.2.1.2 FDA Reviewer's Results

This section will present the current Agency findings on survival and tumorigenicity in male and female Rats.

Survival analysis:

Kaplan-Meier plots comparing treatment groups in both studies are given in Appendix 1, along with more details of the analysis. The following tables (Table 16 for male Rats, Table 17 for female Rats) summarize the mortality results for the dose groups. The data were grouped for the specified time period, and present the number of deaths during the time interval over the number at risk at the beginning of the interval. The percentage cited is the percent survived at the end of the interval.

Table 16. Summary of Male Rats Mortality

Period	0.H2O	1.Vehicle	2.Vehicle	3.Low	4.Medium	5.High
0-52	5/80[1] 93.8%[2]	17/80 78.8%	13/80 83.8%	14/80 82.5%	6/80 92.5%	10/80 87.5%
53-78	18/75 71.3%	19/63 55.0%	12/67 68.8%	17/66 61.3%	14/74 75.0%	23/70 58.8%
79-91	16/57 51.3%	12/44 40.0%	14/55 51.3%	10/49 48.8%	21/60 48.8%	12/47 43.8%
92-105	15/41 32.5%	10/32 27.5%	17/41 30.0%	17/39 27.5%	20/39 23.8%	12/35 28.8%
terminal	26	22	24	22	19	23

[1] number deaths / number at risk
[2] per cent survival to end of period

Table 17. Summary of Female Rats Mortality

Period	0.H2O	1.Vehicle	2.Vehicle	3.Low	4.Medium	5.High
1-52	2/80[1] 97.5%[2]	5/80 93.8%	7/80 91.3%	6/80 92.5%	7/80 91.3%	6/80 92.5%
53-78	26/78 65.0%	19/75 70.0%	22/73 63.8%	28/74 57.5%	27/73 57.5%	18/74 70.0%
79-91	18/52 42.5%	19/56 46.3%	12/51 48.8%	20/46 32.5%	13/46 41.3%	17/56 48.8%
92-99	14/34 25.0%	12/37 31.3%	16/39 28.8%	6/26 25.0%	8/33 31.3%	9/39 37.5%
terminal	20	25	23	20	25	30

[1] number deaths / number at risk
[2] per cent survival to end of period.

Table 18 (a repeat of Table 2 and Table A.1.1 in Appendix 1), below, displays the results of survival tests.

Table 18. Study 126-501: Statistical Significances of Tests of Homogeneity and Trend in Survival in Rats

Hypotheses	Males		Females	
	Logrank	Wilcoxon	Logrank	Wilcoxon
Homogeneity over groups 1+2, 3, 4, 5.	0.9938	0.6931	0.2190	0.1623
No Trend over four groups 1+2, 3, 4, 5.	0.8793	0.8229	0.1675	0.2169
No difference between groups 1+2 & 5	0.9350	0.9825	0.3760	0.5254
No difference between groups 0 & 1+2.	0.2491	0.0996	0.4537	0.6368

From Figure A.1.1, it seems that in male rats the medium group seems to have the roughly the highest survival with the vehicle group generally the lowest, and the other groups

largely intertwined. Although lack of proof does not imply proof of lack, this is not sufficient to results in any statistically significant tests of overall homogeneity, trend, or differences between the high dose and controls (i.e., all six logrank or Wilcoxon p ≥ 0.6931). Again, these tests are based on the pooled vehicle controls. The test of differences in survival between the pooled vehicle groups and the water group is the closest to the usual 0.05 statistical significance (logrank p = 0.2491, Wilcoxon p= 0.0996).

Among females, from the Kaplan-Meier Figure A.1.2 in Appendix 1, at the end of the study the low dose group seems to eventually tend toward somewhat higher mortality than the other dose groups, which in turn are largely intertwined. However, none of the eight statistical tests are particularly close to being statistically significant at Sir R.A. Fisher's 0.05 level (i.e., all 8 p ≥ 0.1623).

Tumorigenicity analysis:

Tables 19 and 20, below, repeats of Tables 6 and 7 above, and repeats of Tables A.2.1 and A.2.2 in Appendix 2, show the tumors that had at least one non-multiplicity adjusted test that was statistically significant at a 0.10 level.

Table 19. Potentially Statistically Significant Neoplasms in Male Rats

Organ/Tumor	H2O	Veh1	Veh2	Low	Med	High	ptrend	phigh vsVeh	pmed vsVeh	plow vsVeh
Male Rats										
liver										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.0	50.7	56.8	49.1				
CARCINOMA, HEPATOCELLULAR	1	0	1	1	3	3	.0591	.1080	.1369	.5631
Adj. # at Risk	57.1	45.9	53.0	50.7	57.3	49.1				
Adenoma/Carcinoma, Hepato.	1	2	2	2	7	4	.2020	.2538	.0580	.6623
multicentric neoplasm										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.0	51.3	56.1	49.8				
Leukemia, any	0	0	0	1	1	2	.0541	.1096	.3636	.3423
pituitary gland										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.7	45.7	53.0	50.7	56.1	49.1				
CARCINOMA, PARS DISTALIS	2	1	0	0	0	2	.0969	.2577	1	1
skin										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.7	46.7	53.0	50.7	56.6	49.8				
Kerato./Sq.Cell Papilloma/ Carcinoma	4	3	1	0	2	4	.0668	.2492	.7065	1
Adj. # at Risk	57.1	45.7	53.0	50.7	56.4	49.1				
PAPILLOMA, SQUAMOUS CELL	0	0	0	0	1	2	.0316	.1096	.3636	.
Adj. # at Risk	57.4	45.7	53.0	50.7	56.4	49.1				
Papilloma/Carc.Squamous Cell	2	0	0	0	1	2	.0316	.1096	.3636	.
skin, subcutis										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.6	50.7	56.3	49.5				
SCHWANNOMA	0	0	1	0	1	2	.0922	.2547	.5935	1

None of the tests of trend in male rats are statistically significant (i.e., whether one considers the tumor to be common, so p > 0.005, or rare, p > 0.025). Thus no joint tests would

be statistically significant. The only pairwise test that even achieves a significance level below 0.10 is the comparison between the medium dose group in pooled hepatocellular adenoma and carcinoma, and even allowing the increase in error due to applying the Haseman-Lin-Rahman rules to this comparison, it would not be considered as statistically significant (p = 0.0580 > 0.01, since it would be classified as a common tumor).

Table 20. Study 126-501: Tumor Incidence and Results in Female Rats

Organ/Tumor	H2O	Veh1	Veh2	Low	Med	High	ptrend	phigh vsVeh	pmed vsVeh	plow vsVeh
Systemic										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.2	55.0	53.2	47.5	50.4	55.4				
HEMANGIOSARCOMA	0	0	0	0	2	0	.4154	.	.0988	.
thyroid gland										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.2	55.0	53.2	48.1	50.7	55.4				
ADENOMA, FOLLICULAR CELL	0	0	0	2	2	2	.1268	.1125	.0988	.0933
Adj. # at Risk	53.8	55.0	53.2	49.2	50.7	55.4				
Adenoma/Carc. Foll. Cell	1	0	2	4	2	2	.4507	.4147	.3771	.0765
vagina										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.2	55.0	53.3	47.5	49.7	56.4				
GRANULAR CELL TUMOR	0	0	2	0	0	4	.0208	.1042	1	1

In rats the test of trend in granular cell tumor in the vagina would be classified as barely statistically significant (p = 0.0208 < 0.025, since tumor would be classified as rare). No other tests achieved statistical significance.

Complete results of statistical poly-k tests of tumor trend and differences between dose groups in rats are given in Tables A.2.3 and A.2.4 in Appendix 2.

3.2.2 Study 126-641: 26 –Week Oral Gavage Oncogenicity Study with A-1043422 and A-84538 in Model 001178-t (Hemizygous) CBYB6F1-TG(HRAS)2JIC Mice

STUDY DURATION: 26 Weeks.
DOSING STARTING DATE: 21 February 2012.
NECROPSY COMPLETED: 24 August 2012
STUDY ENDING DATE (Final Report dated): 16 May 2013.
ROUTE: Daily Oral Gavage

In Tg.rash2 study 126- 641 the compound being tested is identified as A-84530 (ritonavir) with A-1043422 (described as free acid) and vehicle Cremophor EL™: PEG 400:oleic acid at (10:10:80). The basic study design (a copy of Table 2) is summarized below:

Table 21. Design of Study 126-641: Tg,RASH2 Mice Study (dose volume 2 mL/kg)

Treatment Group	# Main study animals/gender	Nominal Dose (mg/kg/day)	Nominal Dosing Concentation (mg/mL)
0.H2O	35	0	0
1.Vehicle	35	0	0
2.Low	35	6/30	3/15
3..Medium	35	60/30	30/15
4. High	35	300/30	150/15

The Sponsor summarizes the study conduct as follows : "Five respective groups of 35 male and 35 female Taconic model 001178-T (hemizygous, CByB6F1-Tg(HRAS)2Jic mice were administered water (deionized water), vehicle [Cremophor EL®:PED 400:oleic acid (10:10:80, w/w/w)], or test articles (A-1043422/ A-84538) at dose levels of 6/30, 60/30, and 300/30 mg/kg/day once a day via oral gavage for at least 182 days. A positive control group of 15 male and 15 female Taconic model 001178-T (hemizygous), CByB6F1-Tg(HRAS)2Jic mice was administered the positive control article (N-Nitroso-N-methylurea (NMU) in citrate buffered saline at pH 4.5) once on Day 1 via an intraperitoneal injection. An additional group of 10 male and 10 female Taconic model 001178-T (hemizygous), CByB6F1-Tg(HRAS)2Jic mice served as sentinel animals for serological health screen purposes." (page 10 of Study 126-641 report)

"Animals considered suitable for study were weighed prior to treatment. After excluding animals that did not meet the inclusion criterion based on weight, 380 animals were randomized, by sex, into six treatment groups using a standard, by weight, measured value randomization procedure according to the following design: (page 7 of statistic report, page 1873 of overall report).

"The water control, vehicle control, or test article was administered orally by gavage. The positive control article was administered via an intraperitoneal injection. The control animals were given the control article at the same volume as the treated animals. Individual doses were based on the most recent body weights.

"Microscopic examination of fixed hematoxylin and eosin-stained paraffin sections was performed on sections of tissues indicated in the list of the organs or tissues table specified in the protocol for all animals indicated, including any animal dying spontaneously or euthanized in extremis." (page 7 of statistical report, 1873 of overall report).

3.2.2.1 Sponsor's Results and Conclusions

Survival Analysis:

For both males and females, no dose-response trend in mortality was detected, nor were

A-1043422/A-84538 dose groups different in mortality from the control group. No mortality difference was detected between the water control group and the vehicle control group for either males or females (page 10 of statistical report, 1876 of overall report).

Carcinogenicity Analysis:

The Sponsor's reports indicate that tumor data were analyzed using Peto's methods applied to "tumor incidence rates off [sic] five groups for all tumor types that occurred, along with the results of statistical analysis of the trend test, the water control group versus the vehicle control group comparison test, and the high dose group versus vehicle control group comparison test.

"None of the statistical tests showed significant results (all unadjusted p-values >0.05)." (page 10 of statistic report, 1876 of overall report).

3.2.2.2 FDA Reviewer's Results

This section will present the current Agency findings on survival and tumorigenicity in male and female Rats.

Survival analysis:

The following table presents the weeks of death or sacrifice in this study. When more than one one animal has such an event in the week, the number of animals experiencing the event is noted in the following parentheses. Thus, for example, "14 (2)" indicates that 2 animals in that dose group died in the 14th week. An asterisk is used to denote sacrifice, and, thus, for example, "*27 (34)" indicates that 34 animals in that dose group were sacrificed in week 27. In each study, data for the positive control was not included in the provided data and hence are not included here.

Table 22.: Study 126-641 Male Tg.rasH2 Mice Survival

Dose Group	Dosage	Survival Time in Weeks	
0.H2O	0	11	*27 (34)
1. Vehicle	0	17, 18, 20	*27 (32)
2. Low	6/30	14 (2)	*27 (33)
3. Medium.	60/30	6, 11, 15, 16 (2), 17 (2)	*27 (28)
4. High.	300/30	14	*27 (33)

(Homogeneity in Survival p = 0.1312, No trend in ranks p = 0.0656)

Table 23.: Study 126-641 Female Tg.rasH2 Mice Survival

Dose Group	Dosage	Survival Time in Weeks	
0.H2O	0	4, 25	*27 (33)
1. Vehicle	0	4, 13, 25	*27 (32)
2. Low	6/30	14	*27 (34)
3. Medium.	60/30	18, 25	*27 (33)
4. High.	300/30	14	*27 (34)

(Homogeneity in Survival p = 0.7908, No trend in ranks p = 0.2320)

Although, this reviewer would argue that no actual statistical tests are needed, the test of homogeneity in survival using permutation based logrank tests over the five study groups was not statistically significant in either gender (Males p = 0.1312, Females p = 0.7908).

Tumorigenicity analysis:

Complete incidence tables are presented in tables A.2.5 and A.2.6. From these tables it is apparent that no test of tumorgenicity achieved the 0.05 level of significance. Again, lack of proof is not proof of lack, but it is consistent with the hypothesis of no particular carcinogenic effect.

3.2.3 Study 126-701: 26-Week Oral Gavage Carcinogenicity Study with A-998821 Sodium in Model 001178-T (Hemizygous) CBYB6F1-TG(HRAS)2JIC Mice

STUDY DURATION: 26 Weeks.
DOSING STARTING DATE: 15 March 2012.
NECROPSY COMPLETED: 14 September 2012
STUDY ENDING DATE (Final Report dated): 16 January 2013.
ROUTE: Daily Oral Gavage

In Tg.rash2 study 126-701 the test compound is identified A-998821, sodium salt, with vehicle, 0.2% hydroxypropyl methylcellulose (HPMC) in distilled water. The basic design of this study (a copy of Table 3) is summarized below:

Table 24. Design of Study 126-701: Tg.RASH2 Mice Study (dose volume 10 mL/kg)

Treatment Group	# Main study animals	Nominal Dose (mg/kg/day)	Nominal Dosing Concentation (mg/mL)
0.Vehicle	25	0	0
1. Low	25	200	20
2..Medium	25	600	60
3. High	25	2000	200

The Sponsor summarizes the study conduct as follows: “Three treatment groups of 25 Taconic model 001178-T (hemizygous), CByB6F1-Tg(HRAS)2Jic mice/sex were administered A-998821 at respective dose levels of 200, 600, and 2000 mg/kg/day. One additional group of 25 hemizygous animals/sex served as the control and received the vehicle, 0.2% hydroxypropyl methylcellulose (HPMC) prepared in distilled water. The formulations were administered via oral gavage once a day for 26 consecutive weeks, at a dose volume of 10 mL/kg. An additional group of 15 hemizygous animals/sex served as the positive control group and received the positive control article, N-Nitroso-N-methylurea, at a dose level of 75 mg/kg via intraperitoneal injection once on Day 1 at a dose volume of 10 mL/kg.” (page 10 of rat report)

“Animals considered suitable for study were weighed prior to treatment. After the appropriate number of animals were excluded, 230 animals were randomized, by sex, into five groups using a standard, by weight, measured value randomization procedure” (page 7 of statistics report, 1209 of rat report)

“The test article and vehicle formulations were administered orally by gavage. The positive control article was administered via an intraperitoneal injection. Individual doses were based on the most recent body weights.

“Microscopic examination of fixed hematoxylin and eosin-stained paraffin sections were performed on sections of tissues indicated in the list of the organs or tissues table specified in the protocol for all animals indicated, including any animal dying spontaneously or euthanized in extremis.” (page 7 of statistics report, 1209 of rat report)

3.2.3.1 Sponsor’s Results and Conclusions

Survival analysis:

The Sponsor summarizes mortality results using lifetable methods: “For both males and females, no dose-response trend in mortality was detected, nor were A-998821 dose groups different in mortality from the control group.” (page 10 of statistics report, 1212 of rat report)

Tumor Incidence

As with the other Tg.rash2 studies the Sponsor summarizes results of the Peto tests of difference in tumor “rates of four groups for all tumor types that occurred, along with the results of statistical analysis of the trend test, and the high dose group versus vehicle control group comparison test.

“None of the statistical tests showed significant results (all unadjusted p-values>0.05).” (page 10 of statistics report, 1212 of rat report)

3.2.3.2 FDA Reviewer's Results

This section will present the current Agency findings on survival and tumorigenicity in male and female mice in study 126-701:.

Survival analysis:

The following table presents the weeks of death or sacrifice in this study. Again,when more than one one animal has such an event in the week, the number of animals experiencing the event is noted in the following parentheses. Thus, for example, "26 (2)" indicates that 2 animals in that dose group died in the 26th week. An asterisk is used to denote sacrifice, and, thus, for example, "*27 (25)" indicates that all 25 animals in that dose group were sacrificed in week 27. In each study, data for the positive control was not included in the provided data and hence are not included here.

Table 25. Study 126-701 Male Tg.rasH2 Mice Survival

Dose Group	Dosage	Survival Time in Weeks	
0. Vehicle	0	17, 19, 26 (2)	*27 (21)
1. Low	200	-	*27 (25)
2. Medium.	600	4, 23	*27 (23)
3. High.	2000	8	*27 (24)

(Homogeneity in Survival p = 0.1906, No trend in ranks p = 0.1510)

Table 26. Study 126-701 Female Tg.rasH2 Mice Survival

Dose Group	Dosage	Survival Time in Weeks	
0. Vehicle	0	-	*27 (25)
1. Low	200	17, 18	*27 (23)
2. Medium.	600	15, 20, 24, 25	*27 (21)
3. High.	2000	-	*27 (25)

(Homogeneity in Survival p = 0.0448, No trend in ranks p = 0.3568)

In the permutation based log rank tests in female mice, note that he relatively large number of somewhat early deaths in the medium dose group was sufficient to generate a slightly statistically significant test of homogeneity, but not a clear trend. No such differences are apparent in male mice.

Tumorigenicity analysis:

As with the other Tg.rash2 studies, no test of tumorgenicity in study 126-701 achieved the 0.05 level of significance, consistent with the hypothesis of no particular carcinogenic effect. Complete incidence tables are presented in tables A.2.7 and A.2.8.

3.2.4 Study 126-712: 26-Week Oral Gavage Carcinogenicity Study with A-1233617 Free Form in Model 001178-T HEMIZYGOUS) CBYB6F1-TG(HRAS)2JIC Mice

STUDY DURATION: 26 Weeks.
DOSING STARTING DATE: 19 June 2012
NECROPSY COMPLETED: 21 December 2012
STUDY ENDING DATE (Final Report dated): 3 April 2013.
ROUTE: Daily Dermal Dosing

In Tg.rash2 study 126-712 the test compound is identified A-1233617 in vehicle, 40% Phosal 53 MCT: 20% Polyethylene Glycol 400: 20% Poloxamer 124: 20% Cremophor RH40. The basic design (a copy of Table 4) is summarized below:

Table 27. Design of Study 126-712: Tg.RASH2 Mice Study (dose volume 2 mL/kg)

Treatment Group	# Main study animals	Nominal Dose (mg/kg/day)		Nominal Dosing Concentation (mg/mL)	
		Males	Females	Males	Females
0. H2O	28	0		0	
1.Vehicle	28	0		0	
2.Vehicle	28	0		0	
3.Low	28	2.5	5	1.25	2.5
4..Medium	28	10	20	5	10
5. High	28	150		75	

The Sponsor reports that "This study was conducted for AbbVie Inc. to evaluate the carcinogenic potential of A-1233617 when administered daily via oral gavage for 26 weeks in Model 001178-T (Hemizygous) CByB6F1-Tg(HRAS)2Jic [transgenic] mice. Three test article groups of 28 male transgenic mice were administered A-1233617 at dose levels of 2.5, 10, and 150 mg/kg/day and three test article groups of 28 female transgenic mice were administered A-1233617 at dose levels of 5, 20, and 150 mg/kg/day. Another group of 28 transgenic animals/sex served as the vehicle control group and received the vehicle control, 40% Phosal 53 MCT: 20% Polyethylene Glycol 400: 20% Poloxamer 124: 20% Cremophor RH40, by weight. One additional group of 28 transgenic animals/sex served as the water control group and received deionized water. The water control, vehicle control, or test article formulations were administered to all groups via oral gavage once a day for 26 consecutive weeks, at a dose volume of 2 mL/kg. An additional group of 15 transgenic animals/sex served as the positive control group and received the positive control article, N-Nitroso-N-methylurea (MNU), at a dose level of 75 mg/kg via intraperitoneal injection once on Day 1 at a dose volume of 10 mL/kg. Three groups of 21 Taconic model 001178-W (nontransgenic), CByB6F1-Tg(HRAS)2Jic [nontransgenic] mice/sex served as toxicokinetic (TK) animals and received A-1233617 in the same manner and at the same dose levels as the main study groups for 13 consecutive weeks.

Another group of 6 nontransgenic animals/sex served as the TK vehicle control group and received the vehicle control in the same manner and dose volume as the main study vehicle control group. An additional group of 10 nontransgenic mice/sex served as sentinel animals and was used for health screening purposes only. Three of the assigned animals/sex in all groups (excluding MNU and sentinel groups) served as possible replacement animals. If not used as replacement animals, the animals were euthanized via carbon dioxide inhalation and the carcasses were discarded without further evaluation.

"Observations for morbidity, mortality, injury, and the availability of food and water were conducted twice daily for all animals. Toxicity was assessed by detailed clinical observations, body weights, and food consumption. Blood samples for determination of plasma concentrations of A-1233617and toxicokinetic (TK) evaluations were collected from designated TK animals at designated time points during Week 13. After blood collection, the TK animals were euthanized and the carcasses were discarded. At study termination, necropsy examinations were performed and tissues were microscopically examined for main study animals." (page 10 of report)

3.2.4.1 Sponsor's Results and Conclusions

Survival Analyses

The Sponsor's report indicates that mortality data were analyzed using life table methods. These results were summarized as follows: "No dose-response trend in mortality was detected in males or females, nor were A-1233617 dose groups different in mortality from the vehicle control group for either sex. No mortality difference was detected between the water control group and the vehicle control group for either sex." (page 10 of statistics report, 1545 of overall report).

Tumor Incidence

Results of the Peto analysis of tumor incidence using a trend test, the high dose group versus vehicle control group comparison test, and the water control group versus the vehicle control group were that: "None of these statistical tests performed showed a significant result (all unadjusted p-values > 0.05)." (page 10 of statistics report, 1545 of overall report)

3.2.4.2 FDA Reviewer's Results

This section will present the current Agency findings on survival and tumorigenicity in male and female Rats.

Survival analysis:

The tables below summarize mortality results:

Table 28. Study 126-712 Male Tg.rasH2 Mice Survival

Dose Group	Dosage	Survival Time in Weeks	
0. H2O	0	24	*27 (24)
1. Vehicle	0	-	*27 (25)
2. Low	2.5	10	*27 (24)
3. Medium.	10	-	*27 (25)
4. High.	150	16	*27 (24)

(Homogeneity in Survival p = 1.0, No trend in ranks p = 0.4503)

Table 29. Study 126-712 Female Tg.rasH2 Mice Survival

Dose Group	Dosage	Survival Time in Weeks	
0. H2O	0	13, 16	*27 (23)
1. Vehicle	0	9	*27 (24)
2. Low	5	9, 24	*27 (23)
3. Medium.	20	-	*27 (25)
4. High.	150	27	*27 (24)

(Homogeneity in Survival p = 0.6546, No trend in ranks p = 0.3412)

It is clear from the tables above, and confirmed by the results of permutation based logrank tests, that there are no particular differences in mortality across study groups.

Tumorigenicity analysis:

As with the other Tg.rash2 studies, no test of tumorgenicity in study 126-712 achieved the 0.05 level of significance, consistent with the hypothesis of no particular carcinogenic effect. Complete incidence tables are presented in tables A.2.9 and A.2.10..

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

NA

5. SUMMARY AND CONCLUSIONS

5.1. Statistical Issues and Collective Evidence

Please see Section 1.3 above.

5.2. Conclusions and Recommendations

Please see section 1.1.

APPENDICES:

Appendix 1. Survival Analysis

Simple summary life tables in mortality are presented in the report (Tables 16 and 17, above). Kaplan-Meier estimated survival curves across study groups for each gender are displayed below in Figures A.1.1 and A.1.2. The plots include 95% confidence intervals around each survival curve (colored area around each curve). These plots are also supported by tests of homogeneity in survival over the treatment groups, with the porovision that the actual tests are based on the pooled control groups, while the Kaplan-meier plots display the two vehicle controls separately. The statistical significance levels (i.e., p-values) are provided in Table A.1.1., below. Please recall that group 0 denotes the H_2O group, 1 and 2 the vehicle group, with 1+2 the pooled vehicle, and 3, 4, 5 increasing dose groups. One might note that the log rank tests place greater weight on later events, while the Wilcoxon test tends to weight them more equally, and thus, in the rat test, places more weight on differences in earlier events than does the log rank test.

Table A.1.1 Study 126-501: Statistical Significances of Tests of Homogeneity and Trend in Survival in Rats

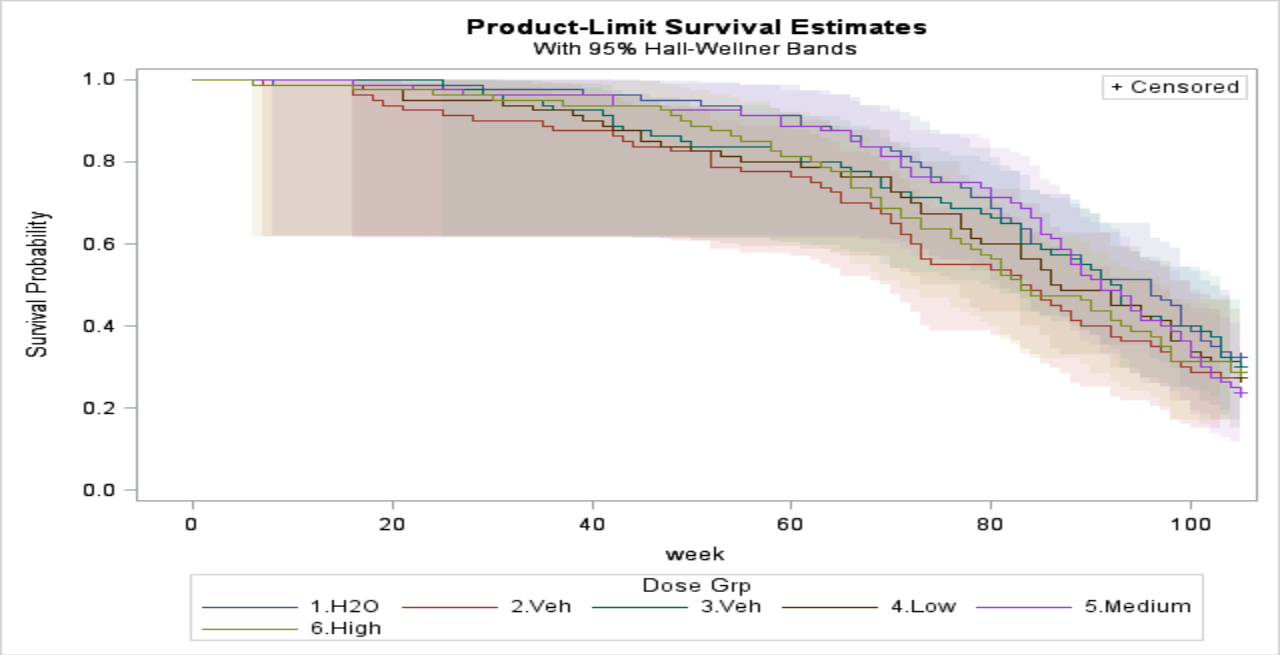
Hypotheses	Males		Females	
	Logrank	Wilcoxon	Logrank	Wilcoxon
Homogeneity over groups 1+2, 3, 4, 5.	0.9938	0.6931	0.2190	0.1623
No Trend over four groups 1+2, 3, 4, 5.	0.8793	0.8229	0.1675	0.2169
No difference between groups 1+2 & 5	0.9350	0.9825	0.3760	0.5254
No difference between groups 0 & 1+2.	0.2491	0.0996	0.4537	0.6368

From Figure A.1.1, it seems that in male rats the medium group seems to have the roughly the highest survival with the vehicle group generally the lowest, and the other groups largely intertwined. Although lack of proof does not imply proof of lack, this is not sufficient to results in any statistically significant tests of overall homogeneity, trend, or differences between the high dose and controls (i.e., all six logrank or Wilcoxon $p \geq 0.6931$). Again, these tests are based on the pooled vehicle controls. The test of differences in survival between the pooled vehicle groups and the water group is the closest to the usual 0.05 statistical significance (logrank $p = 0.2491$, Wilcoxon $p= 0.0996$).

Among females, from Figure A.1.2 below, at the end of the study the low dose group seems to eventually tend toward somewhat higher mortality than the other dose groups, which in turn are largely intertwined. However, none of the eight statistical tests are particularly close to being statistically significant at Sir R.A. Fisher's 0.05 level (i.e., all 8 $p \geq 0.1623$).

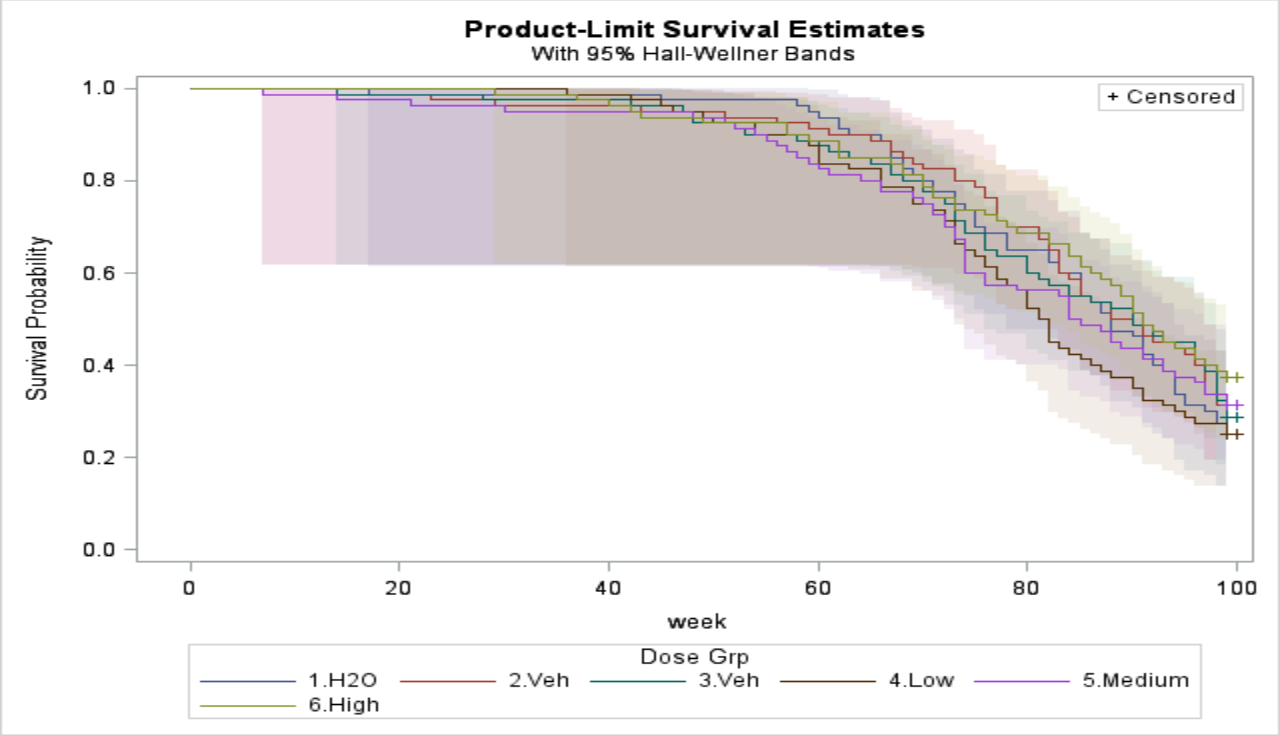
The Kaplan-Meier survival curves are displayed below:

Figure A.1.1 Kaplan-Meier Survival Curves for Male Rats



The corresponding survival curves in female Rats are given below:

Figure A.1.2 Kaplan-Meier Survival Curves for Female Rats



The following tables summarize the time to death or sacrifice in the three different Tg.rasH2 studies, resulting in six tables when stratifying also on gender. These are displayed in the form of a variation on a dot plot, where the weeks of death of individual animals are listed with each dose group. When more than one one animal has such an event in the week, the number of animals experiencing the event is noted in the following parentheses. Thus, for example, "14 (2)" indicates that 2 animals in that dose group died in the 14th week. An asterisk is used to denote sacrifice, and, thus, for example, "*27 (34)" indicates that 34 animals in that dose group were sacrificed in week 27. In each study, data for the positive control was not included in the provided data and hence are not included here.

This reviewer feels these tables are themselves adequately informative. But in case the reader needs actual p-values, results of a test of homogeneity in survival over all groups (including water) and a test of trend over the dose ranks are also provided below each table. It was felt that the within group sample sizes and event counts were likely to be too small for the usual asymptotic tests as used above in the rat study (126-501). These tests are based on the estimated permutation distribution of the actual computed statistic.

Table A.1.2: Study 126-641 Male Tg.rasH2 Mice Survival

Dose Group	Dosage	Survival Time in Weeks	
0.H2O	0	11	*27 (34)
1. Vehicle	0	17, 18, 20	*27 (32)
2. Low	6/30	14 (2)	*27 (33)
3. Medium.	60/30	6, 11, 15, 16 (2), 17 (2)	*27 (28)
4. High.	300/30	14	*27 (33)

(Homogeneity in Survival p = 0.1312, No trend in ranks p = 0.0656)

Table A.1.3: Study 126-641 Female Tg.rasH2 Mice Survival

Dose Group	Dosage	Survival Time in Weeks	
0.H2O	0	4, 25	*27 (33)
1. Vehicle	0	4, 13, 25	*27 (32)
2. Low	6/30	14	*27 (34)
3. Medium.	60/30	18, 25	*27 (33)
4. High.	300/30	14	*27 (34)

(Homogeneity in Survival p = 0.7908, No trend in ranks p = 0.2320)

Table A.1.4: Study 126-701 Male Tg.rasH2 Mice Survival

Dose Group	Dosage	Survival Time in Weeks	
0. Vehicle	0	17, 19, 26 (2)	*27 (21)
1. Low	200	-	*27 (25)
2. Medium.	600	4, 23	*27 (23)
3. High.	2000	8	*27 (24)

(Homogeneity in Survival p = 0.1906, No trend in ranks p = 0.1510)

Table A.1.5: Study 126-701 Female Tg.rasH2 Mice Survival

Dose Group	Dosage	Survival Time in Weeks	
0. Vehicle	0	-	*27 (25)
1. Low	200	17, 18	*27 (23)
2. Medium.	600	15, 20, 24, 25	*27 (21)
3. High.	2000	-	*27 (25)

(Homogeneity in Survival p = 0.0448, No trend in ranks p = 0.3568)

Table A.1.6: Study 126-712 Male Tg.rasH2 Mice Survival

Dose Group	Dosage	Survival Time in Weeks	
0. H2O	0	24	*27 (24)
1. Vehicle	0	-	*27 (25)
2. Low	2.5	10	*27 (24)
3. Medium.	10	-	*27 (25)
4. High.	150	16	*27 (24)

(Homogeneity in Survival p = 1.0, No trend in ranks p = 0.4503)

Table A.1.7: Study 126-712 Female Tg.rasH2 Mice Survival

Dose Group	Dosage	Survival Time in Weeks	
0. H2O	0	13, 16	*27 (23)
1. Vehicle	0	9	*27 (24)
2. Low	5	9, 24	*27 (23)
3. Medium.	20	-	*27 (25)
4. High.	150	27	*27 (24)

(Homogeneity in Survival p = 0.6546, No trend in ranks p = 0.3412)

Note that the only test that is statistically significant at the usual 0.05 level is the test of homogeneity in females in Study 126-701, and it does not appear tobe dose related, If survival were homogeneous over dose groups the probability that at least one test would be statistically significant at the 0.05 level would be roughly 0.26. So this reviewer would place no particular weight on that result.

.

Thus it appears that there is no particularly strong evidence of dose related survival differences in rats or in any of the three tg.rasH2 mouse studies.

.

Please recall that group 0 denotes the H_2O group, 1 and 2 the vehicle group, and 3, 4, 5 increasing dose groups. One might note that the log rank tests place greater weight on later events, while the Wilcoxon test tends to weight them more equally, and thus, in the rat test, places more weight on differences in earlier events than does the log rank test.

Appendix 2. FDA Poly-k Tumorigenicity Analysis

The poly-k test, here with k=3, modifies the original Cochran-Armitage test to adjust for differences in mortality (please see Bailer & Portier, 1988, Bieler & Williams, 1993). The tests used here are small sample exact permutation tests of tumor incidence. When there were no tumors of the specific type being analyzed in either column of the 2x2 table corresponding to a pairwise comparison an argument could be made that the p-value for this test should be 1.0. However, largely for readability, in the tables below these p-values are considered as missing (i.e., corresponding to a null test), denoted by a period ".". Note that the StatXact program used for these analyses adjusts for the variance, which would be 0. Then the significance levels of the test statistics are based on the result of a division by 0, i.e., undefined, and hence StatXact codes these p-values as missing.

For each gender by organ the number of animals microscopically analyzed is presented first. Note that indicating an organ was not examined requires a specification in the data (please see section 2.2 above). Since no animals in the data sets provided are noted as being excluded, it seems possible that this specification could be missing in some of this data. Then the number of animals at risk could be inflated, and the proportion of animals with tumor would be artificially decreased. Thus, as discussed in Section 1.5 above, for some of these organs it is possibly more appropriate to define the actual endpoint used in the statistical analysis be the condition of being microscopically analyzed AND show the tumor. This does have problems unless treatment groups are not treated equally except for actual dose. The entry for each tumor is preceded by the adjusted number of animals at risk for that endpoint. It seems clear that an animal that dies early without having displaying that endpoint reduces the size of the risk set for that getting that particular endpoint. The poly-k test down weights such animals, and as discussed in Section 1.3.1.4, above, the sum of these poly-k weights seems to be a better estimate of the number of animals at risk of getting that tumor than the simple number of animals analyzed. This sum is given in the row labeled "Adjusted # at risk ". Tumor incidence is presented next, with the significance levels of the tests of trend over the vehicle group or pooled vehicle groups, then the incidence in the low, medium, and high dose group in each study. The next entry is the significance levels (i.e., p-values) of the test of trend over the vehicle and actual dose groups. This is followed by the results of the pairwise tests between the high, medium, and low dose groups versus (possibly pooled) vehicle. In the Tg.RasH2 mice studies 126-641 and 126-712, and rat study 126-501 incidence in the water (H2O) group is only used to assess background tumor incidence, and thus whether a tumor is considered to be rare (background incidence <1%) or common.

To adjust for the multiplicity of tests the so-called Haseman-Lin-Rahman (HLR) rules discussed in Section 1.3.1.5 are often applied. As discussed in this particular case this reviewer would recommend treating this as a single species study. That is, when testing for trend over dose and the difference between the highest dose group with a control group, to control the overall Type I error rate to roughly 10% one compares the unadjusted significance level of the trend test to 0.01 for common tumors and 0.05 for rare tumors, and the pairwise test to 0.05 for

common tumors and 0.10 for rare tumors. Using these adjustments for other tests, like testing the comparisons between the low and medium dose groups versus vehicle can be expected to increase the overall type I error rate to some value above the nominal rough 10% level, possibly considerably higher than the nominal 10% rate.

Tables A.2.1 and A.2.2 below, show the tumors in both genders in rats that had at least one non-multiplicity adjusted test that was statistically significant at a 0.10 level.

Table A.2.1. Study 126-501: Tumor Incidence and Results in Male Rats

Organ/Tumor	H2O	Veh1	Veh2	Low	Med	High	ptrend	phigh vsVeh	pmed vsVeh	plow vsVeh
Male Rats										
liver										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.0	50.7	56.8	49.1				
CARCINOMA, HEPATOCELLULAR	1	0	1	1	3	3	.0591	.1080	.1369	.5631
Adj. # at Risk	57.1	45.9	53.0	50.7	57.3	49.1				
Adenoma/Carcinoma, Hepato.	1	2	2	2	7	4	.2020	.2538	.0580	.6623
multicentric neoplasm										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.0	51.3	56.1	49.8				
Leukemia, any	0	0	0	1	1	2	.0541	.1096	.3636	.3423
pituitary gland										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.7	45.7	53.0	50.7	56.1	49.1				
CARCINOMA, PARS DISTALIS	2	1	0	0	0	2	.0969	.2577	1	1
skin										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.7	46.7	53.0	50.7	56.6	49.8				
Kerato./Sq.Cell Papilloma/ Carcinoma	4	3	1	0	2	4	.0668	.2492	.7065	1
Adj. # at Risk	57.1	45.7	53.0	50.7	56.4	49.1				
PAPILLOMA, SQUAMOUS CELL	0	0	0	0	1	2	.0316	.1096	.3636	.
Adj. # at Risk	57.4	45.7	53.0	50.7	56.4	49.1				
Papilloma/Carc.Squamous Cell	2	0	0	0	1	2	.0316	.1096	.3636	.
skin, subcutis										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.6	50.7	56.3	49.5				
SCHWANNOMA	0	0	1	0	1	2	.0922	.2547	.5935	1

Note that the tests of trend in male rats are never statistically significant (i.e. whether one considers the tumor to be common, $p > 0.005$, or rare, $p > 0.025$). Thus no joint tests would be statistically significant. The only pairwise test that even achieves a significance level below 0.10 is the comparison between the medium dose group in pooled hepatocellular adenoma and carcinoma, and even allowing the increase in error due to applying the Haseman-Lin-Rahman rules to this comparison, it would not be considered as statistically significant ($p = 0.0580 > 0.01$, since it would be classified as a common tumor).

Table A.2.2. provides similar results in female Rats:

Table A.2.2 Study 126-501: Tumor Incidence and Results in Female Rats

Organ/Tumor	H2O	Veh1	Veh2	Low	Med	High	ptrend	phigh vsVeh	pmed vsVeh	plow vsVeh
Systemic										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.2	55.0	53.2	47.5	50.4	55.4				
HEMANGIOSARCOMA	0	0	0	0	2	0	.4154	.	.0988	.
thyroid gland										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.2	55.0	53.2	48.1	50.7	55.4				
ADENOMA, FOLLICULAR CELL	0	0	0	2	2	2	.1268	.1125	.0988	.0933
Adj. # at Risk	53.8	55.0	53.2	49.2	50.7	55.4				
Adenoma/Carc. Foll. Cell	1	0	2	4	2	2	.4507	.4147	.3771	.0765
vagina										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.2	55.0	53.3	47.5	49.7	56.4				
GRANULAR CELL TUMOR	0	0	2	0	0	4	.0208	.1042	1	1

In rats the test of trend in granular cell tumor in the vagina would be classified as barely statistically significant ($p = 0.0208 < 0.025$, since tumor would be classified as rare). Note of the tumors above, only squamous cell carcinoma of the treated skin would be classified as a rare tumor. Then using the Haseman-Lin-Rahman (HLR) rules to adjust for multiplicity, the test of trend in the vehicle to high dose in was statistically significant ($p = 0.0264 < 0.05$), however the corresponding test between the high dose and vehicle was not quite significant ($p = 0.1369 > 0.10$). Accepting the inflation of error from including tests other than the trend and high versus vehicle, the comparison of the low dose and vehicle in both adenoma and pooled adenoma and adenocarcinoma were statistically significant ($p = 0.0051, 0.0031 < 0.01$, respectively). None of the other tests achieved the significance levels required for either the separate or pooled tests to be statistically significant using the HLR rules for a single study.
(please see Section 1.3.3 below).

Complete results of statistical poly-k tests of tumor trend and differences between dose groups in male and female Rats are given in Tables A.2.3 and A.2.4, respectively, below:

Table A.2.3. Tumor Incidence and Results in Male Rats

Organ/Tumor	H2O	Veh1	Veh2	Low	Med	High	ptrend	phigh vsVeh	pmed vsVeh	plow vsVeh
Systemic										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.3	45.7	53.0	50.7	56.1	49.1				
HEMANGIOMA	1	0	0	1	0	0	.6126	.	.	.3378
Adj. # at Risk	58.1	45.7	53.0	51.0	56.8	49.9				
HEMANGIOSARCOMA	3	1	2	1	2	1	.5794	.8065	.6003	.8168
Adj. # at Risk	58.4	45.7	53.0	51.0	56.8	49.9				
Hemangioma/Hemangiosarcoma	4	1	2	2	2	1	.6415	.8065	.6003	.5591
Adj. # at Risk	57.1	45.8	53.0	51.4	56.1	49.1				
LYMPHOMA	0	1	0	2	0	0	.8538	1	1	.2698
adrenal glands										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.8	45.9	53.0	50.7	56.1	49.1				
ADENOMA, CORTICAL	2	3	1	0	0	0	1	1	1	1
Adj. # at Risk	58.3	49.9	54.5	51.3	57.7	50.8				
PHEOCHROMOCYTOMA	10	17	14	6	9	8	.8839	.9819	.9864	.9974
Adj. # at Risk	57.1	45.7	53.0	50.7	56.1	49.1				
PHEOCHROMOCYTOMA, COMPLEX	0	0	0	0	1	0	.4150	.	.3636	.
Adj. # at Risk	58.3	49.9	54.5	51.3	57.7	50.8				
Pheocromocytoma, Any	10	17	14	6	10	8	.8853	.9819	.9732	.9974
brain										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.3	45.7	53.1	50.7	56.1	49.1				
ASTROCYTOMA	2	0	3	1	0	0	.9783	1	1	.8118
Adj. # at Risk	57.1	45.7	53.0	50.7	56.1	49.1				
GRANULAR CELL TUMOR	0	0	1	0	1	0	.6588	1	.5966	1
Adj. # at Risk	57.1	45.7	54.9	51.7	56.1	49.1				
MIXED GLIOMA	0	0	2	1	0	0	.9415	1	1	.7125
Adj. # at Risk	57.1	45.7	53.0	50.7	56.1	49.1				
OLIGODENDROGLIOMA	0	0	0	1	0	0	.6126	.	.	.3378
cavity, abdominal										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.0	51.2	56.1	49.1				
ADENOCARCINOMA (PRIMARY SITE	0	0	0	1	0	0	.6142	.	.	.3423
Adj. # at Risk	57.1	45.8	53.0	50.7	56.5	49.1				
FIBROSARCOMA	0	1	0	0	1	0	.6588	1	.5966	1
Adj. # at Risk	57.1	46.5	53.0	51.2	56.8	49.1				
SCHWANNOMA	1	1	0	1	1	0	.6987	1	.5935	.5659
cavity, oral										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.0	50.7	56.4	49.1				
CARCINOMA, SQUAMOUS CELL	1	0	0	0	1	0	.4150	.	.3636	.
epididymides										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.0	51.3	56.1	49.1				
MESOTHELIOMA	1	0	0	2	0	0	.6965	.	.	.1156
heart										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.0	50.7	56.1	49.5				
SCHWANNOMA	0	1	1	1	0	1	.5350	.7068	1	.7127

Table A.2.3. (cont.) Tumor Incidence and Results in Male Rats

Organ/Tumor	H2O	Veh1	Veh2	Low	Med	High	ptrend	phigh vsVeh	pmed vsVeh	plow vsVeh
kidneys										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.0	50.7	56.1	49.1				
ADENOMA, RENAL TUBULE	0	0	1	0	0	0	1	1	1	1
Adj. # at Risk	57.1	45.7	53.0	50.7	56.1	49.1				
Adenoma/Carc.Renal Tub. Any	0	0	1	0	1	0	.6588	1	.5966	1
Adj. # at Risk	57.1	45.7	53.0	50.7	56.1	49.1				
CARCINOMA, RENAL TUBULE,(AV)	0	0	0	0	1	0	.4150	.	.3636	.
Adj. # at Risk	57.5	45.7	53.0	50.7	56.1	49.1				
HEMANGIOSARCOMA	1	0	0	0	0	0	.	.	.	.
Adj. # at Risk	57.1	45.9	53.0	50.7	56.6	49.1				
LIPOSARCOMA	2	1	0	1	1	0	.7017	1	.5966	.5631
Adj. # at Risk	57.1	45.7	53.0	50.7	56.1	49.6				
RENAL MESENCHYMAL TUMOR	0	0	0	0	0	1	.1937	.3333	.	.
liver										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.9	53.0	50.7	56.6	49.1				
ADENOMA, HEPATOCELLULAR	0	2	1	1	4	1	.5872	.8065	.2182	.8118
Adj. # at Risk	57.1	45.7	53.0	50.7	56.8	49.1				
CARCINOMA, HEPATOCELLULAR	1	0	1	1	3	3	.0591	.1080	.1369	.5631
Adj. # at Risk	57.1	45.9	53.0	50.7	57.3	49.1				
Adenoma/Carcinoma, Hepato.	1	2	2	2	7	4	.2020	.2538	.0580	.6623
lung										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.0	50.7	56.1	49.1				
ADENOMA,BRONCHIOLAR ALVEOLAR	0	0	0	1	0	0	.6126	.	.	.3378
Adj. # at Risk	57.1	45.7	53.2	50.7	56.1	49.1				
CARCINOMA, SQUAMOUS CELL	0	0	1	0	0	0	1	1	1	1
lymph node, mesenteric										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.3	45.7	53.0	50.7	56.1	49.1				
HEMANGIOMA	1	0	0	1	0	0	.6126	.	.	.3378
Adj. # at Risk	57.7	45.7	53.0	50.7	56.1	49.1				
HEMANGIOSARCOMA	2	0	1	0	0	0	1	1	1	1
mammary gland										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.0	50.7	56.1	49.1				
ADENOCARCINOMA	0	0	1	0	0	0	1	1	1	1
Adj. # at Risk	57.1	45.7	53.0	50.7	56.1	49.1				
ADENOMA	1	0	0	0	0	0	.	.	.	.
Adj. # at Risk	57.2	45.7	53.2	50.7	56.1	49.1				
Adenocarc./Adenoma/Fibro-adenoma	1	0	2	0	0	1	.4773	.7068	1	1
Adj. # at Risk	57.2	45.7	53.2	50.7	56.1	49.1				
Adenoma/Fibroadenoma	1	0	2	0	0	1	.4773	.7068	1	1
Adj. # at Risk	57.2	45.7	53.2	50.7	56.1	49.1				
FIBROADENOMA	1	0	1	0	0	1	.3505	.5571	1	1
multicentric neoplasm										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.0	51.3	56.1	49.8				
LEUKEMIA, GRANULOCYTIC	0	0	0	1	0	1	.1998	.3333	.	.3423
Adj. # at Risk	57.1	45.7	53.0	50.7	56.1	49.1				
LEUKEMIA,LARGE GRANULAR LYMP	0	0	0	0	1	1	.1230	.3333	.3636	.
Adj. # at Risk	57.1	45.8	53.0	51.4	56.1	49.1				
LYMPHOMA	0	1	0	2	0	0	.8538	1	1	.2698
Adj. # at Risk	57.1	45.7	53.0	51.3	56.1	49.8				
Leukemia, any	0	0	0	1	1	2	.0541	.1096	.3636	.3423
Adj. # at Risk	57.1	46.4	53.0	51.1	56.1	49.1				
SARCOMA, HISTIOCYTIC	1	2	1	1	1	3	.1460	.3128	.8373	.8142

Table A.2.3. (cont.) Tumor Incidence and Results in Male Rats

Organ/Tumor	H2O	Veh1	Veh2	Low	Med	High	ptrend	phigh vsVeh	pmed vsVeh	plow vsVeh
nerve, sciatic										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.0	50.7	56.1	49.1				
SCHWANNOMA	1	0	0	1	0	0	.6126	.	.	.3378
pancreas										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.6	46.1	53.0	50.8	56.3	49.1				
ADENOMA, ACINAR CELL	3	3	0	2	2	2	.3833	.5369	.5958	.5458
Adj. # at Risk	59.8	47.8	54.9	51.1	56.5	49.8				
ADENOMA, ISLET CELL	13	9	14	7	3	7	.7874	.9235	.9994	.9378
Adj. # at Risk	57.7	46.1	53.0	50.8	56.3	49.1				
Adenoma/Carc. Acinar Cell	4	3	1	2	3	2	.4731	.6489	.4948	.6578
Adj. # at Risk	59.9	47.8	55.0	51.7	56.8	50.2				
Adenoma/Carc. Islet Cell	14	10	19	10	5	8	.9147	.9726	.9994	.9177
Adj. # at Risk	57.1	45.7	53.0	50.7	56.1	49.1				
CARCINOMA, ACINAR CELL	1	0	1	1	1	0	.7017	1	.5966	.5631
Adj. # at Risk	57.2	45.7	53.1	51.2	56.4	49.5				
CARCINOMA, ISLET CELL	1	1	5	4	2	1	.9018	.9457	.8571	.4659
parathyroid glands										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.2	50.7	56.1	49.1				
ADENOMA	1	2	3	0	3	1	.6677	.9167	.6085	1
pituitary gland										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	72.0	59.9	63.2	61.5	68.5	64.3				
ADENOMA, PARS DISTALIS	62	44	54	47	52	47	.8115	.8753	.7597	.7288
Adj. # at Risk	57.1	45.7	53.0	50.7	56.1	49.7				
ADENOMA, PARS INTERMEDIA	0	0	1	1	1	1	.3459	.5571	.5966	.5631
Adj. # at Risk	72.7	59.9	63.2	61.5	68.5	64.3				
Adenoma/Carc. Pars Distalis	64	45	54	47	52	49	.6797	.7939	.8001	.7708
Adj. # at Risk	72.7	59.9	63.2	61.5	68.5	64.3				
Adenoma/Carcinoma	64	45	54	47	53	49	.6869	.7939	.7297	.7708
Adj. # at Risk	57.7	45.7	53.0	50.7	56.1	49.1				
CARCINOMA, PARS DISTALIS	2	1	0	0	0	2	.0969	.2577	1	1
Adj. # at Risk	57.1	45.7	53.0	50.7	56.1	49.1				
CRANIOPHARYNGIOMA	1	0	0	0	0	0	.	.	.	.
prostate gland										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.8	50.7	56.1	49.1				
ADENOCARCINOMA	0	0	1	0	0	0	1	1	1	1
Adj. # at Risk	57.1	45.7	53.2	50.7	56.7	49.1				
ADENOMA	0	0	1	0	1	0	.6588	1	.5966	1
skin										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.0	50.7	56.1	49.1				
ADENOMA, BASAL CELL	0	0	1	0	0	1	.3505	.5571	1	1
Adj. # at Risk	57.4	45.7	53.0	50.7	56.1	49.1				
CARCINOMA, SQUAMOUS CELL	2	0	0	0	0	0	.	.	.	.
Adj. # at Risk	57.1	45.7	53.0	50.7	56.1	49.1				
HAIR FOLLICLE TUMOR	0	0	0	0	0	1	.1937	.3333	.	.
Adj. # at Risk	57.4	46.7	53.0	50.7	56.3	49.8				
KERATOACANTHOMA	3	3	1	0	1	2	.3507	.6489	.8976	1
Adj. # at Risk	57.7	46.7	53.0	50.7	56.6	49.8				
Kerato./Sq.Cell Papilloma/ Carcinoma	4	3	1	0	2	4	.0668	.2492	.7065	1
Adj. # at Risk	57.1	45.7	53.0	50.7	56.4	49.1				
PAPILLOMA, SQUAMOUS CELL	0	0	0	0	1	2	.0316	.1096	.3636	.
Adj. # at Risk	57.4	45.7	53.0	50.7	56.4	49.1				
Papilloma/Carc.Squamous Cell	2	0	0	0	1	2	.0316	.1096	.3636	.

Table A.2.3. (cont.) Tumor Incidence and Results in Male Rats

Organ/Tumor	H2O	Veh1	Veh2	Low	Med	High	ptrend	phigh vsVeh	pmed vsVeh	plow vsVeh
skin, subcutis										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	46.4	53.0	50.7	56.1	49.3				
FIBROMA	1	2	1	1	0	1	.6390	.8038	1	.8091
Adj. # at Risk	57.5	46.5	53.4	51.1	56.1	49.1				
FIBROSARCOMA	1	1	1	2	0	2	.2867	.4034	1	.4194
Adj. # at Risk	57.1	45.7	53.0	50.7	56.5	49.1				
FIBROUS HISTIOCYTOMA	0	0	0	0	2	0	.3988	.	.1307	.
Adj. # at Risk	57.1	45.7	53.0	50.7	56.8	49.9				
HEMANGIOSARCOMA	0	0	0	0	1	1	.1230	.3333	.3636	.
Adj. # at Risk	57.8	45.9	53.0	50.7	56.1	49.1				
LIPOMA	2	2	0	0	1	0	.8015	1	.7452	1
Adj. # at Risk	57.1	45.7	53.0	50.7	56.1	49.1				
LIPOSARCOMA	0	0	1	0	0	0	1	1	1	1
Adj. # at Risk	57.1	45.7	53.3	50.7	56.1	49.1				
SARCOMA, UNDIFFERENTIATED	0	0	1	0	0	0	1	1	1	1
Adj. # at Risk	57.1	45.7	53.6	50.7	56.3	49.5				
SCHWANNOMA	0	0	1	0	1	2	.0922	.2547	.5935	1
small intestine, jejunum										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.0	50.7	56.1	49.1				
ADENOCARCINOMA	1	0	0	0	0	0	.	.	.	.
spleen										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.0	51.0	56.1	49.1				
HEMANGIOSARCOMA	0	1	1	1	1	0	.8333	1	.7452	.7185
stomach, nonglandular										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.0	50.7	56.1	49.1				
CARCINOMA, SQUAMOUS CELL	1	0	0	1	0	0	.6126	.	.	.3378
testes										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.4	50.7	56.1	49.1				
ADENOMA, INTERSTITIAL CELL	2	1	1	0	0	0	1	1	1	1
thyroid gland										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	58.2	46.2	53.5	50.7	57.3	49.5				
ADENOMA, C-CELL	8	5	10	5	12	8	.3449	.5144	.2345	.8713
Adj. # at Risk	57.1	45.7	53.3	51.2	56.1	49.1				
ADENOMA, FOLLICULAR CELL	0	1	3	1	0	1	.7187	.8728	1	.8812
Adj. # at Risk	57.1	46.0	53.3	51.2	56.1	49.2				
Adenoma/Carc. Foll. Cell	1	2	3	1	0	2	.4758	.7388	1	.9216
Adj. # at Risk	58.2	46.2	53.6	50.7	57.3	49.5				
Adenoma/Carcinoma C-Cell	10	6	12	5	12	9	.3431	.5723	.4060	.9430
Adj. # at Risk	57.1	45.7	53.1	50.7	56.1	49.1				
CARCINOMA, C-CELL	2	1	3	0	0	1	.6629	.8728	1	1
Adj. # at Risk	57.1	46.0	53.0	50.7	56.1	49.2				
CARCINOMA, FOLLICULAR CELL	1	1	1	0	0	1	.4773	.7068	1	1
tongue										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.0	50.7	56.2	49.1				
CARCINOMA, SQUAMOUS CELL	0	0	0	0	1	0	.4150	.	.3636	.
urinary bladder										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.0	50.7	56.1	49.1				
GRANULAR CELL TUMOR	0	0	0	0	2	0	.3988	.	.1307	.

Table A.2.3. (cont.) Tumor Incidence and Results in Male Rats

Organ/Tumor	H2O	Veh1	Veh2	Low	Med	High	ptrend	phigh vsVeh	pmed vsVeh	plow vsVeh
zymbal`s gland										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.1	45.7	53.7	51.2	56.1	49.1				
CARCINOMA, SEBACEOUS CELL	0	0	1	1	0	0	.8502	1	1	.5659
Adj. # at Risk	57.1	45.7	53.7	51.4	56.1	49.1				
CARCINOMA, SQUAMOUS CELL	0	0	1	1	0	0	.8502	1	1	.5659
Adj. # at Risk	57.1	45.7	54.3	51.9	56.1	49.1				
Carc.Seb.Cell/Squamous Cell	0	0	2	2	0	0	.9316	1	1	.4194

Table A.2.4. Tumor Incidence and Results in Female Rats

organ/tumor	H2O	Veh1	Veh2	Low	Med	High	ptrend	phigh vsVeh	pmed vsVeh	plow vsVeh
Systemic										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.2	55.0	53.2	47.5	50.4	55.4				
HEMANGIOSARCOMA	0	0	0	0	2	0	.4154	.	.0988	.
Adj. # at Risk	53.2	55.5	53.2	47.5	50.0	55.4				
LYMPHOMA	1	1	0	0	1	0	.6428	1	.5282	1
adrenal glands										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.4	55.1	53.6	47.5	49.7	55.7				
ADENOMA, CORTICAL	1	1	3	0	0	1	.7002	.8763	1	1
Adj. # at Risk	53.4	55.1	53.6	47.5	49.7	55.7				
Adenoma/Carcinoma, Cortical	1	1	3	0	1	2	.3859	.6611	.8505	1
Adj. # at Risk	53.2	55.0	53.2	47.5	49.7	55.4				
CARCINOMA, CORTICAL	0	0	0	0	1	1	.1251	.3374	.3121	.
Adj. # at Risk	54.7	56.2	53.5	47.5	50.2	55.6				
PHEOCHROMOCYTOMA	5	3	3	1	3	2	.6031	.8150	.5795	.9234
Adj. # at Risk	53.2	55.0	53.2	47.5	49.7	55.4				
PHEOCHROMOCYTOMA, COMPLEX	0	0	1	0	0	0	1	1	1	1
Adj. # at Risk	54.7	56.2	53.5	47.5	50.2	55.6				
Pheocromocytoma, Any	5	3	4	1	3	2	.6743	.8670	.6617	.9476
brain										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	54.7	55.0	53.2	47.5	49.7	55.4				
ASTROCYTOMA	3	1	0	0	0	0	1	1	1	1
Adj. # at Risk	53.2	55.0	53.2	47.5	49.7	55.7				
GRANULAR CELL TUMOR	0	1	0	0	0	1	.3803	.5624	1	1
cavity, abdominal										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.2	55.0	53.2	48.0	49.7	55.4				
ADENOCARCINOMA(PRIMARY SITE	0	0	0	1	0	0	.5830	.	.	.3032
cavity, oral										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.2	55.6	53.2	47.5	49.7	55.4				
CARCINOMA, SQUAMOUS CELL	0	1	0	0	0	0	1	1	1	1
eyes, optic nerves										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.2	55.0	53.2	47.6	49.7	55.4				
SCHWANNOMA	0	0	0	1	0	0	.5830	.	.	.3032

Table A.2.4. (cont.) Tumor Incidence and Results in Female Rats

organ/tumor	H2O	Veh1	Veh2	Low	Med	High	ptrend	phigh vsVeh	pmed vsVeh	plow vsVeh
Kidneys										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.2	55.0	53.2	47.5	49.7	55.4				
ADENOMA, RENAL TUBULE, (AV) TYPE	0	0	0	1	0	0	.5830	.	.	.3032
Adj. # at Risk	53.2	55.0	53.2	47.5	49.7	55.4				
Adenoma, Renal Tub. Any	0	0	0	1	0	0	.5830	.	.	.3032
Adj. # at Risk	53.2	55.0	53.2	47.5	49.7	55.4				
Adenoma/Carc.Renal Tub. Any	0	1	0	1	0	0	.8271	1	1	.5159
Adj. # at Risk	53.2	55.0	53.2	47.5	49.7	55.4				
CARCINOMA, RENAL TUBULE,(AV)	0	1	0	0	0	0	1	1	1	1
Adj. # at Risk	54.1	55.0	53.2	47.5	49.7	55.4				
NEPHROBLASTOMA	1	0	0	0	0	0	.	.	.	.
liver										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.2	55.0	53.2	47.5	49.7	55.4				
ADENOMA, HEPATOCELLULAR	1	1	0	0	0	0	1	1	1	1
Adj. # at Risk	53.2	55.0	53.2	47.5	49.7	55.4				
Adenoma/Carcinoma, Hepato.	1	1	0	0	1	0	.6428	1	.5282	1
Adj. # at Risk	53.2	55.0	53.2	47.5	49.7	55.4				
CARCINOMA, HEPATOCELLULAR	0	0	0	0	1	0	.4015	.	.3121	.
Adj. # at Risk	53.2	55.0	53.2	47.5	49.7	55.4				
CHOLANGIOCARCINOMA	0	0	1	0	0	0	1	1	1	1
lung										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.2	55.0	53.2	47.5	49.7	55.4				
ADENOMA,BRONCHIOLAR ALVEOLAR	0	0	2	0	0	0	1	1	1	1
lymph node, mesenteric										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.2	55.0	53.2	47.5	49.8	55.4				
HEMANGIOSARCOMA	0	0	0	0	1	0	.4015	.	.3121	.
Adj. # at Risk	53.2	55.0	53.2	47.5	50.0	55.4				
LYMPHANGIOMA	0	0	0	0	1	0	.4038	.	.3165	.
mammary gland										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	57.8	60.0	58.2	51.5	55.3	57.8				
ADENOCARCINOMA	19	20	15	13	17	9	.9759	.9870	.5010	.7680
Adj. # at Risk	54.2	56.3	53.2	48.0	50.6	55.8				
ADENOMA	4	4	0	3	2	3	.3533	.4341	.6144	.3544
Adj. # at Risk	63.0	63.6	63.3	57.4	58.6	61.3				
Adenocarc./Adenoma/Fibro-adenoma	39	40	39	31	34	34	.7330	.8581	.7558	.8897
Adj. # at Risk	63.0	63.6	63.3	57.4	58.6	61.3				
Adenoma/Fibroadenoma	39	40	39	31	34	34	.7330	.8581	.7558	.8897
Adj. # at Risk	60.9	59.8	60.3	54.3	54.5	59.1				
FIBROADENOMA	29	27	29	21	25	27	.4335	.6072	.5824	.8689
multicentric neoplasm										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.2	55.0	53.3	47.5	49.7	55.6				
LEUKEMIA, LARGE GRANULAR LYMP	0	0	1	0	0	2	.1150	.2634	1	1
Adj. # at Risk	53.2	55.5	53.2	47.5	50.0	55.4				
LYMPHOMA	1	1	0	0	1	0	.6428	1	.5282	1
Adj. # at Risk	53.2	55.0	53.3	47.5	49.7	55.6				
Leukemia, any	0	0	1	0	0	2	.1150	.2634	1	1
Adj. # at Risk	53.3	55.5	53.2	48.0	49.8	55.4				
SARCOMA, HISTIOCYTIC	1	1	0	1	1	1	.3488	.5624	.5282	.5221

Table A.2.4. (cont.) Tumor Incidence and Results in Female Rats

organ/tumor	H2O	Veh1	Veh2	Low	Med	High	ptrend	phigh vsVeh	pmed vsVeh	plow vsVeh
ovaries										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.4	55.0	53.2	47.5	49.7	55.4				
CARCINOMA, YOLK SAC	1	0	0	0	0	0	.	.	.	.
Adj. # at Risk	53.2	55.0	53.2	47.5	49.7	55.4				
SERTOLI CELL TUMOR	0	0	0	0	1	0	.4015	.	.3121	.
pancreas										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.2	55.0	53.2	47.9	49.7	55.4				
ADENOMA, ACINAR CELL	0	0	0	1	0	0	.5830	.	.	.3032
Adj. # at Risk	54.0	55.6	53.2	48.5	49.7	55.4				
ADENOMA, ISLET CELL	5	3	2	4	3	1	.8635	.9194	.4804	.2839
Adj. # at Risk	53.2	55.0	53.2	47.9	49.7	55.4				
Adenoma/Carc. Acinar Cell	0	0	0	1	0	0	.5830	.	.	.3032
Adj. # at Risk	54.0	56.5	53.2	48.5	50.0	55.4				
Adenoma/Carc. Islet Cell	5	5	2	5	4	2	.8328	.8670	.4611	.2858
Adj. # at Risk	53.2	55.8	53.2	47.5	50.0	55.4				
CARCINOMA, ISLET CELL	0	2	0	1	1	1	.5020	.7091	.6745	.6617
parathyroid glands										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.3	55.4	53.2	47.5	49.7	55.6				
ADENOMA	1	1	1	0	0	4	.0197	.1001	1	1
pituitary gland										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	73.4	73.3	72.8	71.1	72.9	77.4				
ADENOMA, PARS DISTALIS	66	66	68	64	67	71	.4303	.5671	.4861	.7485
Adj. # at Risk	53.2	55.6	53.2	47.5	49.7	55.4				
ADENOMA, PARS INTERMEDIA	0	1	0	0	0	0	1	1	1	1
Adj. # at Risk	74.4	73.9	73.6	71.1	73.8	77.8				
Adenoma/Carc. Pars Distalis	69	68	70	65	69	72	.4937	.6629	.5569	.8229
Adj. # at Risk	74.4	74.5	73.6	71.1	73.8	77.8				
Adenoma/Carcinoma	69	69	70	65	69	72	.4973	.6676	.5617	.8263
Adj. # at Risk	54.2	55.5	54.0	47.5	50.5	55.7				
CARCINOMA, PARS DISTALIS	3	2	2	1	2	1	.6885	.8743	.6144	.8381
skin										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.2	55.0	53.4	47.5	49.7	55.4				
ADENOMA, BASAL CELL	0	0	1	0	0	0	1	1	1	1
Adj. # at Risk	53.2	55.0	53.6	47.5	50.0	55.4				
CARCINOMA, SQUAMOUS CELL	0	0	1	0	1	0	.6428	1	.5282	1
Adj. # at Risk	53.7	55.0	53.2	47.5	49.7	55.4				
KERATOACANTHOMA	1	0	0	0	0	0	.	.	.	.
Adj. # at Risk	53.7	55.0	53.6	48.1	50.0	55.7				
Kerato./Sq.Cell Papilloma/ Carcinoma	1	0	1	1	1	1	.3488	.5624	.5282	.5221
Adj. # at Risk	53.2	55.0	53.2	48.1	49.7	55.7				
PAPILLOMA, SQUAMOUS CELL	0	0	0	1	0	1	.2026	.3374	.	.3077
Adj. # at Risk	53.2	55.0	53.6	48.1	50.0	55.7				
Papilloma/Carc.Squamous Cell	0	0	1	1	1	1	.3488	.5624	.5282	.5221

Table A.2.4. (cont.) Tumor Incidence and Results in Female Rats

organ/tumor	H2O	Veh1	Veh2	Low	Med	High	ptrend	phigh vsVeh	pmed vsVeh	plow vsVeh
skin, subcutis										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.2	55.0	53.4	47.5	49.8	55.4				
FIBROMA	0	0	1	0	1	1	.2836	.5624	.5282	1
Adj. # at Risk	53.2	55.0	53.2	47.5	49.7	56.1				
FIBROSARCOMA	1	1	2	0	0	2	.2939	.5573	1	1
Adj. # at Risk	53.2	55.0	53.2	47.5	49.7	55.4				
FIBROUS HISTIOCYTOMA	1	0	0	0	0	0	.	.	.	.
Adj. # at Risk	53.2	55.0	53.2	47.5	50.3	55.4				
HEMANGIOSARCOMA	0	0	0	0	1	0	.4038	.	.3165	.
Adj. # at Risk	53.2	55.2	53.2	47.5	49.7	55.7				
LIPOMA	1	1	0	0	0	1	.3803	.5624	1	1
Adj. # at Risk	53.2	55.0	53.2	47.5	49.7	56.0				
LIPOSARCOMA	0	0	0	0	0	1	.2154	.3415	.	.
Adj. # at Risk	53.2	55.1	53.2	47.5	49.7	55.4				
OSTEOSARCOMA	0	1	0	0	0	0	1	1	1	1
small intestine, ileum										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.2	55.0	53.2	47.5	49.7	55.4				
LEIOMYOMA	0	0	0	1	0	0	.5830	.	.	.3032
stomach, glandular										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.2	55.2	53.2	47.5	49.7	55.4				
LEIOMYOMA	0	1	0	0	0	0	1	1	1	1
stomach, nonglandular										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.2	55.0	53.2	47.5	49.7	55.4				
CARCINOMA, SQUAMOUS CELL	1	0	1	0	0	0	1	1	1	1
Adj. # at Risk	53.3	55.0	53.2	47.5	49.7	55.4				
PAPILLOMA, SQUAMOUS CELL	1	0	0	0	0	0	.	.	.	.
thyroid gland										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	55.7	55.8	54.8	48.1	50.5	55.7				
ADENOMA, C-CELL	6	5	9	3	7	3	.8889	.9636	.5030	.9381
Adj. # at Risk	53.2	55.0	53.2	48.1	50.7	55.4				
ADENOMA, FOLLICULAR CELL	0	0	0	2	2	2	.1268	.1125	.0988	.0933
Adj. # at Risk	53.8	55.0	53.2	49.2	50.7	55.4				
Adenoma/Carc. Foll. Cell	1	0	2	4	2	2	.4507	.4147	.3771	.0765
Adj. # at Risk	55.8	55.8	54.8	48.6	50.5	55.7				
Adenoma/Carcinoma C-Cell	8	5	11	4	7	3	.9407	.9817	.6230	.9143
Adj. # at Risk	53.3	55.0	53.2	47.9	49.7	55.4				
CARCINOMA, C-CELL	2	0	2	1	0	0	.9287	1	1	.6646
Adj. # at Risk	53.8	55.0	53.2	48.6	49.7	55.4				
CARCINOMA, FOLLICULAR CELL	1	0	2	2	0	0	.9185	1	1	.3614

Table A.2.4. (cont.) Tumor Incidence and Results in Female Rats

organ/tumor	H2O	Veh1	Veh2	Low	Med	High	ptrend	phigh vsVeh	pmed vsVeh	plow vsVeh
uterus with cervix										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.9	55.0	53.2	47.5	49.7	55.4				
ADENOCARCINOMA	1	0	0	0	0	0	.	.	.	.
Adj. # at Risk	53.2	55.0	53.2	47.8	50.1	56.1				
GRANULAR CELL TUMOR	0	1	0	1	1	2	.1499	.2688	.5341	.5159
Adj. # at Risk	53.2	55.0	53.2	47.5	49.7	55.4				
LEIOMYOMA	0	0	0	0	0	1	.2124	.3374	.	.
Adj. # at Risk	53.2	55.0	53.2	47.5	50.0	55.4				
LEIOMYOSARCOMA	0	0	0	0	1	1	.1251	.3374	.3121	.
Adj. # at Risk	54.8	56.4	53.2	49.1	51.7	55.4				
POLYP, ENDOMETRIAL STROMAL	3	7	3	4	7	1	.9687	.9907	.2706	.6851
Adj. # at Risk	53.2	55.0	53.2	47.5	49.7	55.4				
SARCOMA, STROMAL	0	0	0	1	1	0	.4844	.	.3121	.3032
Adj. # at Risk	53.2	55.0	53.2	47.9	49.7	55.4				
SCHWANNOMA	0	0	0	1	0	0	.5830	.	.	.3032
vagina										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.2	55.0	53.3	47.5	49.7	56.4				
GRANULAR CELL TUMOR	0	0	2	0	0	4	.0208	.1042	1	1
Adj. # at Risk	53.2	55.9	53.2	47.5	50.0	55.4				
POLYP, STROMAL	0	1	0	0	1	0	.6409	1	.5254	1
zymbal`s gland										
# Evaluated	80	80	80	80	80	80				
Adj. # at Risk	53.2	55.2	53.2	47.5	49.7	55.4				
CARCINOMA, SQUAMOUS CELL	0	1	0	0	0	0	1	1	1	1
Adj. # at Risk	53.2	55.2	53.2	47.5	49.7	55.4				
Carc.Seb.Cell/Squamous Cell	0	1	0	0	0	0	1	1	1	1

Tables A.2.5 through A.2.10 provide tumor incidence and the significance levels associated with the three Tg.rasH2 studies. Note that none of the tests of trend or pairwise comparisons achieved a 0.10 level, let alone the approximate multiplicity adjusted .0.05 level.

Table A.2.5 Study 126-641 Tumor Incidence and Results in Male Mice

organ/tumor	H2O	Veh	Low	Med	High	ptrend	phigh vsVeh	pmed vsVeh	plow vsVeh
Systemic									
# Evaluated	35	35	35	35	34				
Adj. # at Risk	34.1	33.0	33.3	29.2	32.3				
HEMANGIOSARCOMA	1	2	1	1	1	.6677	.8810	.8622	.8864
Adj. # at Risk	34.1	33.0	33.3	29.2	31.4				
LYMPHOMA	0	0	1	0	0	.7440	.	.	.5077
bone, sternum									
# Evaluated	35	35	35	35	34				
Adj. # at Risk	34.1	33.0	33.3	29.2	31.4				
LYMPHOMA	0	0	1	0	0	.7440	.	.	.5077
epididymides									
# Evaluated	35	35	35	35	34				
Adj. # at Risk	34.1	33.0	33.3	29.2	31.4				
HEMANGIOSARCOMA	0	1	0	0	0	1	1	1	1
harderian glands									
# Evaluated	35	35	35	35	34				
Adj. # at Risk	34.1	33.0	33.3	29.2	31.4				
ADENOMA	0	2	0	0	0	1	1	1	1
heart									
# Evaluated	35	35	35	35	34				
Adj. # at Risk	34.1	33.0	33.3	29.2	31.4				
LYMPHOMA	0	0	1	0	0	.7440	.	.	.5077
lung									
# Evaluated	35	35	35	35	34				
Adj. # at Risk	34.1	33.0	33.3	29.2	31.4				
ADENOMA,BRONCHIOLAR ALVEOLAR	3	1	3	3	1	.6910	.7460	.2696	.3183
Adj. # at Risk	34.1	33.5	33.3	29.2	31.4				
Adenoma & Carc. Bronch. Alv.	3	2	3	3	1	.7751	.8690	.4376	.5000
Adj. # at Risk	34.1	33.5	33.3	29.2	31.4				
CARCINOMA,BRONCHIOLAR ALV.	0	1	0	0	0	1	1	1	1
multicentric neoplasm									
# Evaluated	35	35	35	35	34				
Adj. # at Risk	34.1	33.0	33.3	29.2	32.3				
HEMANGIOSARCOMA	1	2	1	1	1	.6677	.8810	.8622	.8864
Adj. # at Risk	34.1	33.0	33.3	29.2	31.4				
LYMPHOMA	0	0	1	0	0	.7440	.	.	.5077
skeletal muscle, hypaxial									
# Evaluated	35	35	35	35	34				
Adj. # at Risk	34.1	33.0	33.3	29.2	32.3				
HEMANGIOSARCOMA	0	0	0	0	1	.2540	.5000	.	.
spleen									
# Evaluated	35	35	35	35	34				
Adj. # at Risk	34.1	33.0	33.3	29.2	31.4				
HEMANGIOSARCOMA	1	2	1	1	0	.9153	1	.8622	.8864
stomach, nonglandular									
# Evaluated	35	35	35	35	34				
Adj. # at Risk	34.1	33.0	33.3	29.2	31.4				
PAPILLOMA, SQUAMOUS CELL	0	0	1	0	0	.7440	.	.	.5077
thymus									
# Evaluated	35	35	35	35	34				
Adj. # at Risk	34.1	33.0	33.3	29.2	31.4				
LYMPHOMA	0	0	1	0	0	.7440	.	.	.5077

Table A.2.6 Study 126-641 Tumor Incidence and Results in Female Mice

organ/tumor	H2O	Veh	Low	Med	High	ptrend	phigh vsVeh	pmed vsVeh	plow vsVeh
Systemic									
# Evaluated	35	35	35	35	35				
Adj. # at Risk	33.9	32.9	34.1	34.1	34.1				
HEMANGIOMA	0	0	1	0	0	.7612	.	.	.5152
Adj. # at Risk	34.0	33.8	34.1	34.1	34.1				
HEMANGIOSARCOMA	1	3	1	0	0	.9969	1	1	.9466
Adj. # at Risk	34.0	33.8	34.1	34.1	34.1				
Hemangioma/-sarcoma	1	3	2	0	0	.9953	1	1	.8314
clitoral glands									
# Evaluated	35	35	35	35	35				
Adj. # at Risk	33.9	32.9	34.1	34.3	34.1				
CARCINOMA, SQUAMOUS CELL	0	0	0	1	0	.5075	.	.5152	.
harderian glands									
# Evaluated	35	35	35	35	35				
Adj. # at Risk	33.9	32.9	34.1	34.1	34.1				
ADENOMA	1	2	1	1	0	.9307	1	.8916	.8916
lung									
# Evaluated	35	35	35	35	35				
Adj. # at Risk	33.9	32.9	34.1	34.1	34.1				
ADENOMA,BRONCHIOLAR ALVEOLAR	2	2	2	0	2	.4613	.7161	1	.7161
Adj. # at Risk	33.9	32.9	34.1	34.1	34.1				
Adenoma & Carc. Bronch. Alv.	2	3	2	0	2	.5635	.8407	1	.8407
Adj. # at Risk	33.9	32.9	34.1	34.1	34.1				
CARCINOMA,BRONCHIOLAR ALV.	0	1	0	0	0	1	1	1	1
mammary gland									
# Evaluated	35	35	35	35	35				
Adj. # at Risk	33.9	32.9	34.1	34.1	34.1				
ADENOCARCINOMA	0	0	0	0	1	.2537	.5152	.	.
multicentric neoplasm									
# Evaluated	35	35	35	35	35				
Adj. # at Risk	33.9	32.9	34.1	34.1	34.1				
HEMANGIOMA	0	0	1	0	0	.7612	.	.	.5152
Adj. # at Risk	34.0	33.8	34.1	34.1	34.1				
HEMANGIOSARCOMA	1	3	1	0	0	.9969	1	1	.9466
skin, subcutis									
# Evaluated	35	35	35	35	35				
Adj. # at Risk	33.9	32.9	34.1	34.1	34.1				
HEMANGIOMA	0	0	1	0	0	.7612	.	.	.5152
Adj. # at Risk	33.9	33.8	34.1	34.1	34.1				
HEMANGIOSARCOMA	0	1	1	0	0	.9416	1	1	.7612
spleen									
# Evaluated	35	35	35	35	35				
Adj. # at Risk	34.0	32.9	34.1	34.1	34.1				
HEMANGIOSARCOMA	1	2	0	0	0	1	1	1	1
stomach, nonglandular									
# Evaluated	35	35	35	35	35				
Adj. # at Risk	33.9	32.9	34.1	34.1	34.1				
PAPILLOMA, SQUAMOUS CELL	0	0	0	0	1	.2537	.5152	.	.
uterus with cervix									
# Evaluated	35	35	35	35	35				
Adj. # at Risk	33.9	32.9	34.1	34.1	34.1				
POLYP	0	1	0	0	0	1	1	1	1

Table A.2.7 Study 126-701 Tumor Incidence and Results in Male Mice

organ/tumor	Veh	Low	Med	High	ptrend	phigh vsVeh	pmed vsVeh	plow vsVeh
Systemic								
# Evaluated	25	25	25	25				
Adj. # at Risk	24.3	25.0	23.6	24.0				
HEMANGIOSARCOMA	2	0	0	1	.5826	.8830	1	1
harderian glands								
# Evaluated	25	25	25	25				
Adj. # at Risk	23.6	25.0	23.6	24.0				
ADENOMA	1	0	1	1	.3915	.7660	.7556	1
kidneys								
# Evaluated	25	25	25	25				
Adj. # at Risk	23.6	25.0	23.6	24.0				
HEMANGIOSARCOMA	0	0	0	1	.2526	.5106	.	.
lung								
# Evaluated	25	25	25	25				
Adj. # at Risk	23.6	25.0	23.6	24.0				
ADENOMA, BRONCHIOLAR ALVEOLAR	0	1	0	0	.7579	.	.	.5208
Adj. # at Risk	23.6	25.0	23.6	24.0				
Adenoma & Carc. Bronch. Alv.	0	2	1	0	.7375	.	.5000	.2660
Adj. # at Risk	23.6	25.0	23.6	24.0				
CARCINOMA, BRONCHIOLAR ALVEOLAR	0	1	1	0	.6289	.	.5000	.5208
Adj. # at Risk	23.6	25.0	23.6	24.0				
HEMANGIOSARCOMA	0	0	0	1	.2526	.5106	.	.
multicentric neoplasm								
# Evaluated	25	25	25	25				
Adj. # at Risk	24.3	25.0	23.6	24.0				
HEMANGIOSARCOMA	2	0	0	1	.5826	.8830	1	1
spleen								
# Evaluated	25	25	25	25				
Adj. # at Risk	24.3	25.0	23.6	24.0				
HEMANGIOSARCOMA	2	0	0	0	1	1	1	1
stomach, nonglandular								
# Evaluated	25	25	25	25				
Adj. # at Risk	23.6	25.0	23.6	24.0				
CARCINOMA, SQUAMOUS CELL	1	0	0	0	1	1	1	1

Table A.2.8 Study 126-701 Tumor Incidence and Results in Female Mice

organ/tumor	Veh	Low	Med	High	ptrend	phigh vsVeh	pmed vsVeh	plow vsveh
Systemic								
# Evaluated	25	25	25	25				
Adj. # at Risk	25.0	23.5	23.2	25.0				
HEMANGIOMA	1	0	0	0	1	1	1	1
Adj. # at Risk	25.0	24.3	23.7	25.0				
HEMANGIOSARCOMA	0	2	3	2	.2746	.2449	.1024	.2347
Adj. # at Risk	25.0	24.3	23.7	25.0				
Hemangioma/-sarcoma	1	2	3	2	.3953	.5000	.2730	.4844
harderian glands								
# Evaluated	25	25	25	25				
Adj. # at Risk	25.0	23.5	23.2	25.0				
ADENOMA	1	1	0	0	.9342	1	1	.7340
lung								
# Evaluated	25	25	25	25				
Adj. # at Risk	25.0	23.5	23.2	25.0				
ADENOMA, BRONCHIOLAR ALVEOLAR	3	0	1	0	.9574	1	.9350	1
Adj. # at Risk	25.0	23.5	23.2	25.0				
Adenoma & Carc. Bronch. Alv.	3	0	1	0	.9574	1	.9350	1

Table A.2.8 (cont.) Study 126-701 Tumor Incidence and Results in Female Mice

organ/tumor	Veh	Low	Med	High	ptrend	phigh vsVeh	pmed vsVeh	plow vsveh
lymph node								
# Evaluated	25	25	25	25				
Adj. # at Risk	25.0	23.5	23.2	25.0				
SARCOMA, HISTIOCYTIC	0	0	1	0	.5000	.	.4792	.
lymph node, inguinal								
# Evaluated	25	25	25	25				
Adj. # at Risk	25.0	23.5	23.2	25.0				
SARCOMA, HISTIOCYTIC	0	0	1	0	.5000	.	.4792	.
lymph node, mesenteric								
# Evaluated	25	25	25	25				
Adj. # at Risk	25.0	23.5	23.2	25.0				
SARCOMA, HISTIOCYTIC	0	0	1	0	.5000	.	.4792	.
multicentric neoplasm								
# Evaluated	25	25	25	25				
Adj. # at Risk	25.0	24.3	23.7	25.0				
HEMANGIOSARCOMA	0	2	3	2	.2746	.2449	.1024	.2347
Adj. # at Risk	25.0	23.5	23.2	25.0				
SARCOMA, HISTIOCYTIC	0	0	1	0	.5000	.	.4792	.
ovaries								
# Evaluated	25	25	25	25				
Adj. # at Risk	25.0	23.5	23.2	25.0				
HEMANGIOMA	1	0	0	0	1	1	1	1
pancreas								
# Evaluated	25	25	25	25				
Adj. # at Risk	25.0	23.5	23.7	25.0				
HEMANGIOSARCOMA	0	0	1	0	.5000	.	.4792	.
skin								
# Evaluated	25	25	25	25				
Adj. # at Risk	25.0	24.2	23.2	25.0				
PAPILLOMA, SQUAMOUS CELL	0	1	1	0	.6239	.	.4792	.4898
skin, subcutis								
# Evaluated	25	25	25	25				
Adj. # at Risk	25.0	23.5	23.2	25.0				
HEMANGIOSARCOMA	0	0	0	1	.2604	.5000	.	.
spleen								
# Evaluated	25	25	25	25				
Adj. # at Risk	25.0	23.5	23.7	25.0				
HEMANGIOSARCOMA	0	1	2	1	.3345	.5000	.2243	.4792
thymus								
# Evaluated	25	25	25	25				
Adj. # at Risk	25.0	23.5	23.2	25.0				
THYMOMA	2	3	1	0	.9639	1	.8670	.4592
uterus with cervix								
# Evaluated	25	25	25	25				
Adj. # at Risk	25.0	23.5	23.2	25.0				
POLYP	0	0	0	1	.2604	.5000	.	.
vagina								
# Evaluated	25	25	25	25				
Adj. # at Risk	25.0	24.3	23.2	25.0				
HEMANGIOSARCOMA	0	1	1	0	.6239	.	.4792	.4898

Table A.2.9 Study 126-712 Tumor Incidence and Results in Male Mice

organ/tumor	H2O	Veh	Low	Med	High	ptrend	phigh vsVeh	pmed vsVeh	plow vsveh
Systemic									
# Evaluated	25	25	25	25	25				
Adj. # at Risk	24.7	25.0	24.0	25.0	24.2				
HEMANGIOSARCOMA	0	1	0	1	1	.3882	.7449	.7551	1
harderian glands									
# Evaluated	25	25	25	25	25				
Adj. # at Risk	24.7	25.0	24.0	25.0	24.2				
ADENOMA	1	0	1	1	0	.6211	.	.5000	.4898
lung									
# Evaluated	25	25	25	25	25				
Adj. # at Risk	24.7	25.0	24.0	25.0	24.2				
ADENOMA, BRONCHIOLAR ALVEOLAR	1	3	2	1	3	.3550	.6465	.9451	.8129
Adj. # at Risk	24.7	25.0	24.0	25.0	24.2				
Adenoma & Carc. Bronch. Alv.	1	3	2	1	4	.1850	.4762	.9451	.8129
Adj. # at Risk	24.7	25.0	24.0	25.0	24.2				
CARCINOMA, BRONCHIOLAR ALVEOL	0	0	0	0	1	.2449	.4898	.	.
multicentric neoplasm									
# Evaluated	25	25	25	25	25				
Adj. # at Risk	24.7	25.0	24.0	25.0	24.2				
HEMANGIOSARCOMA	0	1	0	1	1	.3882	.7449	.7551	1
Adj. # at Risk	24.7	25.0	24.0	25.0	24.2				
SARCOMA, HISTIOCYTIC	0	0	0	0	1	.2449	.4898	.	.
spleen									
# Evaluated	25	25	25	25	25				
Adj. # at Risk	24.7	25.0	24.0	25.0	24.2				
SARCOMA, HISTIOCYTIC	0	0	0	0	1	.2449	.4898	.	.
stomach, nonglandular									
# Evaluated	25	25	25	25	25				
Adj. # at Risk	24.7	25.0	24.0	25.0	24.2				
PAPILLOMA, SQUAMOUS CELL	0	0	0	1	0	.5000	.	.5000	.
testes									
# Evaluated	25	25	25	25	25				
Adj. # at Risk	24.7	25.0	24.0	25.0	24.2				
HEMANGIOSARCOMA	0	1	0	1	1	.3882	.7449	.7551	1
thymus									
# Evaluated	25	25	25	25	25				
Adj. # at Risk	24.7	25.0	24.0	25.0	24.2				
THYMOMA	1	0	0	1	0	.5000	.	.5000	.

Table A.2.10 Study 126-712 Tumor Incidence and Results in Female Mice

organ/tumor	H2O	Veh	Low	Med	High	ptrend	phigh vsVeh	pmed vsVeh	plow vsveh
Systemic									
# Evaluated	25	25	25	25	25				
Adj. # at Risk	24.2	24.0	23.8	25.0	25.0				
HEMANGIOSARCOMA	4	3	2	2	2	.6440	.8384	.8384	.8130
bone marrow, femur									
# Evaluated	25	25	25	25	25				
Adj. # at Risk	23.3	24.0	23.8	25.0	25.0				
HEMANGIOSARCOMA	0	0	0	0	1	.2577	.5102	.	.
bone marrow, sternum									
# Evaluated	25	25	25	25	25				
Adj. # at Risk	24.2	24.0	23.8	25.0	25.0				
HEMANGIOSARCOMA	1	0	0	0	0	.	.	.	.

Table A.2.10 (cont.) Study 126-712 Tumor Incidence and Results in Female Mice

organ/tumor	H2O	Veh	Low	Med	High	ptrend	phigh vsVeh	pmed vsVeh	plow vsveh
cavity, thoracic									
# Evaluated	25	25	25	25	25				
Adj. # at Risk	24.2	24.0	23.8	25.0	25.0				
HEMANGIOSARCOMA	1	0	0	0	0	.	.	.	.
harderian glands									
# Evaluated	25	25	25	25	25				
Adj. # at Risk	23.3	24.0	23.8	25.0	25.0				
ADENOMA	0	1	0	2	0	.7067	1	.5156	1
heart									
# Evaluated	25	25	25	25	25				
Adj. # at Risk	24.2	24.0	23.8	25.0	25.0				
HEMANGIOSARCOMA	1	0	0	0	0	.	.	.	.
large intestine, rectum									
# Evaluated	25	25	25	25	25				
Adj. # at Risk	23.3	24.0	23.8	25.0	25.0				
HEMANGIOSARCOMA	0	0	1	0	0	.7526	.	.	.4894
lung									
# Evaluated	25	25	25	25	25				
Adj. # at Risk	24.1	24.0	23.8	25.0	25.0				
ADENOMA, BRONCHIOLAR ALVEOLAR	1	0	0	1	0	.5155	.	.5102	.
multicentric neoplasm									
# Evaluated	25	25	25	25	25				
Adj. # at Risk	24.2	24.0	23.8	25.0	25.0				
HEMANGIOSARCOMA	4	3	2	2	2	.6440	.8384	.8384	.8130
skin									
# Evaluated	25	25	25	25	25				
Adj. # at Risk	23.3	24.0	24.0	25.0	25.0				
CARCINOMA, SQUAMOUS CELL	0	0	1	0	0	.7551	.	.	.5000
spleen									
# Evaluated	25	25	25	25	25				
Adj. # at Risk	23.3	24.0	23.8	25.0	25.0				
HEMANGIOSARCOMA	3	2	0	1	1	.5861	.8901	.8901	1
thymus									
# Evaluated	25	25	25	25	25				
Adj. # at Risk	24.2	24.0	23.8	25.0	25.0				
HEMANGIOSARCOMA	1	0	0	0	0	.	.	.	.
Adj. # at Risk	23.3	24.0	23.8	25.0	25.0				
THYMOMA	2	3	3	2	2	.7161	.8384	.8384	.6460
urinary bladder									
# Evaluated	25	25	25	25	25				
Adj. # at Risk	23.3	24.0	23.8	25.0	25.0				
HEMANGIOSARCOMA	0	1	1	0	0	.9407	1	1	.7447
Adj. # at Risk	23.3	24.0	23.8	25.0	25.0				
LEIOMYOSARCOMA	0	0	1	0	0	.7526	.	.	.4894
uterus with cervix									
# Evaluated	25	25	25	25	25				
Adj. # at Risk	23.3	24.0	23.8	25.0	25.0				
HEMANGIOSARCOMA	0	0	1	0	0	.7526	.	.	.4894
vagina									
# Evaluated	25	25	25	25	25				
Adj. # at Risk	23.3	24.0	23.8	25.0	25.0				
HEMANGIOSARCOMA	0	0	1	1	0	.6390	.	.5102	.4894

Appendix 3. References

Bailer, A. and Portier, C. (1988), "Effects of Treatment-Induced Mortality on Tests for Carcinogenicity in Small Samples", *Biometrics*, **44**, 4, 417-431.

Bieler, G.S., and Williams, R.L. (1993), "Ratio Estimates, the Delta Method, and Quantal Response Tests for Increased Carcinogenicity", *Biometrics*, **49**, 4, 793-801.

Brix, A.E., Hardisty, J.F., and McConnell, E.E. (2010), "Combining Neoplasms for Evaluation of Rodent Carcinogenesis Studies", *Cancer Risk Assessmen*, ed. Tsu, C-H & Stedeford, T., pp 699-715. John Wiley, New York.

Chu, K.C., Ceuto, C., and Ward, J.M. (1981), "Factors in the Evaluation of 200 National Cancer Institute Carcinogen Bioassays", *Journal of Toxicology and Environmental Health*, **8**, 251-280.

Greaves, P. (2007), "Neoplasia of Adrenal Medulla," In: *Histopathology of Preclinical Toxicity Studies.* 3rd edition, pp 819. Oxford, UK: Academic Press, Elsevier Inc.

Haseman, J. K. (1983), "A Reexamination of False-positive Rates for Carcinogenicity Studies", *Fundamental and Applied Toxicology*, **3**, 334-339.

Jara, A. (2007), "Applied Bayesian Non- and Semi-parametric Inference using DPpackage", *Rnews*, **7**, 3, 17-26.

Lin, K. K. and Ali, M.W. (2006), "Statistical Review and Evaluation of Animal Tumorigenicity Studies", *Statistics in the Pharmaceutical Industry, Third Edition*, edited by C.R. Buncher and J.Y. Tsay, Marcel Dekker, Inc. New York.

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.

McConnell, E.E., Solleveld, H.A., Swenberg, J.A., and Boorman, G.A. (1986), "Guidelines for Combining Neoplasms for Evaluation of Rodent Carcinogenesis Studies", *Journal of the National Cancer Institute*. **76**, 283-289.

Parola, A, and Jacobs, A. (2010). "Combining Tumors for Statistical Analysis", online FDA handout.

Peto, R., Pike, M.C., Day, N.E., Gray, R.G., Lee, P.N., Parrish, S., Peto, J., Richards, S., and Wahrendorf, J. (1980). "Guidelines for sample sensitive significance tests for carcinogenic effects in long-term animal experiments", *IARC Monographs on the Evaluation of Carcinogenic*

Risk of Chemicals to Humans, supplement 2: Long term and Short term Screening Assays for Carcinogens: A Critical Appraisal, International Agency for Research Against Cancer, 311-426.

R Development Core Team (2009). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

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**, 949-958.

STP Peto Working Group (2002), "Statistical Methods for Carcinogenicity Studies", *Toxicologic Pathology*. **30** (3), 403-414.

U.S. Department of Health and Human Services (2013), Guidance for Industry Statistical Aspects of the Design, Analysis, and Interpretation of Chronic Rodent Carcinogenicity Studies of Pharmaceuticals (DRAFT GUIDANCE), Center for Drug Evaluation and Research, Food and Drug Administration

---

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

/s/

STEVEN F THOMSON
09/04/2014
Statistical Carcinogenicity Review

KARL K LIN
09/04/2014
Concur with review

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

NDA Number: 206619 **Applicant:** AbbVie **Stamp Date:** April 21, 2014

Drug Name: Ombitasvir/ABT-450/Ritonavir tablets copackaged with Dasabuvir tablets

NDA/BLA Type: Priority with 8 month clock for this NME

EDR Location: \\CDSESUB1\evsprod\NDA206619\206619.enx

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.	✓			
2	ISS, ISE, and complete study reports are available (including original protocols, subsequent amendments, etc.)	✓			
3	Safety and efficacy were investigated for gender, racial, and geriatric subgroups investigated (if applicable).	✓			
4	Data sets in EDR are accessible and do they conform to applicable guidances (e.g., existence of define.pdf file for data sets).	✓			A sampling of the datasets suggests no stat issues

IS THE STATISTICAL SECTION OF THE APPLICATION FILEABLE? __YES_____

If the NDA/BLA is not fileable from the statistical perspective, state the reasons and provide comments to be sent to the Applicant. No filing issues from a statistical perspective.

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

Content Parameter (possible review concerns for 74-day letter)	Yes	No	NA	Comment
Designs utilized are appropriate for the indications requested.	✓			
Endpoints and methods of analysis are specified in the protocols/statistical analysis plans.	✓			
Interim analyses (if present) were pre-specified in the protocol and appropriate adjustments in significance level made. DSMB meeting minutes and data are available.			✓	One Phase 2 study was an adaptive design
Appropriate references for novel statistical methodology (if present) are included.			✓	
Safety data organized to permit analyses across clinical trials in the NDA/BLA.	✓			Recoding of the numeric codes for arms is necessary to do analyses using both ISE and ISS data
Investigation of effect of dropouts on statistical analyses as described by applicant appears adequate.	✓			Flanking Imputation

File name: 5_Statistics Filing Checklist for a New NDA_BLA110207

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

A table summarizing the six Phase 3 trials and schematics illustrating the trial designs are found on the next five pages of this document. In addition to reviewing the six Phase 3 trials, this statistical reviewer plans to review the Phase 2 trial M11-652; this trial has 14 arms with a sample size of 438 subjects and about 40 patients in each arm.

Joy Mele	5/22/14
Reviewing Statistician	Date
Greg Soon	
Supervisor/Team Leader	Date

File name: 5_Statistics Filing Checklist for a New NDA_BLA110207

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

NDA 206619 Summary of Phase 3 trials to assess safety & efficacy of AbbVie 3 DAA regimen to treat HCV

Study OL or DB	3DAA?	RBV?	ITT N Duration	Patient Population	Primary Efficacy Endpoint and Comparison	Stratifier	Applicant's Findings SVR_{12} Rates (95%CI)
M11-646 SAPPH-I DB	Y	Y	473 12 wks	GT 1 Treatment Naïve No cirrhotic subjects	SVR_{12} Compare to historical SVR rate of telaprevir plus pegIFN/ RBV therapy Threshold of 70% for NI Threshold of 80% for superiority	GT 1a and non-1a IL28B genotype CC or non-CC	3DAA+RBV: 455/473 96% (94.5%, 98%)
	PLA	PLA	158 12 wks				
M13-098 SAPPH-II DB	Y	Y	297 12 wks	GT 1 Treatment Experienced No cirrhotic subjects	SVR_{12} Compare to historical SVR rate of telaprevir plus pegIFN/ RBV therapy Threshold of 60% for NI Threshold of 70% for superiority	GT 1a and non-1a null-responder, non-responder/partial responder, or relapser	3DAA+RBV: 286/297 96% (94%, 98%)
	PLA	PLA	97 12 wks				
M13-099 TURQ-II OL	Y	Y	208 12 wks	GT 1 Treatment Naïve & Experienced All cirrhotic subjects	SVR_{12} Compare to historical SVR rate of telaprevir plus pegIFN/ RBV therapy Threshold of 43% for NI Threshold of 54% for superiority Secondary analysis planned to compare two treatment arms	treatment-naïve subjects stratified by GT 1a and non-1a and IL28B genotype (CC, non-CC) treatment-experienced subjects stratified by GT 1a and non-1a null responder, partial responder, or relapser	97.5% CI 3DAA+RBV 12 weeks: 191/208 92% (88%, 96%) 3DAA+RBV 24 weeks: 165/172 96% (93%, 99%)
	Y	Y	172 24 wks				

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

Study OL or DB	3DAA?	RBV?	ITT N Duration	Patient Population	Primary Efficacy Endpoint and Comparison	Stratifier	Applicant's Findings SVR_{12} Rates (95%CI)
M13-389 PEARL-II OL	Y	Y	88 12 wks	GT1b Treatment Experienced No cirrhotic subjects	SVR_{12} Compare to historical SVR rate of telaprevir plus pegIFN/ RBV therapy Threshold of 64% for NI Threshold of 75% for superiority Test arm without RBV first Secondary SVR_{12} analysis of 2 arms	By type of responder null-responder, non-responder/partial responder, or relapser	3DAA+RBV: 85/88 97% (93%, 100%) 3DAA: 91/91 100% (96%, 100%)
	Y	PLA	91 12wks				
M13-961 PEARL-III DB	Y	Y	210 12 wks	GT1b Treatment Naïve No cirrhotic subjects	SVR_{12} Compare to historical SVR rate of telaprevir plus pegIFN/ RBV therapy Threshold of 73% for NI Threshold of 84% for superiority Test arm without RBV first Secondary SVR_{12} analysis of 2 arms	by IL28B genotype (CC versus non-CC)	3DAA+RBV: 209/210 99.5% (99%, 100%) 3DAA: 207/209 99% (98%, 100%)
	Y	PLA	209 12 wks				
M14-002 PEARL-IV DB	Y	Y	100 12 wks	GT1a Treatment Naïve No cirrhotic subjects	SVR_{12} Compare to historical SVR rate of telaprevir plus pegIFN/ RBV therapy Threshold of 65% for NI Threshold of 75% for superiority Test arm without RBV first Secondary SVR_{12} analysis of 2 arms	by IL28B genotype (CC versus non-CC)	3DAA+RBV: 97/100 97% (94%, 100%) 3DAA: 185/205 90% (86%, 94%)
	Y	PLA	205 12 wks				

ITT population includes patients randomized & taking at least one dose of randomized treatment 3DAA = ABT-450/r(150/100)+ABT-267(25)+ABT-333 (250)
PLA= placebo r =Ritonavir pegIFN = pegylated interferon RBV = ribavirin wt based dosing < 75 kg 1000mg BID or ≥ 75 kg 1200mg BID
SVR_{12} = sustained virologic response [HCV RNA < lower limit of quantification (LLOQ) of 25 IU/mL 12 weeks after the last actual dose of study drug]
CC = IL28B genotype; *CC genotypes may be easier to treat than non-CC genotypes*

File name: 5_Statistics Filing Checklist for a New NDA_BLA110207

Study design schematics for the 6 Phase 3 trials

Study M13-098 and Study M11-646 have the same trial design

Figure 1. Study M11-646 Schematic

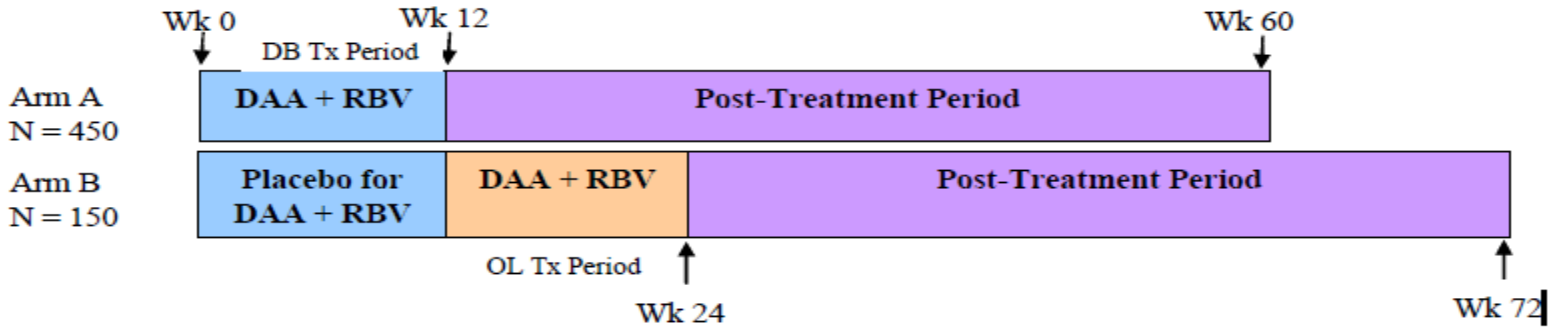
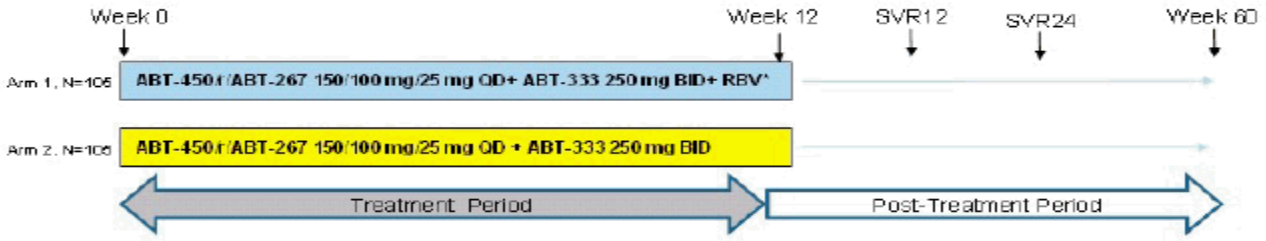


Figure 1. Study Design of Study M13-389



Study M13-099

Figure 1. Study Schematic

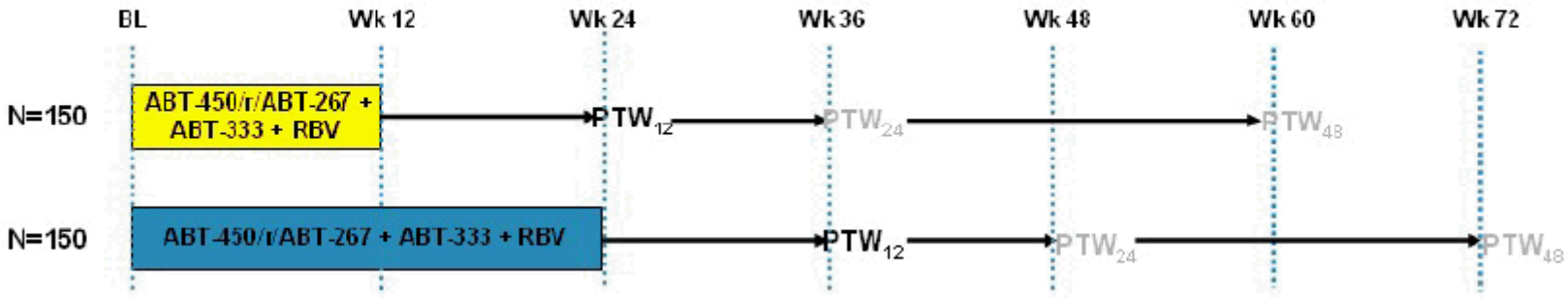


Figure 1. Study Design of Study M13-961

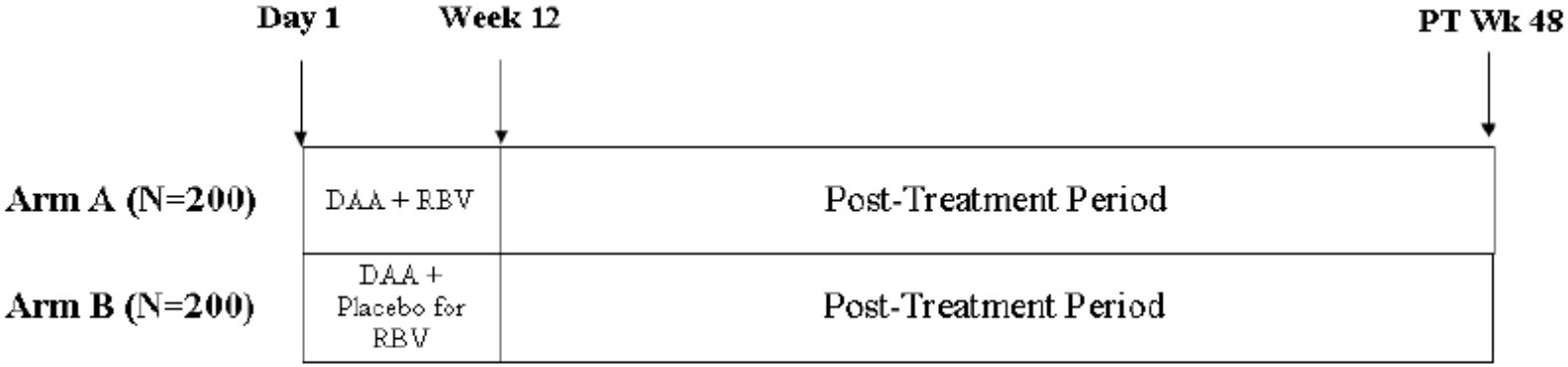
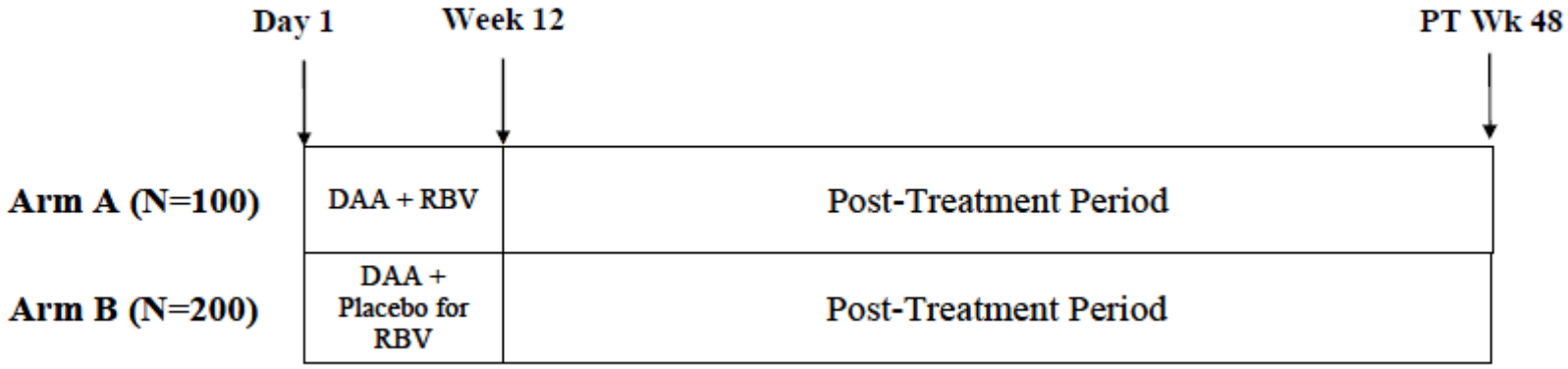


Figure 1. Study Design of Study M14-002



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

/s/

JOY D MELE
05/29/2014

GUOXING SOON
06/12/2014