CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER: NDA 20-671/S-004

STATISTICAL REVIEW(S)

CATTERSON

Statistical Review and Evaluation

JUN 4 1998

NDA Number:	20,671 suppl. SE1
Applicant:	SmithKline Beecham
Name of Drug:	Hycamtin (topotecan HCl)
Indication:	Treatment of small cell lung cancer
Documents Reviewed:	Vols. 4.157-158, 4.176, 4.185, 4.194, 4.212 submission dated 5 Dec 1997; Various electronic volumes in the SmithKline Beecham Hycamtin CANDA, submission dated 22 Jan 1998
Medical Reviewer:	Steven Hirschfeld, M.D.
Statistical Reviewer:	David Smith, Ph.D.

APPEARS THIS WAY
ON ORIGINAL

1. Background and Overview

In order to support labeling for the indication of treatment of small cell lung cancer, the sponsor submitted a supplemental NDA which is comprised of one Phase III trial and five Phase II trials. The sponsor's submission included the reports of the three pivotal Phase II studies, one pivotal Phase III study and two supportive Phase II studies. We will only consider the pivotal studies in this review.

A brief summary of the pivotal studies appears below.

Study	Type	N	Arms
090	Randomized Ph III	107	Hycamtin
		104	CAV*
014	Non-comparative Ph II	101	Hycamtin
053	Non-comparative Ph II	99	Hycamtin
092	Non-comparative Ph II	119	Hycamtin

* "CAV" herein is 1000 mg/m^2 cyclophosphamide, 45 mg/m^2 doxorubicin and 2 mg vincristine.

The next section includes relevant statistical issues for these studies. The following sections will discuss these studies in more detail and will follow the following format:

1. General description of study
2. Efficacy endpoints and results
3. Summary and conclusions

The last two sections will include an integrated summary of efficacy and overall conclusions and recommendations for the submission.

References will follow the review.

2. Statistical Issues

- In the Phase II trial 092, the protocol stated that if none of the first 14 patients enrolled in either the chemotherapy-sensitive or refractory group failed to respond, then no more patients would be enrolled in that respective group. One patient in the chemotherapy-sensitive group responded, but no patients in the refractory group responded. In accordance with the protocol, no further patients should have been enrolled to the refractory group. However, refractory patients continued to be enrolled.
- In the Phase III Trial (090), although two primary endpoints were considered, there was no corresponding adjustment in Type I error apportionment. If efficacy is to be concluded by a statistically significant finding on only one of the two primary endpoints, then there needs to be a corresponding adjustment for Type I error. However, it is not stated in the protocol whether the co-primary endpoints should be considered simultaneously or independently.
- In study 090, improvement in symptoms was assessed and analyzed using a Pearson's χ^2 statistic. This analysis relies heavily on the assumption that patients are missing completely at random and the effects of repeated measurements may be ignored. Additionally, there were large lags between consecutive visits. The sponsor originally planned on performing such a longitudinal analysis by utilizing Generalized Estimating Equation (GEE) methodology. However, convergence problems were encountered due to the prevalence of missing observations, necessitating the Pearson's χ^2 alternative analyses.
- Other than objective response, it is not clear which endpoints in Study 090 should be tested with a hypothesis of superiority or a hypothesis of non-inferiority. If one tests for superiority and concludes the null hypothesis, and then tests for non-inferiority, additional Type I error

adjustments must be made. The sponsor only tested for superiority for endpoints other than objective response, and so conclusions about non-inferiority of Hycamtin may not be justified.

3. Description and Efficacy of Studies 014, 053, and 092

These three pivotal studies enrolled a total of 319 patients with recurrent small cell lung cancer who had been previously treated with first-line chemotherapy. Each of these studies were open-label, non-comparative Phase II trials.

3.1. Description of Studies 014, 053, and 092

Study Objectives: To determine if partial or complete responses could be achieved with topotecan in small cell lung cancer in patients who were either refractory or sensitive to prior chemotherapy; to assess the probability of an actual response rate warranting further evaluation of the therapeutic effectiveness of topotecan; to further characterize the toxic effects of topotecan in this group of patients.

Study Enrollment Periods:

014: December 1992 - February 1996
053: June 1994 - June 1996
092: September 1993 - April 1995

Study Design: Open label, non-comparative Phase II. No stratification variables were defined.

Sample Size: Patients fell into two groups: those who were refractory to prior chemotherapy and those who were sensitive to prior chemotherapy.

The planned enrollment for Study 053 was 100 patients. With this number of patients, the 95% confidence interval would provide an estimate which was within 9% of the true response rate in the overall study population, if the true response rate was 35%.

For each group (either refractory or sensitive to chemotherapy), a minimum response rate of 20% was assumed. For Studies 014 and 092, sample size calculations for objective response were based on Gehan (1961). Initially, 14 patients were to be entered into each group in the trial. If none of the first 14 evaluable patients responded, the treatment was to be rejected as being less than 20% effective with a chance of false rejection error of 5%. If at least one responder was observed in the first 14 patients (CR or PR), then an additional 26 evaluable patients were to be entered for a total of 40 evaluable patients in each study group so that the true response rate of topotecan could be estimated with a standard error of no greater than 8%.

Interim Analysis: No adjustments to Type I error were necessary since each Phase II trial was non-comparative.

014: Out of the first 14 evaluable patients in the chemo-sensitive group, three patients showed a partial response and one patient showed a complete response. In the refractory group, two out of the first 14 evaluable patients showed a partial response. Therefore, the interim condition was satisfied and the sponsor enrolled the specified number of patients.

092: Out of the first 14 evaluable patients in the chemo-sensitive group, one patient showed a partial response. No patients out of the original 14 in the refractory group responded. The trial continued to enroll patients to the refractory group despite the interim analysis plan outlined in the protocol.

Dose: Hycamtin was administered at a dose of 1.5 mg/m^2/day as a 30 minute infusion for 5 days every 21 days. The number of courses of treatment depended on disease response.

Criteria for Evaluation: The primary efficacy endpoint was the proportion of patients with a complete or partial response; the secondary endpoints were time to response, response duration, time to progression and survival. On study 053, the patients assessed their disease symptoms at baseline and again prior to every course of therapy on the following symptom subscales: Shortness of Breath, Cough, Chest Pain, Coughing Up Blood, Loss of Appetite, Interference with Sleep, Hoarseness. Each subscale was assessed on the following 4-point ordinal scale: "Not at All," "A Little," "Quite a Bit," and "Very Much."

3.2. Efficacy Endpoints and Results of Studies 014, 053, and 092

Objective Response

The primary endpoint for each of these Phase II trials was objective response rate, and Tables 3.1-3.3 show the sponsor's results. In study 014, 18 out of 101 patients (18 %) of the intent-to-treat population had a complete or partial response. In studies 053 and 092, 9 patients each (8-9%) had a complete or partial response.

In patients who were considered to be sensitive to prior chemotherapy, the response rates for study 014, 053, and 092 were 31%, 15%, and 11% respectively. For patients refractory to prior chemotherapy, the response rates for study 014, 053, and 092 were 7%, 2%, and 2% respectively.

Table 3.1. Study 014: Best objective response, intent-to-treat population.

Best Response	Refractory		Sensitive		All Pts.	
Complete	2	4 %	5	11 %	7	7 %
Partial	2	4 %	9	20 %	11	11 %
Total Response	**4**	**7 %**	**14**	**31 %**	**18**	**18 %**
Stable Disease	11	20 %	8	18 %	19	19 %
Progression	33	60 %	20	44 %	54	54 %
Unknown	7	13 %	3	7 %	10	10 %
Total Non-Resp.	**51**	**93 %**	**31**	**69 %**	**83**	**82 %**

Table 3.2. Study 053: Best objective response, intent-to-treat population.

Best Response	Refractory		Sensitive		All Pts.	
Complete	0	0 %	2	4 %	2	2 %
Partial	1	2 %	6	12 %	7	7 %
Total Response	**1**	**2 %**	**8**	**15 %**	**9**	**9 %**
Stable Disease	11	23 %	14	27 %	25	25 %
Progression	30	64 %	22	42 %	52	53 %
Unknown	5	11 %	8	15 %	13	13 %
Total Non-Resp.	**46**	**98 %**	**44**	**85 %**	**90**	**91 %**

Table 3.3. Study 092: Best objective response, intent-to-treat population.

Best Response	Refractory		Sensitive		All Pts.	
Complete	0	0 %	3	4 %	3	3 %
Partial	1	2 %	5	7 %	6	5 %
Total Response	**1**	**2 %**	**8**	**11 %**	**9**	**8 %**
Stable Disease	2	4 %	14	20 %	16	13 %
Progression	37	77 %	43	61 %	80	67 %
Unknown	8	17 %	6	9 %	14	12 %
Total Non-Resp.	**47**	**98 %**	**63**	**89 %**	**110**	**92 %**

3.3. Secondary Endpoints

In all studies, the time to response was calculated from the time of the first Hycamtin infusion until the time of initial documented response. Similarly, time to progression was calculated from the time of the first Hycamtin infusion until documented progression or until an alternate therapy was started. Survival was measured from the time of first Hycamtin infusion until death of any cause. Duration of response was calculated as the time from the first response until disease progression.

In these non-comparative Phase II trials, time to event endpoints were considered to be secondary. Descriptive statistics on the time to event endpoints in the intent-to-treat population appear in Table 3.4.

Table 3.4. Time to event endpoints (in weeks) for the intent-to-treat population. The summary statistics presented are in the form: median (first quartile - third quartile). "Sensitive" refers to chemo-sensitive patients.

Study		Time to Resp.	Resp. Duration	Time to Prog.	Survival
014	Refractory	5.6 (1.6-5.7)	30.7 (20.6-47.1)	8.3 (4.1-17.1)	20.9 (7.4-35.1)
	Sensitive	7.0 (5.4-9.4)	20.7 (13.3-38.6)	17.7 (8.4-27.4)	35.6 (20.7-59.6)
	All Pts.	6.3 (5.4-9.4)	30.7 (17.0-38.6)	11.9 (6.4-22.0)	26.0 (13.1-41.4)
053	Refractory	5.4 (5.4-5.4)	24.9 (24.9-24.9)	9.6 (5.1-14.1)	21.4 (11.0-32.6)
	Sensitive	6.1 (5.1-6.1)	23.1 (14.6-31.3)	13.1 (7.3-26.7)	28.3 (16.0-51.3)
	All Pts.	6.1 (5.4-6.6)	23.7 (21.7-31.1)	11.3 (6.1-19.6)	27.3 (15.0-39.6)
092	Refractory	5.7 (5.7-5.7)	22.0 (22.0-22.0)	6.4 (4.0-10.0)	15.9 (9.6-29.0)
	Sensitive	5.6 (3.9-6.6)	21.9 (12.6-23.3)	10.3 (5.9-18.6)	26.4 (13.7-51.0)
	All Pts.	5.7 (5.4-6.6)	22.0 (17.1-23.3)	8.1 (5.3-14.1)	21.7 (10.9-42.9)

In study 014, there were 4 refractory patients and 14 chemotherapy-sensitive patients who responded. The refractory patients who responded had a shorter time to response and longer duration of response. In studies 053 and 092, there was little difference between refractory and chemo-sensitive patients with respect to time to response and response duration, although these estimates are based on small numbers of patients. In study 014, there is a larger difference between refractory and chemo-sensitive patients with respect to time to progression and survival than the other two studies. It should be noted that these results must be interpreted with caution because of the small amount of data in which these statistics are based.

Patients, particularly those who were chemotherapy-sensitive, generally had longer duration of response, survival, and time to progression in study 014 than those patients in both 053 and 092. This may be due to the fact that patients in study 014 received a larger median cumulative dose than patients in the other two studies (see Table 3.5).

Table 3.5. Cumulative doses (mg/m^2) in each of the intent-to-treat populations in each of the Phase II studies.

Study	Population	Mean	Median	Min	Max
014	Refractory	24.32	15.00		
	Sensitive	38.83	37.15		
	All Pts.	30.62	26.25		
053	Refractory	27.21	20.00		
	Sensitive	32.02	25.00		
	All Pts.	29.74	22.50		
092	Refractory	18.46	15.00		
	Sensitive	24.22	22.50		
	All Pts.	21.89	15.00		

Study 053 assessed symptoms of shortness of breath, cough, chest pain, hemoptysis, anorexia, insomnia and hoarseness. Each symptom was scored on a 4 point ordinal scale: "Not at All," "A Little," "Quite a Bit," and "Very Much." The sponsor defined an improvement in disease-related symptoms as two consecutive visits at which a symptom score indicated improvement in one ordinal point over the baseline assessment. The baseline assessments appear in Table 3.6. The sponsor examined the changes from baseline in terms of improvement, no change, or worsening for both chemo-sensitive and refractory patients. These frequencies appear in Table 3.7.

Table 3.6. Frequency of symptom occurrence at baseline, Study 053.

		Sensitive Patients				Refractory Patients			
Symptom	$n_{baseline}$	Not at All	A Little	Quite a Bit	Very Much	Not at All	A Little	Quite a Bit	Very Much
Short of breath	93	12	18	13	5	12	22	10	1
Cough	92	13	27	6	1	9	22	11	3
Chest Pain	91	29	12	5	1	28	14	2	0
Hemoptysis	93	47	1	0	0	41	3	0	1
Anorexia	91	23	10	9	5	26	8	5	5
Insomnia	91	27	11	6	2	25	12	5	3
Hoarseness	92	33	12	1	1	32	10	2	1

$n_{baseline}$ is the number of patients with baseline assessments for each symptom.

Table 3.7. Sponsor's tabulation of changes from baseline to follow-up assessments (all patients).

		Sensitive Patients			Refractory Patients		
Symptom	$n_{baseline+}$	Worsened	Unchanged	Improved	Worsened	Unchanged	Improved
Short of breath	64	8	14	8	13	16	5
Cough	65	9	16	6	6	19	9
Chest Pain	65	6	19	6	6	19	9
Hemoptysis	59	0	25	0	2	32	0
Anorexia	64	8	15	8	8	17	8
Insomnia	64	8	13	9	7	16	11
Hoarseness	63	3	23	3	5	24	5

$n_{baseline+}$ is the number of patients with baseline and at least one post-baseline assessment.

Although the collection and tabulation of this information is useful for determining the extent of palliation Hycamtin patients experience, it is difficult to make formal inferences on these measurements because of the small number of observations.

3.4. Summary and Conclusions

The results of these open-labeled non-comparative Phase II trials show that the objective response rates range between 8% to 18% in the intent-to-treat populations across all three studies. When one considers only those patients who were chemotherapy-sensitive, the response rate across the three studies ranged between 11% to 31%. Refractory patients had response rates that ranged between 2% and 11% across the three studies.

Time to event endpoints varied across these three studies, but were generally consistent. Because the number of refractory patients who responded was small, one should understand that the estimates of the time to event endpoints for refractory patients are subject to large amounts of variability.

There may be an association between cumulative dose and the endpoints that we presented in this section. Study 014, the study with the highest proportion of responses and the "best" time to event endpoints, also had patients with the highest median cumulative dose. The nature of this association is difficult to quantify formally, however.

Symptom of disease measurements were collected in Study 053. Although such an assessment is useful, formal inference, such as comparing symptoms during follow-up to baseline symptoms, should be discounted since there are very few patients in which to base such inference.

4. Study 090

4.1. Description of Study 090

Study Objective: To evaluate response rate and response duration of patients with small cell lung cancer that have failed first-line therapy.

Study Enrollment Period: June 1995 - March 1997

Study Design: Open label, multi-center randomized Phase III study. The stratification factors were baseline performance status and extent of disease at relapse.

Sample Size: Out of 211 patients randomized, 107 were assigned to the Hycamtin arm and 104 were assigned to the CAV arm. All patients were determined to be sensitive to chemotherapy.

Sample size calculations for this study were based on the method described by Makuch and Simon (1978). The sponsor assumed that the response rates following second-line treatment with CAV or Hycamtin in patients with small-cell lung cancer were 28% and 33%, respectively. Therefore, a sample of 200 patients (100 per group) provides at least a 90% probability that the lower 95% (two-sided) confidence limit for the difference between treatments excludes values larger than 14 percentage points in favor of CAV. These assumptions were based on the sponsor's experience with Hycamtin, the highest reported response rate for CAV as second-line treatment of patients with small-cell lung cancer, and the magnitude of a clinically important difference in this population. This design is similar to that of a non-inferiority trial for objective response, although the protocol did not specify a test for non-inferiority for duration of response or any other endpoint.

Interim Analysis: The sponsor conducted a preliminary analysis of the clinical data available on all patients who were randomized for at least 20 weeks before 18 November 1996. In December 1996, an interim analysis of 125 patients with at least 20 weeks of treatment determined that the median time to

progression was 11.1 weeks (range 0.7-43.1 weeks) for Hycamtin treated patients and 11.9 weeks (range 0.1-57.3 weeks) for CAV treated patients. In a protocol modification dated 19 March 1997, the sponsor changed the final analysis of results from a follow-up period of 20 weeks to a follow-up of 12 weeks. The sponsor stated in the protocol that the interim analysis was not strictly an analysis for efficacy because no modifications to the experimental design were based on the interim analysis.

Dose: The Hycamtin arm received 1.5 mg/m^2/day as a 30-minute intravenous infusion for 5 consecutive days every 21 days. The CAV arm received cyclophosphamide 1000 mg/m^2, doxorubicin 45 mg/m^2, and vincristine 2 mg, administered once on day 1, every 21 days.

Criteria for Evaluation: The primary endpoints were response rate and response duration. Response rate was defined as the percentage of the total number of evaluable patients who responded to treatment (CR+PR). Response duration was defined as the time from initial documented response until the first sign of disease progression (this variable was undefined for non-responders). The secondary endpoints were time to response, time to progression, and survival. Patients assessed their disease symptoms at baseline and again prior to every course of therapy on the following symptom subscales: Shortness of Breath, Cough, Chest Pain, Coughing Up Blood, Loss of Appetite, Interference with Sleep, Hoarseness, Fatigue, and Interference with Daily Activities. Each subscale was assessed on the following 4-point ordinal scale: "Not at All," "A Little," "Quite a Bit," and "Very Much."

4.2. Efficacy Endpoints and Results

Response Rate

Out of 107 patients on 422 Hycamtin courses, 85% had a defined response of CR, PR, SD, or PD. The remaining 15% had a final response of NE (not evaluable) by the protocol definition. Out of 104 patients who received CAV in 343 courses, 81% had a defined response of CR, PR, SD, or PD and the remaining 19% had a final response of NE. Overall best response was listed as NE for patients whose response could not be determined due to premature withdrawal from the study or whose lesion assessments were insufficient to confirm a response.

In the intent-to-treat population, the objective response rate (CR+PR) for the Hycamtin arm was 24.3% (95% CI 16.2-32.4%). The objective response rate for the CAV arm was 18.3% (95% CI 10.8-25.7%). The difference in response rates was 6.0% (95% CI -5.0-17.0%). The p-value for the test of superiority of response rate was not statistically significant ($p = 0.283$). This reviewer calculated a 95% confidence interval for the difference in proportions using the continuity correction recommended on p. 29 of Fliess (1981) and this resulted in a slight increase in confidence interval width: 6.0% (95% continuity corrected CI -5.9-18.0%). The p-value using this continuity correction was $p = 0.366$.

Note that the lower confidence limits of both confidence intervals are not less than -14.0%. The conclusion we may draw from this endpoint, therefore, is that there is evidence that the difference between treatments excludes values larger than 14 percentage points in favor of CAV, which is what this trial was powered to detect.

Interpretation of these two primary endpoints (objective response and duration of response) is difficult since there was no adjustment for Type I error pre-specified in the protocol. If efficacy is to be concluded by a statistically significant finding on only one of the two primary endpoints, then there needs to be a corresponding adjustment for Type I error. However, it is not stated in the protocol whether the co-primary endpoints should be considered simultaneously or independently. A simple conservative method of adjusting for the two multiple endpoints is to construct the 97.5% confidence interval for the difference in objective response. This 97.5% continuity corrected confidence interval for the difference in objective response is -9.0% to 21.0%. Therefore, we still may conclude that there

is evidence that the difference between treatments excludes values larger than 14 percentage points in favor of CAV.

Duration of Response

The median duration of response was measured from the time of documented response to the time until progression. In the ITT population, the median duration of response for the Hycamtin arm was 14.4 weeks (95% CI 13.1-18.0 weeks). The median duration of response for the CAV arm was 15.3 weeks (95% CI 13.1-23.1 weeks). The difference between duration of response in the two arms was not statistically significant (logrank p = 0.592). The proportional hazards ratio was 1.29 (95% CI 0.58-2.85). The protocol did not specify a test for non-inferiority of duration of response (no difference in duration of response between CAV and Hycamtin was specified for testing purposes). However, the confidence interval of the hazards ratio was very wide and therefore does not warrant a non-inferiority claim.

Time to Response

Time to response was defined as the time from the first dose administration to the time of initial documented response. In the ITT population, the median time to response was 6.1 weeks (95% CI 5.6-6.9 weeks) in the CAV group and 6.2 weeks (95% CI 5.7-6.4 weeks) in the Hycamtin group. There was no statistically significant difference in time to response (logrank p = 0.680). The proportional hazards ratio for time to response was 0.96 (95% CI 0.48-1.91).

Time to Progression

Time to progression was defined as the time from the first dose administration until the first documented progression. The median time to progression in the Hycamtin arm was 13.3 weeks (95% CI 11.4-16.4 weeks). The median time to progression in the CAV arm was 12.3 weeks (95% CI 11.0-14.1 weeks). There was no statistically significant difference between arms with respect to time to progression (logrank p=0.696). The proportional hazards ratio for time to progression was 0.95 (95 % CI 0.72-1.25).

Survival

Survival was not defined formally in the protocol, although this reviewer based his assessment of survival on the following definition: survival is the time from the first dose administration until death due to any cause or loss to follow-up. Patients who had not died were censored at their last follow-up.

There were discrepancies among 13 patients with respect to last follow-up date. This reviewer found that additional follow-up information was available from these patients' case report forms that was not reflected in the electronic data sets submitted. A table of these discrepancies appears in Table 4.1.

Table 4.1. Follow-up / last visit discrepancies between the case report forms (CRF) and the submitted electronic data. The date format is mm/dd/yy and "Missing" designates a missing observation for follow-up.

Patient ID	Last Follow-up or Visit As Reported in Elect. Data	Last Follow-up or Visit As Reported in CRF
	05/07/97	06/30/97
	05/08/97	06/13/97
	Missing	02/19/97
	05/03/97	06/16/97
	Missing	02/24/96
	12/20/96	07/11/97
	Missing	05/06/97
	Missing	04/16/97
	11/20/96	05/16/97
	Missing	04/15/97
	05/13/97	06/06/97
	Missing	04/15/97
	11/06/96	05/12/97

Presumably these discrepancies affect the sponsor's results on survival. Note that the discrepancies ranged from a few days to half a year.

For the ITT population, median survival in the Hycamtin arm was 24.7 weeks (95% CI 19.7-28.9 weeks). The median survival in the CAV arm was 24.4 weeks (95% CI 21.3-29.3 weeks). The hazard ratio for survival was 1.14 (95% CI 0.86-1.53). The test for a difference in survival between treatments was not statistically significant (logrank p = 0.367).

If one defines survival as the time from randomization until death or censoring, median survival in the Hycamtin arm of the ITT population was 25.3 weeks (95% CI 20.3-29.4 weeks). The median survival in the CAV arm was 25.1 weeks (95% CI 20.3-29.4 weeks). The hazard ratio for survival was 1.15 (95% CI 0.86-1.54). The test for a difference in survival between treatments was not statistically significant (logrank p = 0.349).

Disease Symptoms

Symptoms of Disease questionnaires were administered at screening and immediately prior to each subsequent course of chemotherapy. Nine symptoms were assessed: shortness of breath, cough, chest pain, hemoptysis, anorexia, insomnia, hoarseness, fatigue and interference with daily activity. Each symptom was scored on a 4 point ordinal scale: "Not at All," "A Little," "Quite a Bit," and "Very Much." The baseline assessments appear in Table 4.2.

Table 4.2. Frequency of symptom occurrence at baseline for study 090.

	Hycamtin					CAV				
Symptom	$n_{baseline}$	Not at All	A Little	Quite a Bit	Very Much	$n_{baseline}$	Not at All	A Little	Quite a Bit	Very Much
Short of breath	78	13	30	26	9	69	17	29	17	6
Cough	78	18	30	24	6	69	15	32	15	7
Chest Pain	78	47	22	7	2	67	36	20	7	4
Hemoptysis	77	70	6	1	0	69	62	7	0	0
Anorexia	78	30	20	13	15	69	28	20	16	5
Insomnia	78	42	18	15	3	68	29	19	12	8
Hoarseness	78	48	20	8	2	69	44	13	8	4
Fatigue	77	10	25	29	13	69	15	19	20	15
Interf. with Daily Activity	77	19	21	23	14	69	21	14	21	13

$n_{baseline}$ is the number of patients with baseline assessments for each symptom.

The sponsor defined an improvement in disease-related symptoms as two consecutive visits at which a symptom score indicated improvement in one ordinal point over the baseline assessment. Patients included in the sponsor's analysis were required to have both baseline and post-baseline measurements in order to assess a change. In the event that a patient had a missing baseline measurement and a non-missing post-baseline measurement of "A Little" or worse, the sponsor imputed "Not at All" for the baseline value and the patient was included in the analyses. When symptom assessments were missing during any course of treatment, the "worst" of the two known non-missing symptom assessments recorded either prior to or after the missing assessment was imputes at the missing symptom assessment. The sponsor examined the changes from baseline in terms of improvement, no change, or worsening for both chemo-sensitive and refractory patients. The frequencies and proportions of those patients that improved appear in Table 4.3.

Table 4.3. Sponsor's analysis of percentage of patients with symptoms of disease improvement and Pearson's χ^2 p-values for comparing Hycamtin to CAV.

	Hycamtin			CAV			
Symptom	$n_{baseline+}$	$n_{improved}$	%	$n_{baseline+}$	$n_{improved}$	%	χ^2 p-val.
Shortness of Breath	66	19	28.8	61	4	6.6	0.001
Cough	66	17	25.8	61	8	13.1	0.073
Chest Pain	42	11	26.2	41	7	17.1	0.314
Hemoptysis	15	4	26.7	11	4	36.4	0.597
Anorexia	55	16	29.1	57	10	17.5	0.148
Insomnia	55	18	32.7	53	10	18.9	0.100
Hoarseness	38	12	31.6	37	4	10.8	0.028
Fatigue	68	16	23.5	65	6	9.2	0.027
Interf. with Daily Activity	64	18	28.1	63	7	11.1	0.016

$n_{baseline+}$ is the number of patients with baseline and at least one post-baseline assessment.
$n_{improved}$ is the number of patients with improvement in symptoms; "improvement" is defined as two consecutive improvements over the baseline assessment.

The sponsor presented formal comparisons for the hypothesis that patients on Hycamtin experience more palliation of disease symptoms than patients on the CAV arm. There are several statistical issues that made interpretation of these comparisons problematic for this reviewer:

- The sponsor's analysis consisted of comparing the frequency of those symptoms that had improved from baseline (as described above) using a Pearson's χ^2 statistic. The assumptions of this analysis rely heavily on the behavior of patient dropout. Namely, it assumes that patients are missing completely at random. A longitudinal quality of life analysis is more appropriate in determining trends over time and variance inflation due to non-random patient dropout. It should be noted that the sponsor's originally planned on performing such a longitudinal analysis by utilizing Generalized Estimating Equation (GEE) methodology. However, convergence problems were encountered due to the prevalence of missing observations, necessitating the Pearson's χ^2 alternative analyses. The assumption of dropout occurring completely at random is a rather strong one to justify, particularly in cancer trials.
- An analysis based on Pearson's χ^2 statistic does not account for within-patient correlation across the repeatedly measured symptom assessments.
- Imputing baseline values for patients without baseline measurements may be a confounding factor for the treatment comparisons for symptom improvement. However, treatment of missing baseline observations in this manner is the worst-case scenario.
- The interpretation of "consecutive visits" in the above analysis must be made with caution, as the maximum lag between "consecutive" visits was four.
- If one were to overlook the potential bias inherent in the p-values, there should be some adjustment to Type I error due to the multiplicity of symptom subscales.

It is this reviewer's opinion that p-values based on the Pearson's χ^2 statistic in this setting may be biased.

4.3. Summary and Conclusions

This study was designed as a non-inferiority trial for the objective response endpoint. In the intent-to-treat population of this Phase III trial, the objective response rate (CR+PR) for the Hycamtin arm was 24.3% (95% CI 16.2-32.4%). The objective response rate for the CAV arm was 18.3% (95% CI 10.8-25.7%). The difference in response rates was 6.0% (95% CI -5.9-18.0%). Since the lower confidence limit of this confidence interval is not less than -14.0%, we may conclude that there is evidence that the difference between treatments excludes values larger than 14 percentage points in favor of CAV, which is what this trial was powered to detect.

On the other primary efficacy endpoint, duration of response, the Hycamtin arm had a median of 14.4 weeks (95% CI 13.1-18.0 weeks) and the median duration of response for the CAV arm was 15.3 weeks (95% CI 13.1-23.1 weeks). The difference between duration of response in the two arms was not statistically significant (logrank p = 0.592). The proportional hazards ratio was 1.29 (95% CI 0.58-2.85). The protocol did not specify a test for non-inferiority of duration of response (no difference in duration of response between CAV and Hycamtin was specified for testing purposes). However, the confidence interval for duration of response was very wide and therefore does not warrant a non-inferiority claim.

There was no prospective Type I error apportionment for the two primary endpoints, so conclusions on the interpretation of these two endpoints combined should be considered as post hoc if efficacy is to be concluded by a statistically significant finding on only one of the two primary endpoints. However, it is not stated in the protocol whether the co-primary endpoints should be considered simultaneously or independently. A simple conservative method of adjusting for the two multiple endpoints is to construct the 97.5% confidence interval for the difference in objective response. This 97.5% continuity corrected confidence interval for the difference in objective response is -9.0% to 21.0%.

Therefore, we still may conclude that there is evidence that the difference between treatments excludes values larger than 14 percentage points in favor of CAV.

The other secondary endpoints included time to first response, time to progression and survival. The median time to response was 6.1 weeks (95% CI 5.6-6.9 weeks) in the CAV group and 6.2 weeks (95% CI 5.7-6.4 weeks) in the Hycamtin group. There was no statistically significant difference in time to response (logrank p = 0.680) and the proportional hazards ratio for time to response was 0.96 (95% CI 0.48-1.91). The median time to progression in the Hycamtin arm was 13.3 weeks (95% CI 11.4-16.4 weeks). The median time to progression in the CAV arm was 12.3 weeks (95% CI 11.0-14.1 weeks). There was no statistically significant difference between arms with respect to time to progression (logrank p=0.696) and the proportional hazards ratio for time to progression was 0.95 (95 % CI 0.72-1.25). Finally, the median survival in the CAV arm was 24.4 weeks (95% CI 21.3-29.3 weeks). The test for a difference in survival between treatments was not statistically significant (logrank p = 0.367) and the proportional hazards ratio for survival was 1.14 (95% CI 0.86-1.53).

Although the assessment of symptom improvement provides important information on symptom palliation of patients on the Hycamtin and CAV arms, it is this reviewer's opinion that the formal comparisons presented may suffer from bias for the reasons outlined at the end of Section 4.2.

5. Integrated Summary of Hycamtin Efficacy

Since study 090 and the three pivotal phase II studies 014, 053 and 092 were similar with respect to Hycamtin dosage and study design, the medical reviewer requested the combined efficacy results across all studies. In the analyses that follow, Hycamtin efficacy is tabulated by chemo-sensitivity. The CAV arm of study 090 was not considered. This integrated summary of Hycamtin efficacy appears in Table 5.1.

Table 5.1. Integrated summary of efficacy for patients who received Hycamtin across the four pivotal trials. Time to event endpoints are given in weeks. The numbers in parentheses are 95% confidence intervals for the point estimate.

	Prior Chemotherapy (Sensitive / Refractory)		
Endpoint	Refractory (150 pts.)	Sensitive (275 pts.)	All Pts. Combined (426 pts.)*
Objective Response	4.0% (0.9-7.1%)	20.4% (15.6-25.1%)	14.3% (11.2-17.9%)
Time to Response	5.7 (5.4-5.7)	6.4 (5.7-6.6)	6.1 (5.7-6.6)
Duration of Response	24.9 (20.6-47.1)	17.1 (14.1-22.9)	18.4 (14.6-22.9)
Time to Progression	7.7 (6.3-9.6)	12.6 (11.3-14.3)	11.1 (9.6-11.9)
Survival	14.0 (11.3-17.4)	24.9 (21.7-27.6)	20.7 (18.1-23.6)

* Sensitivity data was missing for one patient (014.018.02324). This patient's best response was progression.

The results in Table 5.1 are consistent with the integrated summary of efficacy included in the sponsor's submission. It should be noted that in the refractory group, the estimates of objective response, time to response, and duration of response were based on six patients.

6. Overall Recommendations and Conclusions

In all of the pivotal trials included in this submission, objective response rate was either the primary or co-primary endpoint. Based on the results of these trials, we can estimate with some certainty that the response rate for Hycamtin is approximately 14.3% for patients who are either refractory or sensitive to chemotherapy. The response rate is estimated to be approximately 4.0% for patients who are refractory to chemotherapy and approximately 20.4% for patients who are sensitive to chemotherapy. In the comparative Phase III trial, the difference in response rate favored Hycamtin over CAV; this

estimated difference was 6.0% (95% CI -6.6-19.2%). Since the lower confidence limit of this continuity corrected confidence interval is not less than -14.0%, we may conclude that there is evidence that the difference between treatments excludes values larger than 14 percentage points in favor of CAV, which is what the comparative Phase III trial was powered to detect. A simple conservative method of adjusting for the two multiple endpoints is to construct the 97.5% confidence interval for the difference in objective response. This 97.5% continuity corrected confidence interval for the difference in objective response is -10.5% to 22.5%. Therefore, we still may conclude that there is evidence that the difference between treatments excludes values larger than 14 percentage points in favor of CAV.

In Study 090, the other co-primary endpoint of duration of response was not statistically significant based on a logrank test ($p = 0.592$). The proportional hazards ratio was 1.29 (95% CI 0.58-2.85). The protocol did not specify a test for non-inferiority of duration of response (no difference in duration of response between CAV and Hycamtin was specified for testing purposes). However, the confidence interval of the hazards ratio was very wide and therefore does not warrant a non-inferiority claim.

Other than objective response, it is not clear which endpoints in Study 090 should be tested with a hypothesis of superiority or a hypothesis of non-inferiority. If one tests for superiority and concludes the null hypothesis, and then tests for non-inferiority, additional Type I error adjustments must be made. The sponsor only tested for superiority for endpoints other than objective response, and so conclusions about non-inferiority of Hycamtin may not be justified.

There was no statistically significant difference between CAV and Hycamtin on time to progression, time to first response and survival. Assessment of disease symptoms were tabulated in study 053 and 090. The information provided by these assessments are useful, but it is this reviewer's opinion that the formal comparisons and hypothesis tests that were presented suffer from large amounts of missing data and depend on strong assumptions that are otherwise difficult to justify.

It is this reviewer's opinion that the evidence in favor of Hycamtin efficacy rests solely on whether ruling out a difference in objective response greater than 14% of CAV over Hycamtin is of interest to clinicians. There is evidence to support this conclusion, based on the lower bound of the 95% confidence interval for the difference in objective response between CAV and Hycamtin. There is little evidence to conclude a Hycamtin efficacy advantage over CAV, since all other endpoints are either not statistically significant (as is the case of objective response and the time to event endpoints) or are based on strong underlying assumptions that were not justified (as in the case of symptom improvement).

/S/

David Smith, Ph.D.
Mathematical Statistician

Concur:

Dr. Chen

Dr. Chi 6/4/98

cc: Archival NDA 20-671

HFD-150 / Ms. D. Catterson, Project Manager

HFD-150 / Dr. S. Hirschfeld

HFD-150 / Dr. G. Williams

HFD-344 / Dr. B. Barton

HFD-710 / Dr. G. Chi

HFD-710 / Dr. G. Chen

HFD-710 / Dr. D. Smith

HFD-710 / Chron

Smith / 18 May 1998 / Word / c:\data\wordfiles\20671.doc

This review consists of 15 pages of text.

References

Gehan E. The Determination Of The Number Of Patients Required In A Preliminary And A Follow-up Trial Of A New Chemotherapeutic Agent. *J. Chronic Diseases* **13** 346-353, 1961.

Makuch R and Simon R. Sample size requirements for Evaluating a Conservative Therapy. *Cancer Treatment Reports* **62**(7) 1037-40, 1978.

Fliess, J.L. *Statistical Methods for Rates and Proportions*, 2nd Edition. John Wiley and Sons, Inc. New York. 1981.