Food and Drug Administration

Center for Drug Evaluation and Research

Oncologic Drugs Advisory Committee
May 3-4, 2004

Briefing Material: May 3, 2004
AM Session - Genasense

Table of Contents

1 EXECUTIVE SUMMARY: 1

2 AGENCY APPROVAL REQUIREMENTS 2

2.1 EFFECTIVENESS REQUIREMENT FOR APPROVAL 2
2.2 APPROVAL MECHANISMS 3

3 MELANOMA 3

3.1 APPROVED THERAPIES FOR METASTATIC DISEASE 3
3.1.1 Hydroxyurea 4
3.1.2 Dacarbazine (DTIC) 4
3.1.3 Aldesleukin (IL-2) 4
3.2 NON-APPROVED THERAPIES FOR METASTATIC DISEASE 4
3.2.1 Interferon 4
3.2.2 Temozolomide 4
3.2.3 Histamine Dihydrochloride 7

4 NDA SUBMISSION 8

4.1 ELIGIBILITY CRITERIA 9
4.2 EXCLUSION CRITERIA 10
4.3 STRATIFICATION 10
4.4 SCREENING ASSESSMENT 11
4.5 CONCOMITANT TREATMENT 11
4.5.1 Treatment 12
4.5.2 Efficacy Parameters 12
4.5.2.1 Primary Efficacy Parameter: 12
4.5.2.2 Secondary Efficacy parameter definitions 12
4.5.2.3 Missing and Incomplete Data 13
4.5.2.4 Confirmed response 14
4.5.2.5 Independent, Blinded Assessment of Response 14

5 RESULTS 14

5.1 ANALYSIS POPULATIONS 14
5.2 PATIENT DEMOGRAPHICS 15
5.3 BASELINE DISEASE CHARACTERISTICS 17
5.4 EFFICACY 21
5.4.1 Overall Survival 21
5.4.2 Progression-Free Survival (1): Sponsor's Results and Reviewer's Comments 25
5.4.3 Progression-Free Survival (2): Reviewer's Evaluations and Exploratory Analyses 29
5.4.3.1 Evaluation of The Impact of Assessment Schedules 29
5.4.3.2 Evaluation of The Impact of Missing Data 34
5.4.3.3 Exploratory Analyses of Interval-Censored Data 37
5.4.4 Antitumor Response Rate: Sponsor's Analysis and Reviewer's Comments 39
5.4.5 Duration of Antitumor Response: Sponsor's Analysis and Reviewer's Comments 41
5.4.6 Findings in Special/Subgroup Populations 41
5.4.6.1 Gender, Race and Age 41
5.4.6.2 Other Special/Subgroup Populations 44
5.4.7 Exploratory FDA Analysis of Performance Status 46
5.5 SAFETY 46
5.5.1 Dosing summary for GM301 46
5.5.2 AEs AND SAEs 47

6 SUMMARY AND CONCLUSIONS 51

6.1 STATISTICAL ISSUES AND COLLECTIVE EVIDENCE 51

6.2 CONCLUSIONS AND RECOMMENDATIONS 52
APPENDIX 53
6.3 SIMULATION RESULTS OF DIFFERENT LESION ASSESSMENT SCHEDULES 53

List of Tables

TABLE 1: SUMMARY OF FDA ANALYSIS: TMZ VERSUS DTIC 6
TABLE 2 DEMOGRAPHIC AND BACKGROUND CHARACTERISTICS (ITT) 15
TABLE 3 SYMPTOMS AT BASELINE 16
TABLE 4 BASELINE DISEASE CHARACTERISTICS 19
TABLE 5 PRIOR THERAPY 19
TABLE 6 PATIENT DISPOSITION 20
TABLE 7 TREATMENT DISCONTINUATION THROUGH FIRST TWO CYCLES 21
TABLE 8 DISPOSITION OF SUBJECTS BY TREATMENT GROUP: INTENT-TO-TREAT POPULATION 21
TABLE 9: SURVIVAL TIME ANALYSIS (INTENT-TO-TREAT POPULATION) 22
TABLE 10: SURVIVAL TIME ANALYSIS: INTENT-TO-TREAT COHORT RANDOMIZED ON OR BEFORE 01 AUG 2002 (N = 520) 23
TABLE 11: PROGRESSION-FREE SURVIVAL ANALYSIS (INTENT-TO-TREAT POPULATION) 25
TABLE 12: SUMMARY OF MISSING DATA IN ANALYSIS OF PROGRESSION-FREE SURVIVAL 27
TABLE 13: ANALYSIS OF PROGRESSION-FREE SURVIVAL WITH DATA CENSORED AT DATE OF LAST COMPLETE OBSERVATION 27
TABLE 14: SUMMARY OF NUMBER OF PATIENTS BY NUMBER OF LESION ASSESSMENTS 29
TABLE 15: SUMMARY OF TIME FROM RANDOMIZATION TO THE 1ST, 2ND AND 3RD LESION ASSESSMENTS(1) (REVIEWER'S TABLE) 30
TABLE 16: SUMMARY OF TIME FROM TREATMENT START TO THE 1ST, 2ND AND 3RD LESION ASSESSMENTS(1) (REVIEWER'S TABLE) 32
TABLE 17: SUMMARY OF NUMBER OF PATIENTS WITH MISSING LESION MEASUREMENTS PRIOR TO DOCUMENTED DISEASE PROGRESSION OR DEATH AMONG THOSE WHO HAD DISEASE PROGRESSION OR DEATH (REVIEWER'S TABLE) 36
TABLE 18: SUMMARY OF NUMBER OF PATIENTS WITH MISSING LESION MEASUREMENTS AT THE SAME TIME OF DOCUMENTED DISEASE PROGRESSION OR DEATH AMONG THOSE WHO HAD DISEASE PROGRESSION OR DEATH (REVIEWER'S TABLE) 36
TABLE 19: SUMMARY OF NUMBER OF PATIENTS WITH MISSING LESION MEASUREMENTS AMONG THOSE WHO WERE CENSORED AND HAD BASELINE AND POST BASELINE LESION MEASUREMENTS (REVIEWER'S TABLE) 37
TABLE 20: ILLUSTRATION OF APPROACHES TO INTERVAL-CENSORING (REVIEWER'S TABLE) 38
TABLE 21: RESULTS OF GENERALIZED LOGRANK TEST WITH INTERVAL-CENSORED DATA (REVIEWER'S TABLE) 39
TABLE 22: SPONSOR-CONFIRMED ANTITUMOR RESPONSE RATES DURING THE TREATMENT PHASE 40
TABLE 23: RADPHARM-CONFIRMED ANTITUMOR RESPONSE RATES DURING THE TREATMENT PHASE (REVIEWER'S TABLE) 40
TABLE 24: SUMMARY OF SURVIVAL ANALYSIS BY GENDER 42
TABLE 25: SUMMARY OF SURVIVAL ANALYSIS BY AGE 42
TABLE 26: SUMMARY OF PROGRESSION-FREE SURVIVAL ANALYSIS BY GENDER 43
TABLE 27: SUMMARY OF PROGRESSION-FREE SURVIVAL ANALYSIS BY AGE 43
TABLE 28: SUMMARY OF SPONSOR-CONFIRMED ANTITUMOR RESPONSE RATE BY GENDER 44
TABLE 29: SUMMARY OF SPONSOR-CONFIRMED ANTITUMOR RESPONSE RATE BY AGE 44
TABLE 30: SUMMARY OF SURVIVAL ANALYSIS BY COUNTRY 45
TABLE 31: SUMMARY OF PROGRESSION-FREE SURVIVAL ANALYSIS BY COUNTRY 45
TABLE 32: SUMMARY OF SPONSOR-CONFIRMED ANTITUMOR RESPONSE RATE BY COUNTRY 46
TABLE 33 DOSING DURING THE STUDY 46
TABLE 34 OVERVIEW OF ADVERSE EVENTS BY TREATMENT GROUP: SAFETY POPULATION 47
TABLE 35 ADVERSE EVENTS IN ≥ 20% OF PATIENTS 48
TABLE 36 REVIEWER'S SUMMARY COMPILATION OF TREATMENT EMERGENT AES (TEAES) OF CLINICAL INTEREST FROM THE SPONSOR'S TABULATION TABLE 14.3.1.5: SAFETY POPULATION 49
TABLE 37 SERIOUS TREATMENT-EMERGENT ADVERSE EVENTS (SAES) OCCURRING IN = 2% OF SUBJECTS BY PREFERRED TERM AND TREATMENT GROUP: SAFETY POPULATION 50
TABLE 38: ASSESSMENT SCHEDULES IN SIMULATION STUDIES (REVIEWER'S TABLE) 53

TABLE 39: SIMULATION RESULTS UNDER EQUAL PROGRESSION-FREE SURVIVAL DISTRIBUTIONS (REVIEWER'S TABLE) 54
TABLE 40: SIMULATION RESULTS UNDER SLIGHTLY DIFFERENT PROGRESSION-FREE SURVIVAL DISTRIBUTIONS (REVIEWER'S TABLE) 55

List of Figures

FIGURE 1 OVERALL SURVIVAL - ITT POPULATION (P=0.20, NS) FROM THE ODAC FDA PRESENTATION ON MARCH 23, 1999 5
FIGURE 2 PROGRESSION-FREE SURVIVAL - ITT POPULATION FROM THE ODAC FDA PRESENTATION ON MARCH 23, 1999 6
FIGURE 3: KAPLAN-MEIER SURVIVAL CURVES BY TREATMENT GROUP (INTENT-TO-TREAT 22
FIGURE 4: KAPLAN-MEIER SURVIVAL CURVES BY TREATMENT GROUP: INTENT-TO-TREAT COHORT RANDOMIZED ON OR BEFORE 01 AUG 2002 (N = 520) 24
FIGURE 5: KAPLAN-MEIER PROGRESSION-FREE SURVIVAL CURVES BY TREATMENT GROUP (INTENT TO-TREAT POPULATION) – (SPONSOR'S FIGURE) 26
FIGURE 6: KAPLAN-MEIER CURVES OF TIME FROM RANDOMIZATION TO 1ST, 2ND, AND 3RD LESION ASSESSMENTS (REVIEWER'S FIGURES) 31
FIGURE 7: KAPLAN-MEIER CURVES OF TIME FROM TREATMENT START TO 1ST, 2ND, AND 3RD LESION ASSESSMENT (REVIEWER'S FIGURES) 33

1 EXECUTIVE SUMMARY:

Genta, Inc., seeks the following indication: Genasense in combination with dacarbazine is indicated for the treatment of patients with advanced melanoma who have not received prior chemotherapy.

Genta has submitted a single, international, multi-center, open-label, active control, randomized, phase 3 study of Genasense (G3139) plus dacarbazine (DTIC) versus DTIC alone every three weeks as first-line chemotherapy for metastatic melanoma, along with a small, supportive, single arm, phase 1/2 study with various doses and schedules of G3139 plus DTIC. In the phase 3 study, G3139 was given as a continuous intravenous (IV) infusion over 5 days, requiring central venous access and an ambulatory pump. DTIC 1000 mg/m2 IV was given over 60 minutes once every three week cycle in each arm.

The primary endpoint of the trial was survival. The trial was sized to demonstrate improvement in survival from 6 months with DTIC alone to 8 months for the combination. Secondary endpoints were progression-free survival and response rate. A total of 771 patients were enrolled from 9 countries, 56% from the U.S., 11% from Australia, and the remainder from Europe and Canada. Most patients were asymptomatic and were ECOG performance status 0 (56%) at baseline. Despite attempts to stratify and balance prognostic factors at randomization, some imbalances did occur (e.g., fewer patients with visceral disease-LDH elevations (59% versus 67% in the DTIC alone arm)). The majority of patients in both arms went off study after 6 weeks (two cycles) because of progressive disease.

The study failed to show a survival benefit from the combination of G3139 plus DTIC using an unadjusted logrank analysis of survival time for the intention-to-treat population ($p = 0.18$, HR=0.89). The sponsor's analysis of secondary endpoints did show a statistically significant benefit in progression-free survival from a median of 49 days on DTIC to 74 days on the combination, a difference of 25 days ($p=0.0003$, HR=0.73), using a censoring procedure of "last observation carried forward" for missing data. Also, the sponsor reported a significant difference in response rate of 6.8% for DTIC alone versus 11.7% for the combination ($p=0.019$).

The FDA chose to examine the sponsor’s secondary endpoints in more detail. The sponsor was asked to perform a different procedure of censoring at last observation for missing data which resulted in a progression free survival (PFS) of 48 days on DTIC alone versus 61 days for the combination, a difference of 13 days ($p=0.0006$, HR=0.75), also statistically significantly different. Furthermore, simulations conducted by FDA reviewers suggest that in a large study such as the one under review, with a very small systematic study arm bias such as in assessment intervals between the study arms, statistically significant differences may be observed which are in fact false positive.

The sponsor-determined response rates (see above) were based on a computer algorithm using measurements from the investigators. The sponsor had an independent contract

organization review data (radiological films and photographs) for all responding patients. The overall response rates determined by the blinded, independent review organization (IRO) were 3.6 % for DTIC alone and 6.7 % for the combination, a difference of 3.1% (p= 0.056). Of the 5 complete responses reported by Genta, none were verified by the IRO. For all 71 responders identified by investigators and Genta, there was concordance with the IRO for only 49% of the interpretations.

Most patients were asymptomatic at study entry and were performance status zero. Therefore, it was difficult to assess whether patients achieved any symptom benefit from combination therapy over single agent therapy.

The combination arm was associated with increased toxicity and discontinuations due to adverse events (AEs) including 69 (18.6%) patients who discontinued therapy for adverse events on the G3139 arm versus 39 (10.8%) on the DTIC alone arm. The rate of serious adverse events, SAEs, was 40% on the G3139 arm versus 27% on DTIC alone. Since the dosing of DTIC was identical on the two arms; toxicity increases were likely due to the addition of G3139.

In summary, the study failed to demonstrate a survival benefit, the primary trial endpoint. Uncertainty exists regarding the reliability of small benefits claimed for PFS and response rates. Uncertainty also exists regarding whether an improvement in PFS and RR of this magnitude outweighs the increase in toxicity seen with the combination.

In considering the potential approval of Genasense, recent regulatory history is pertinent. To date, the Division has not considered small differences in PFS to represent clinical benefit for metastatic melanoma patients, especially in an asymptomatic patient population. This position was most recently reviewed with ODAC when the Temozolomide application was presented in 1999. The Temozolomide application, like the present Genasense application, did not demonstrate an improvement in survival and had similar effects for PFS. Temozolomide was not approved.

Approval for a melanoma application requires substantial evidence of efficacy. Review of the literature has not shown a correlation of improvement in PFS, TTP, or RR with an improvement in survival. GM-301 is the largest well-conducted phase 3 trial in metastatic melanoma to date. Survival was not improved and toxicity was increased.

2 AGENCY APPROVAL REQUIREMENTS

2.1 EFFECTIVENESS REQUIREMENT FOR APPROVAL

In 1962, Congress amended the Federal Food, Drug, and Cosmetic Act to add a requirement that, to obtain marketing approval, manufacturers must demonstrate effectiveness by providing "substantial evidence." Substantial evidence was defined in

section 505(d) of the Act as "evidence consisting of adequate and well-controlled investigations." With regard to quantity, it has been FDA's position that Congress generally intended to require at least two adequate and well-controlled studies, each convincing on its own, to establish effectiveness. In 1997, the Food and Drug Administration Modernization Act stated that a single trial may suffice if other supportive evidence exists such as evidence from other trials where the drug has been used in different age groups, at different doses, and in different regimens, or different modified release dosage forms. The 1998 Guidance for Industry: Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products states, that to be considered, the single trial must be well-conducted, internally consistent, and demonstrates a compelling result. In general, the FDA has relied on a single adequate and well controlled efficacy study (along with supportive evidence) to support approval in cases in which a single multicenter study of excellent design provided highly reliable and statistically strong evidence of an important clinical benefit, such as an effect on survival, and a confirmatory study would have been difficult to conduct on ethical grounds.

2.2 APPROVAL MECHANISMS

With regard to approval, two mechanisms for approval exist: regular and accelerated. The Agency grants regular approval based on an endpoint which demonstrates clinical benefit or an established surrogate for clinical benefit. The Agency grants accelerated approval based on a surrogate endpoint deemed reasonably likely to predict clinical benefit. For Accelerated Approval, the Agency requires additional post-marketing studies to verify the clinical benefit. The March 2003 Oncology Drug Advisory Committee (ODAC) recommended that the post-marketing studies be ongoing at the time of Accelerated Approval.

3 MELANOMA

Malignant melanoma, if detected in an early stage, is usually curable surgically. However, locally advanced or metastatic disease usually is not resectable or curable and is almost uniformly fatal. Radiation therapy has a very limited role in the treatment of melanoma. For the treatment of locally advanced or metastatic disease, physicians have used chemotherapy, immunotherapy or biochemotherapy. Single agent chemotherapy has been associated with low response rates. Combination chemotherapy has been associated with higher response rates than single agent chemotherapy but without a significant improvement in survival. In numerous phase 3 trials, no single agent or combination therapy has demonstrated a survival benefit for the treatment of melanoma over single agent DTIC. DTIC has not been shown to provide a survival benefit. Immunotherapy has been shown to induce only a few complete durable responses with considerable toxicity.

3.1 APPROVED THERAPIES FOR METASTATIC DISEASE

At the present time, the Agency has approved only single agent therapies for malignant melanoma. No combination therapies have been approved.

3.1.1 Hydroxyurea

In 1967, the Agency granted approval for hydroxyurea for use in malignant melanoma based on the Agency's calculation of an estimated 10% response rate from the published literature and submitted case reports. The response rate included patients who had a 25% or greater reduction in tumor size, which lasted 30 days or more.

3.1.2 Dacarbazine (DTIC)

In 1975, the Agency granted approval for dacarbazine (DTIC – Dome Labs, Inc., NDA 17-575) for the single agent treatment of metastatic melanoma based on data from cooperative group, single arm studies demonstrating an overall response rate (ORR) of 23%, including 6% complete responses (CRs). To date, no evidence for a survival or progression-free survival (PFS) benefit for DTIC has been demonstrated. Various dosing schedules have used IV injections of DTIC of 150 mg/m2 to 250 mg/m2 daily times 5 – 10 days, as well as 800 – 1200 mg/m2 once, each three to four weeks. Reported response rates have ranged from about 10% to 24% with response durations of 3 – 6 months. No dosing regimen has emerged as superior to all others for this alkylating agent.

3.1.3 Aldesleukin (IL-2)

In 1998, the Agency approved Aldesleukin (Proleukin, IL-2, Chiron Corp.), a recombinant form of interleukin-2, a T-lymphocyte growth factor cytokine. The approval basis of IL-2 (BLA97-0501) for adults with metastatic melanoma rested on the demonstration of long term, durable CRs from 8 phase 2 studies. Patient eligibility for some of these trials included a negative thallium stress test. The median age of these highly selected patients was 42 years. The studies were pooled and showed an objective response rate of 16%, with 6% (17) CRs and 10% (26) PRs among 270 patients in a series of NCI trials.[i] Most of the CRs involved responses in skin, lymph node and lung sites. The median duration of response (for all responders) was 8.9 months, and the median duration of response for PRs was 6 months. When reported in 1999,[i] the median duration of CR had not been reached with 10 of 17 CRs continuing from 2 to 5 years. Use of IL-2 is limited due to significant toxicity.

3.2 NON-APPROVED THERAPIES FOR METASTATIC DISEASE

3.2.1 Interferon

In 1997 the Agency approved interferon alfa-2b (Intron A, Schering Corp) for adjuvant use following complete surgical resection in adults who are at high risk of recurrence, based on evidence of improved relapse-free and overall survival. Interferon has not been approved for treatment of metastatic disease, but is often used in this setting and response rates of 10-20% have been reported.

3.2.2 Temozolomide

On August 12, 1998, Schering submitted a New Drug Application for temozolomide (TMZ) for the treatment of metastatic melanoma (NDA 21-051). On March 23, 1999 the

Oncology Drug Advisory Committee reviewed the application. On June 11, 1999, the Agency denied approval for temozolomide (TMZ) (NDA 21051), a structural analog of DTIC, for the treatment of metastatic melanoma based on a review of the main trial and the Oncology Drug Advisory Committee's recommendation.

Schering submitted one major phase 3 trial for the indication. The international, multicenter, active control, open-label, phase 3 trial randomized 305 patients with metastatic melanoma to either DTIC 250 mg/m2 IV daily times 5 days each three weeks or TMZ 200 mg/m2 by mouth daily times 5 days each four weeks. The primary endpoint of this trial was survival. The study was sized to give an 80% power with a two-sided alpha of 5% to detect 3 month superiority in survival, from 6 months for DTIC to 9 months for TMZ therapy. The median survival on the DTIC arm was 6.4 months versus 7.7 months for TMZ, a difference in median survival of 1.3 months. Although 80% of the events had occurred by the time of the analysis, the difference was not statistically significant (p=0.2). An additional post-hoc analysis suggested an improvement in survival at 6 months (61% for temozolomide versus 51% for DTIC); however, the Agency did not accept this new endpoint for full or accelerated approval. In addition to the failure to demonstrate a survival benefit for Temozolomide, the ODAC and the Agency had reservations about the fact that DTIC itself lacked evidence for a survival benefit.

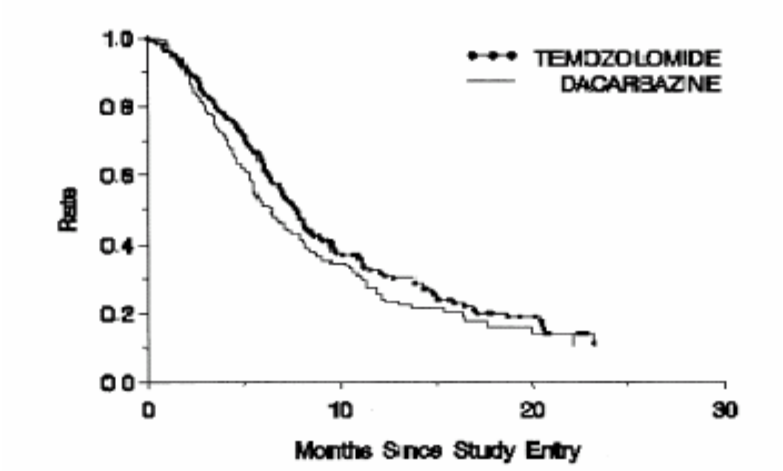


Figure 1 Overall Survival - ITT Population (p=0.20, NS) from the ODAC FDA presentation on March 23, 1999

Additional secondary endpoints reviewed were response rate and PFS. The Agency's analysis showed that median PFS was 1.74 months for Temozolomide versus 1.38 months for DTIC (p=0.002). Additional post-hoc exploratory analyses included possible prolongation in PFS for some patients in the third quartile of the survival curve before the curves came together again at 14 months. No symptomatic benefit was identified in association with this difference. Survival after progression also was similar on both arms.

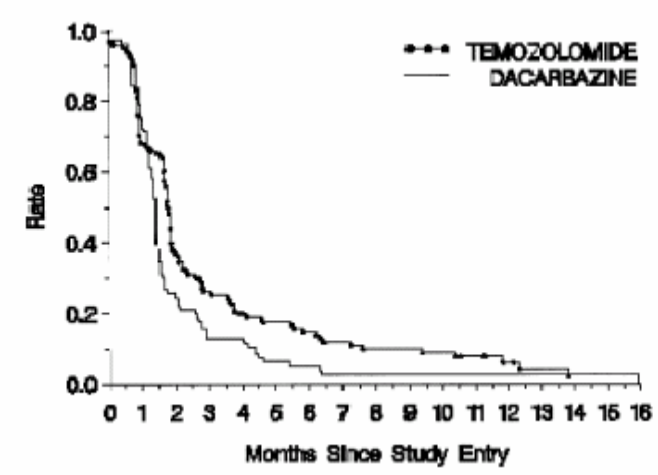


Figure 2 Progression-free Survival - ITT Population from the ODAC FDA presentation on March 23, 1999

The Agency compared response rates for the 2 arms; the ORR for TMZ was 12.2% (2.6% CR), versus 9.4% for DTIC (2.7% CR). The lower 95% confidence limit for the TMZ response rate was 7.1%. Six (75%) of the eight complete responses were in skin or nodes; 73% of the PRs were in skin, nodes, or lung. The safety profiles of the two drugs were similar.

Table 1: Summary of FDA analysis: TMZ versus DTIC

ITT population	TMZ	DTIC	Hazard Ratio	95% C.I. for HR	p value
	N = 156	N = 149			
Median Survival (months)	7.7	6.4	1.18	0.92, 1.52	0.20
Median PFS (months)	1.74	1.38	1.49	1.15, 1.92	0.002
Median PFS – Days	**53**	**42**	1.49	1.15, 1.92	0.002
Number of Responders CR/PR	4 / 15	4 / 10			
Overall Response Rate - %	12.2 %	9.4 %			0.43
95% C.I. for Response Rate	(7.1 – 17.3)	(4.7 -14.1)			

Reviewer's Table

During the Advisory Committee meeting, members considered two questions with regard to the Temozolomide application. The questions were:

1) Do the results of this study, particularly the objective tumor rates and response durations for the temozolomide versus dacarbazine, and the effect on progression-free survival, even in the absence of any effect on survival provide substantial evidence of effectiveness?
2) Does the Committee recommend approval of temozolomide for the treatment of advanced metastatic melanoma?

The voting on the first question was 10 No and one abstention. On the second question the vote was 10 No and 1 Yes.

In conclusion, TMZ was not approved because the trial did not show superiority in survival to DTIC alone. The statistically significant PFS (p-value = 0.002) and slightly higher response rate, in the absence of any clinical evidence of symptomatic benefit, were not thought to represent a meaningful benefit.

3.2.3 Histamine Dihydrochloride

On July 18, 2000, Maxim Pharmaceuticals submitted an NDA for Histamine dihydrochloride (NDA 21-240) based on results from a subgroup analysis from their major histamine trial in metastatic melanoma. On December 13, 2000, an ODAC met and reviewed the efficacy and safety data. On January 18, 2001, the Agency denied approval of Histamine for the adjunctive use with IL-2 in the treatment of adult patients with advanced metastatic melanoma that has metastasized to the liver.

The submission included one major randomized, controlled study which compared the results of low dose subcutaneous IL-2 with the same dose regimen of subcutaneous IL-2 plus Histamine. For the ITT population there was no statistically significant difference in survival; the median survivals were 245 (8 months) days and 272 (8.9 months) days, respectively. For the ITT population, the Agency noted that the response rate observed on each arm was 3%. For the ITT population, the Agency noted that the median time-to-progression (TTP) using censoring for missing data at the time of last observation observed on each arm was 2.7 months. Neither the Agency nor ODAC accepted the subgroup and post-hoc analyses.

In conclusion, no evidence to date exists that available treatment improves survival for patients with metastatic melanoma. With low treatment response rates and short durations of anti-tumor effect, numerous reports conclude that tumor response, similar to those observed in submitted NDAs, do not contribute to a survival benefit in metastatic melanoma. To date, survival statistics from studies of locally advanced or metastatic melanoma may reflect the patient mix of prognostic factors found in the trial, not the effects of therapy and comparisons of treatments must examine the mix of prognostic variables. The overall survival for patients with metastatic melanoma (stage IV) has ranged from 4.7 to 11 months, with a median survival of 8.5 months.[ii] Review of the literature has not shown a correlation of improvement in PFS, TTP, or RR with an improvement in survival.

In a recent analysis of prognostic factors by Balch et al,[iii] few patients with stage IV disease lived beyond one year. A SWOG database review in 1996 of 600 patients noted a 2% five year survival for metastatic melanoma patients who received various treatments.[iv] Lower performance status, increasing number of sites of metastases, elevated LDH, and liver metastases may worsen the prognosis. Female gender is associated with a more favorable prognosis as is a longer disease free interval from original diagnosis to time of recurrence.[v] Certain sites of metastatic disease altered survival slightly; patients with visceral metastases (excluding lung) had a median survival of 8 months, while for those with metastases confined to skin and distant lymph nodes the median survival was 14 months. Patients with lung metastases alone had a median survival of 13 months.[iii]

Treatment has not altered the natural history of the disease thus far for most patients. Treatment with DTIC alone has not been shown to have a survival benefit. No combination therapy has shown any advantage over DTIC alone.

4 NDA SUBMISSION

Two studies were submitted to support the efficacy claims for G3139 for the indication of treatment of advanced and metastatic melanoma in combination with DTIC. The comparative trial, GM301, was an international, multicenter, randomized, open label, active control, phase 3 trial comparing DTIC 1000 mg/m2 IV every three weeks with the same DTIC dose preceded by a 5 day continuous IV infusion of G3139 at 7 mg/kg/day every three weeks in 771 patients. The second supporting study was a single center, single arm, phase 1/2, open label, dose escalation involving 33 patients who received variable doses of G3139 and DTIC.

The primary objective of GM301 was to compare survival in patients with advanced malignant melanoma treated with DTIC alone versus patients treated with DTIC plus G3139. Secondary objectives included comparing Progression-free Survival and antitumor response rate, performance status, patient weight, and tumor-related symptoms between the 2 treatment arms.

The trial enrolled 771 patients who were randomized to either:

- Control group: DTIC (1000 mg/m^2) administered by IV infusion over 60 minutes on Day 1;
- Investigational group: GM3139 (7.0 mg/kg/day) administered by continuous IV infusion for 5 days (Days 1 through 6) and DTIC (1000 mg/m^2) administered by IV infusion over 60 minutes immediately upon completion of the GM3139 infusion.

4.1 ELIGIBILITY CRITERIA

The following criteria were to be met at baseline before randomization, within 4 weeks prior to starting therapy:

1. Histologically confirmed diagnosis of malignant melanoma
2. Progressive disease that was not surgically resectable or metastatic stage IV disease
3. Prior immunotherapy, cytokine, biologic, or vaccine therapy was permitted, so long as no cytotoxic chemotherapy had been administered
4. Measurable disease was required to be present either on physical examination or imaging studies. Lesions that were considered intrinsically non-measurable included the following:
   - Bone lesions;
   - Pleural/pericardial effusion;
   - Lymphangitis cutis/pulmonis;
   - Abdominal masses that are not confirmed and followed by imaging techniques; and
   - Lesions that are situated in a previously irradiated area

   Measurable disease is defined as at least one malignant lesion that can be accurately and serially measured in at least one dimension (longest diameter to be recorded), using a caliper (diameter = 10 mm) for superficial cutaneous disease, or using contrast-enhanced CT or spiral CT (diameter = 20 mm) for visceral or nodal/soft tissue disease.
5. Eastern Cooperative Oncology Group (ECOG) performance status 0-2
6. At least 4 weeks and also recovery from effects of major prior surgery or other therapy, including radiation therapy, immunotherapy, cytokine, biologic or vaccine therapy
7. Adequate organ function determined within 2 weeks prior to randomization, defined as:
   a. Absolute neutrophil count (ANC) = 1500/mm3, platelet counts = 100,000/mm3, and hemoglobin = 8 g/100 mL without need for hematopoietic growth factor or transfusion support
   b. Serum creatinine = 1.5 x ULN, or 24-hour creatinine clearance = 50 cc/min. (Note: Creatinine clearance need not be determined if the baseline serum creatinine is within normal limits.)
   c. Serum bilirubin = 1.5 x ULN; aspartate amino transferase (AST) = 2.5 x ULN; alanine amino transferase (ALT) < 2.5 x ULN; alkaline phosphatase = 2.5 x ULN. Serum albumin must be = 2.5 g/dL. No exceptions will be granted.
   d. Prothrombin time (PT) = 1.5 x ULN (or INR =1.3) and partial thromboplastin time (PTT) = 1.5 x ULN. Elevated PT, > 1.5 x ULN or INR > 1.3 for any reason excludes a patient.
8. Satisfactory venous access
9. Intellectual, emotional, and physical ability to maintain an ambulatory infusion pump (required if the patient is randomized to Arm B).

4.2 EXCLUSION CRITERIA

Exclusion criteria included:

1. Prior cytotoxic chemotherapy, including regional perfusion
2. History of brain metastases or leptomeningeal disease
3. Significant medical disease other than cancer including: uncontrolled congestive heart failure; active symptoms of coronary artery disease (defined as uncontrolled arrhythmias or recurrent chest pain despite prophylactic medication); New York Heart Association class III or IV disease; cardiovascular signs and symptoms = Grade 2 by CTC criteria during the 4-week period before protocol drug therapy; uncontrolled seizure disorder; history of chronic hepatitis or cirrhosis; active infection; uncontrolled diabetes mellitus; requirement for chronic corticosteroid treatment with an average dose =20 mg/day of prednisone (or equivalent); requirement for concurrent immunosuppressive drug(s); active autoimmune disease
4. Organ allografts
5. Prior radiotherapy, or prior intratumor injection therapy, to areas of measurable disease that are used as target indicator lesions, unless progression has occurred at that site or measurable disease has developed outside the treated area
6. Known HIV-infection
7. Pregnancy or lactation
   - Women of childbearing potential and sexually active males were advised to take precautions to prevent pregnancy during treatment
   - All patients were required to use an effective form of birth control while on treatment (except if a woman was post menopausal of = 1 year or surgically sterile)
8. History of second cancer (except for adequately treated basal cell or squamous cell skin cancer, in situ cervical cancer, or other cancer for which the patient has been disease-free for five or more years)
9. Known hypersensitivity to phosphorothioate-containing oligonucleotides or to DTIC
10. Bone-only metastatic disease (without other measurable disease)
11. Primary ocular or mucosal melanoma
12. Use of any experimental therapy within 3 weeks prior to baseline evaluations done prior to randomization
13. Concomitant anticoagulant therapy was not permitted (with the exception of 1 mg/day of warfarin for central line prophylaxis). INR must be = 1.3.

4.3 STRATIFICATION

Patients were stratified according to the following criteria:

1. ECOG performance status = 0 or ECOG = 1 – 2
2. Skin, subcutaneous/lymph node metastases without visceral metastases and normal LDH - versus any visceral metastases or elevated LDH*
3. Liver metastases versus no liver metastases

*LDH elevation had to be at least 10% over the upper limit of normal, and based on the most recent screening evaluation done for this protocol, and was to be without other known medical causes unrelated to metastatic melanoma.

4.4 SCREENING ASSESSMENT

The following evaluations were required within 4 weeks of treatment. All abnormal and normal results were to be noted in the Case Report Form and the source document.

1. Contrast-enhanced CT or MRI of the head, and contrast-enhanced CT of the chest and abdomen. [If the patient had a history of melanoma distal to the thorax, a CT of the pelvic area was also to be performed at baseline.] Ultrasound was not acceptable for measurement of target lesions.

Note: Spiral CT was preferred for measurement of target indicator lesions that are visceral or nodal/soft tissue. Follow-up of target lesions used the same technique established to measure the target lesion(s) at baseline.

2. Chest X-ray
3. Medical history to include determination of tumor-related symptoms
4. Physical examination to include patient weight
5. Calculation of body surface area (BSA)
6. ECOG Performance Status
7. Complete blood count (CBC) with differential and platelet count
8. Serum laboratory studies including: bilirubin, AST, ALT, LDH, alkaline phosphatase, albumin, creatinine, BUN, glucose, calcium, phosphorus
9. PT (or INR) and PTT
10. Pregnancy test (only in women of childbearing potential)
11. Electrocardiogram
12. Urinalysis
13. Determination and measurement of "Target" and "Non-target" qualifying indicator lesions. Tumor measurements of target superficial skin metastases were obtained or updated < 1 week prior to protocol therapy

Measurable disease was defined as the presence of at least one measurable lesion.

Measurable lesions were defined as lesions that could be accurately and serially measured in at least one dimension and for which the greatest diameter was = 20 mm as measured by contrast-enhanced or spiral CT scan for visceral or nodal/soft tissue disease, or = 10 mm as measured by caliper for superficial cutaneous metastases.

4.5 CONCOMITANT TREATMENT

Patients were not allowed to receive any other anti-cancer treatments (such as chemotherapy, radiation, biologic or investigational therapies) while receiving protocol therapy.

4.5.1 Treatment

Treatment was to start within 14 calendar days after randomization. All patients were treated in 21-day cycles. Protocol therapy and related testing were completed in the outpatient setting. Inpatient treatment and testing were permitted at the discretion of the investigator. Only patients in the combination arm had ambulatory infusion pumps, preferably with central venous access (e.g., PICC line, portacath, etc.), in order to deliver the G3139 infusion. Treatment on this protocol continued for 8 cycles in responding or stable patients. Patients who achieved a complete response, partial response or stable disease continued on protocol therapy until one of the following events occurred: disease progression, onset of intolerable drug-related side effects that did not respond to dose modification or delay as noted below, or completion of 8 cycles of therapy.

4.5.2 Efficacy Parameters

4.5.2.1 Primary Efficacy Parameter:

Survival, defined as the time from randomization to death. Patients who had not reached the event were censored at the time of last follow-up or data cut-off date of this analysis.

4.5.2.2 Secondary Efficacy parameter definitions

4.5.2.2.1 Response

Response was defined by standard RECIST criteria. Confirmation was required at least three weeks (one cycle) or up to 6 weeks later. "Durable" responses were those lasting over 6 months. The duration of response was dated from the time that measurement criteria were first met for CR or PR (whichever was first recorded), until the first date that recurrent or progressive disease was objectively documented. A subject was considered to have achieved a response if either complete response or partial response was initially documented within 8 cycles of beginning protocol therapy in the absence of nonprotocol antitumor therapy.

4.5.2.2.2 Progression free survival

Progression was also defined by standard RECIST criteria

- If progression was suggested by increased size of one or more target indicator lesions, without evidence of clearly defined new metastases, then all target lesions were measured in order to confirm progressive disease by use of the sum of the longest diameter (LD).
- If progression was indicated by well-defined development of new metastatic disease, repeat evaluation of all target lesions was not necessary but may be done at the discretion of the investigator.
- If tumor progression or the discovery of metastatic disease was equivocal, the determination of tumor progression, stable disease, or response was determined by sum LD evaluation of target lesions that were documented at baseline. Examples of equivocal tumor progression included detection of

metastatic disease in a body region that was not clearly tumor-free at baseline, or detection of an abnormality that could not be confirmed to be new metastatic melanoma. In this situation, the nature of the equivocal finding, suggesting progressive or new metastatic disease, was recorded in the source document and in the Case Report Form, along with imaging by CT or photography as appropriate.

Patients with a global deterioration of health status that requires discontinuation of treatment without objective evidence of disease progression were reported as "symptomatic deterioration." Every effort was to be made to document the objective progression even after discontinuation of treatment.

Progression-free survival (PFS) was measured from the date of randomization until the date of criteria for progression (date of scan or measurement) was met or the patient died. The sponsor performed sensitivity analyses of progression-free survival by using several different approaches:

- Incomplete lesion measurements were imputed by averaging the 2 measurements that were collected immediately before and after the missing data.
- Time to treatment failure analysis: Subjects who discontinued from study prematurely due to progression/relapse, intercurrent illness, adverse drug toxicity, and death also were included as events.
- Time to progression analysis: Events included only those who met progression-of-disease criteria based on lesion assessment.
- For subjects whose response at last target lesion measurement was complete response, partial response, or stable disease, progression-free survival was censored at 60 days from last lesion measurement.
- When different lesions were measured on different dates in a given cycle, the earliest lesion measurement date instead of the last lesion measurement date was used.
- The number of subjects with progression of disease included those who met progression of disease criteria based on lesion assessment during the treatment phase only or subjects who died within 60 days of the last on-treatment lesion assessment.
- When a new nontarget lesion was observed or a nontarget lesion progressed in the presence of clear cut progression, the overall response was determined as progression of disease when the target lesion measurement was absent or indicated stable disease or response.

4.5.2.3 Missing and Incomplete Data

Missing data patterns were examined by the sponsor by summarizing the number of subjects randomized but not treated, the number of treated subjects who discontinued prematurely, and the corresponding reasons for early discontinuation. For antitumor response-related data that were incomplete, the sum of target lesions was calculated

based on the last available lesion measurements. Progression of disease status was determined by target lesion measurements when at least one target lesion was measured at the visit.

4.5.2.4 Confirmed response

A confirmed response, complete or partial, indicated that the response measurement was confirmed by the investigator by a repeat study done no less than 3 weeks after response was first noted. RECIST criteria normally specify a re-evaluation at an interval of 4 weeks. The sponsor termed the 3 week re-assessment interval as "modified" RECIST. The primary difference between the modified RECIST (used in this protocol) and the RECIST was a change in the confirmatory evaluation period from 4 weeks to 3 weeks in order to coincide with the schedule of study visits specified by protocol.

4.5.2.5 Independent, Blinded Assessment of Response

RadPharm (Princeton, NJ) was contracted to conduct a blinded assessment of response. They independently reviewed the data of subjects who were identified as having a response (i.e., response assessed as complete response or partial response) on the basis of lesion measurements by the investigator.

5 RESULTS

A total of 771 subjects were randomized at 139 study sites in 9 countries: the G3139 plus DTIC arm consisted of 386 patients and the control arm of DTIC alone comprised 385 patients. One hundred U.S. sites enrolled 54% of the patients. The distribution of subjects by age and race was similar in the 2 treatment groups for the Intent-to-Treat Population. There were almost 5% more women in the G3139 arm. The mean age in both groups was nearly 59 years; females represented a little more than one third of the subjects randomized to each group. More than 95% of subjects in both treatment groups were white. A total of 731 patients of the 771 randomized, including 371 (96.1%) subjects in the G3139 plus DTIC group and 360 (93.5% subjects) in the DTIC group, initiated protocol therapy. See Section 5.2, demographics.

5.1 ANALYSIS POPULATIONS

The ITT population was the primary efficacy population. All patients who received any therapy constituted the safety population. A per-protocol population and other subsets were examined by the sponsor for secondary variables and for sensitivity analyses. The ITT population included 386 subjects in the G3139 plus DTIC group and 385 subjects in the DTIC group (total N = 771). The Per-Protocol Population included 357 (92.5%) subjects in the G3139 plus DTIC group and 353 (91.7%) subjects in the DTIC group. The Safety Population of 731 included 371 (96.1%) subjects in the G3139 plus DTIC group and 360 (93.5%) subjects in the DTIC group who received one or more doses of medication. Untreated subjects included 15 (3.9%) subjects in the G3139 plus DTIC group and 25 (6.5%) subjects in the DTIC group. The final data cutoff date was August 1, 2003.

5.2 PATIENT DEMOGRAPHICS

Table 2 Demographic and Background Characteristics (ITT)

Study GM301	G + DTIC N = 386	DTIC N = 385	TOTAL N= 771
Gender	N (%)	N (%)	N (%)
Men	236 (61.1)	253 (65.7)	489 (63.4)
Women	150 (38.9)	132 (34.3)	282 (36.6)
Age (years)			
Median	59.0	60.0	59.0
Range	(17 - 93)	(16 - 89)	(16 - 93)
< 65	239 (61.9)	241 (62.6)	480 (62.3)
≥ 65	147 (38.1)	144 (37.4)	291 (37.7)
≥ 75	47 (12.2)	54 (14.0)	101 (13.1)
≥ 85	4 (1.0)	6 (1.6)	10 (1.3)
Race			
Caucasian	375 (97.2)	371 (96.4)	746 (96.8)
Black	1	0	1 (0.1)
Asian	2	1	3 (0.4)
Hispanic	7 (1.8)	11 (2.9)	18 (2.3)
Other and unknown [a]	1	2	3 (0.4)
Weight (kg)			
N	371	359	730
Median	80.4	80.5	80.4
Range	43 - 174	48 – 143	43 - 174
Body surface area (m^2)			
N	369	358	727
Median	1.96	1.96	1.96
Range	1.35 – 3.09	1.43 – 2.66	1.35 – 3.09
ECOG Performance Status [b]			
Missing	8	2	10
0	207 (55)	220 (57)	427 (56)
1	146 (39)	132 (34)	278 (37)
2	24 (6.3)	29 (7.6)	53 (7.0)
3	1	2	3 (0.4)
Prior Immunotherapy or Cytokine Therapy	156 (40.4)	142 (36.9)	298 (38.7)

[a] Recording of race was not permitted at the study site.
[b] Status at baseline
Revised Sponsor's Table

The majority of the patients were not symptomatic at study entry. There were no reported differences in symptoms by study arm. The following table shows the four pre-study symptoms solicited by the sponsor to assess tumor-related symptoms and disease burden.

Table 3 Symptoms at Baseline

Baseline Tumor-Related Symptoms Intent-to-Treat Population	**G3139 + DTIC** (N=386)	**DTIC Alone** (N=385)	**Overall** (N=771)
	n (%)	n (%)	n (%)
Pain			
No	258 (66.8)	257 (66.8)	515 (66.8)
Yes	125 (32.4)	124 (32.2)	249 (32.3)
CTC Grade 1 - Mild	62 (16.1)	61 (15.8)	123 (16.0)
CTC Grade 2 - Moderate	52 (13.5)	46 (11.9)	98 (12.7)
CTC Grade 3 - Severe	11 (2.8)	17 (4.4)	28 (3.6)
CTC Grade 4 - Disabling	0 (0.0)	0 (0.0)	0 (0.0)
Skin Bleeding			
No	370 (95.9)	377 (97.9)	747 (96.9)
Yes	13 (3.4)	4 (1.0)	17 (2.2)
CTC Grade 1 - Mild	11 (2.8)	1 (0.3)	12 (1.6)
CTC Grade 2 - Moderate	2 (0.5)	1 (0.3)	3 (0.4)
CTC Grade 3 - Severe	0 (0.0)	2 (0.5)	2 (0.3)
CTC Grade 4 - Disabling	0 (0.0)	0 (0.0)	0 (0.0)
Pruritus			
No	376 (97.4)	378 (98.2)	754 (97.8)
Yes	7 (1.8)	3 (0.8)	10 (1.3)
CTC Grade 1 - Mild	4 (1.0)	3 (0.8)	7 (0.9)
CTC Grade 2 - Moderate	3 (0.8)	0 (0.0)	3 (0.4)
CTC Grade 3 - Severe	0 (0.0)	0 (0.0)	0 (0.0)
CTC Grade 4 - Disabling	0 (0.0)	0 (0.0)	0 (0.0)
Other Symptoms			
No	284 (73.6)	290 (75.3)	574 (74.4)
Yes	99 (25.6)	91 (23.6)	190 (24.6)

Sponsor's table 14.1.14, GM301 CSR

5.3 BASELINE DISEASE CHARACTERISTICS

The two tables below show the sponsor's assignment of patients based on enrollment information. There were six strata, and subjects were randomly assigned therapy based on the findings at time of randomization by interactive voice response system (IVRS) into one of these six groups. The first tabulation below was based on the IVRS assignment distribution at the time of randomization. However, a second tabulation of patients' actual characteristics was based on review of the Case Report Form information after treatment had begun. Some differences were found in the patient's prognostic features between the IVRS randomization descriptions and the information from the CRF review on study. The sponsor suggested that the reason for this was the time interval between randomization and start of therapy, in which an additional lab profile was obtained within three days of starting therapy:

"These differences were not unexpected given the lapse of time between collection of data used in the IVRS randomization and study assessments at baseline, the extent of illness in this subject population, and the diagnostic studies required prior to initiation of protocol therapy…. Accordingly, when eCRF data were used to stratify subjects retrospectively, some subjects in both treatment groups shifted to another stratum."

The sponsor also asserted that:

"Nevertheless, subjects remained evenly distributed between treatment groups in each of the 6 strata. In both treatment groups, the stratum with the most subjects included subjects with an ECOG score of 0 and disease in visceral organs other than liver or elevated LDH (G3139 plus DTIC group, 113 [29.3%] subjects; DTIC, 120 [31.2%] subjects). No imbalance was observed between treatment groups."

TABLE 14.1.15.1.1
Stratification Used in IVRS Randomization
Intent-to-Treat Population

Stratum	ECOG Score	Liver Metastases	Disease Distribution and LDH status	G3139 + DTIC (N=386) n (%)	DTIC Alone (N=385) n (%)	Overall (N=771) n (%)
1	0	Present	Not Applicable	61 (15.8)	61 (15.8)	122 (15.8)
2	1 or 2	Present	Not Applicable	66 (17.1)	67 (17.4)	133 (17.3)
3	0	Absent	Disease in Other Visceral Organ or Elevated LDH	113 (29.3)	112 (29.1)	225 (29.2)
4	1 or 2	Absent	Disease in Other Visceral Organ or Elevated LDH	87 (22.5)	86 (22.3)	173 (22.4)
5	0	Absent	Disease in the Skin, Subcutaneous and/or Lymph Without Visceral Metastases and Without Elevated LDH	34 (8.8)	34 (8.8)	68 (8.8)
6	1 or 2	Absent	Disease in the Skin, Subcutaneous and/or Lymph Without Visceral Metastases and Without Elevated LDH	25 (6.5)	25 (6.5)	50 (6.5)

Sponsor's table 14.1.15.1.1, CSR: ITT Population stratification by IVRS Assignment

Stratification Based on CRF Data
Intent-to-Treat Population

Stratum	ECOG Score	Liver Metastases	Disease Distribution and LDH status	G3139 + DTIC (N=386) n (%)	DTIC Alone (N=385) n (%)	Overall (N=771) n (%)
	Unknown	Absent	Disease in Other Visceral Organ or Elevated LDH	0 (0.0)	1 (0.3)	1 (0.1)
		Present	Not Applicable	2 (0.5)	0 (0.0)	2 (0.3)
1	0	Present	Not Applicable	70 (18.1)	63 (16.4)	133 (17.3)
2	1 or 2	Present	Not Applicable	58 (15.0)	73 (19.0)	131 (17.0)
3	0	Absent	Disease in Other Visceral Organ or Elevated LDH	113 (29.3)	120 (31.2)	233 (30.2)
4	1 or 2	Absent	Disease in Other Visceral Organ or Elevated LDH	82 (21.2)	78 (20.3)	160 (20.8)
5	0	Absent	Disease in the Skin, Subcutaneous and/or Lymph Without Visceral Metastases and Without Elevated LDH	36 (9.3)	32 (8.3)	68 (8.8)
6	1 or 2	Absent	Disease in the Skin, Subcutaneous and/or Lymph Without Visceral Metastases and Without Elevated LDH	25 (6.5)	18 (4.7)	43 (5.6)

Sponsor's table 14.1.16.1, CSR: ITT Population stratification by CRF Data

Reviewer's comments: *See next table. There are imbalances in the prognostic factors between the two study arms. Almost half of all patients had favorable baseline sites of disease in skin, soft tissue or lung (366, 47.5%) for the ITT population but more of these were in the combination arm. There were fewer subjects with liver involvement or visceral-disease-other than-lung-and-elevated-LDH in the G3139 plus DTIC group (226, 58.5%) than in the DTIC group (257, 66.8%) (p = 0.02). An analysis of quantity of tumor present was estimated by comparing the baseline sum LD (sum of the long diameters of each measurable lesion) measurements for each study arm. The median sum LD for the G3139 plus DTIC arm was 65 cm versus 75 cm for DTIC alone, suggesting a greater uni-dimensional quantity of tumor at baseline in the DTIC alone group.*

Table 4 Baseline Disease Characteristics

N = 771	G3139 + DTIC	DTIC Alone
Metastasis [a]	(N=386)	(N=385)
	n (%)	n (%)
Local regional disease only	34 (8.8)	22 (5.7)
Distant metastatic disease	352 (91.2)	363 (94.3)
LDH and disease site distribution: [b]		
Non-visceral and non-elevated LDH	61 (15.8)	50 (13.0)
Lung and non-elevated LDH	93 (24.1)	75 (19.5)
Visceral other than lung, or elevated LDH	226 (58.5)	257 (66.8)
Unknown	6 (1.6)	3 (0.8)

Reviewer's Table

[a] Both local and distant were counted as distant.

[b] Based on the AJCC classification.

Bolded = favorable prognostic factor for melanoma

A baseline LDH value was not available for 6 subjects in the G3139 plus DTIC group and for 3 subjects in the DTIC group.

Prior Therapy:

On the basis of prior therapy received, the two study arms appeared to be equally balanced.

Table 5 Prior Therapy

	Prior Surgery, Radiotherapy and Immunotherapy/Cytokine Therapy by Analysis Populations		
	G3139 + DTIC	DTIC Alone	Overall
Intent-to-Treat Population	(N=386)	(N=385)	(N=771)
Prior Surgery n (%)	369 (95.6)	372 (96.6)	741 (96.1)
Prior Radiation Therapy n (%)	73 (18.9)	65 (16.9)	138 (17.9)
Prior Immunotherapy/Cytokine Therapy for Metastatic Disease n (%)	156 (40.4)	142 (36.9)	298 (38.7)
Adjuvant Immunotherapy			
Yes	109 (28.2)	107 (27.8)	216 (28.0)
No	69 (17.9)	45 (11.7)	114 (14.8)

Reviewer's Table

The following table shows the disposition of patients during the trial.

Table 6 Patient Disposition

	G3139 + DTIC	DTIC Alone	Overall
Subject Disposition	(N=386)	(N=385)	(N=771)
Intent-to-Treat Population	n (%)	n (%)	n (%)
Randomized (Intent-to-Treat)	386 (100.0)	385 (100.0)	771 (100.0)
Randomized But Not Treated	15 (3.9)	25 (6.5)	40 (5.2)
Treated (Safety Population)	371 (96.1)	360 (93.5)	731 (94.8)
Per Protocol Population	357 (92.5)	353 (91.7)	710 (92.1)
Completion of Protocol-Specified Therapy	50 (13.0)	56 (14.5)	106 (13.7)
Continuing on protocol-specified therapy (study GM214)	10 (2.6)	1 (0.3)	11 (1.4)
Discontinued Early	326 (84.5)	328 (85.2)	654 (84.8)
Primary Reason for Early Discontinuation	326 (84.5)	328 (85.2)	654 (84.8)
Disease Progression or Relapse	242 (62.7)	266 (69.1)	508 (65.9)
Subject Choice	20 (5.2)	27 (7.0)	47 (6.1)
Intercurrent Illness	3 (0.8)	2 (0.5)	5 (0.6)
Adverse Drug Toxicity	26 (6.7)	8 (2.1)	34 (4.4)
Grade 3 or 4 Hematological Toxicity	6 (1.6)	1 (0.3)	7 (0.9)
Grade 3 or 4 Non Hematological Toxicity	8 (2.1)	5 (1.3)	13 (1.7)
Other	12 (3.1)	5 (1.3)	17 (2.2)
Subject Noncompliance	0 (0.0)	0 (0.0)	0 (0.0)
Subject Ineligible	5 (1.3)	3 (0.8)	8 (1.0)
Protocol Violation	1 (0.3)	5 (1.3)	6 (0.8)
Prohibited Drug or Biologic Agent	0 (0.0)	0 (0.0)	0 (0.0)
Other Prohibited Additional Therapy	0 (0.0)	1 (0.3)	1 (0.1)
Other	1 (0.3)	4 (1.0)	5 (0.6)
Death	19 (4.9)	1(3.9)	34 (4.4)
Lost to Follow-up	0 (0.0)	0 (0.0)	0 (0.0)
Other	5 (1.3)	2 (0.5)	7 (0.9)
Drug Administration Noncompliance	5 (1.3)	0 (0.0)	5 (0.6)

Sponsor's Table from Study Report for GM301

Less than half (44%) of the patients remained on study beyond 2 cycles (see Table 7). Most discontinuations of study treatment by Day 42 reflected disease progression.

Table 7 Treatment Discontinuation through First Two Cycles

	G + D	D	Total
ITT (Randomized) N =	386	385	771
Less number not treated	15	25	40
Discontinued by day 42	185	209	394
Still on study after day 42	186	151	337

Reviewer's table (Compilation of sponsor's tables 14.1.4 and 14.1.5)

The following table shows the disposition of patients by treatment group.

Table 8 Disposition of subjects by treatment group: Intent-to-Treat Population

	G3139 plus DTIC	DTIC	Overall
	(N = 386)	(N = 385)	(N = 771)
	n (%)[a]	n (%)[a]	n (%)[a]
Randomized	386 (100.0)	385 (100.0)	771 (100.0)
Completed protocol therapy (8 cycles)	50 (13.0)	56 (14.5)	106 (13.7)
Discontinued early [b]	326 (84.5)	328 (85.2)	654 (84.8)
Continuing on protocol therapy [c]	10 (2.6)	1 (0.3)	11 (1.4)

Reviewer's Table

[a] Percentages were calculated by using N, the total number of subjects in the group, as the denominator.

[b] Discontinued early was defined as those subjects who did not complete all 8 cycles of protocol therapy.

[c] At time of data cut-off date of Aug. 1, 2003

5.4 EFFICACY

5.4.1 Overall Survival

According to the Statistical Analysis Plan, survival analysis was to be performed after 508 deaths occurred. It was actually performed when 537 deaths (~ 70%) had occurred. Of the 537 deaths, two occurred after 2 years of follow-up, so survival times for these two patients were censored at the end of 2 years. The following table is the sponsor's primary analysis result. There was no statistically significant difference between the two treatment arms with respect to the primary endpoint overall survival (unadjusted log-rank test, p-value = 0.18). The estimated median survivals were 238 and 274 days respectively, in the DTIC and G3139 + DTIC treatment arms. Figure 3 is the sponsor's Kaplan-Meier survival curves. The curves are nearly superimposable for the first 6 months. According to the sponsor's study report, the result based on the per-protocol population was similar to that based on the intent-to-treat population.

Table 9: Survival Time analysis (Intent-to-Treat Population)

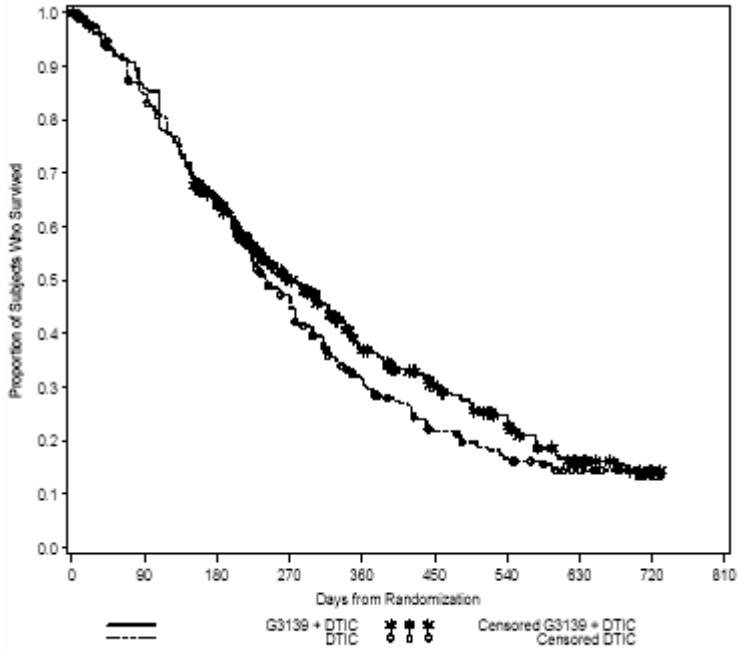
[*Source: Sponsor's Table 16 in Section 11.2.1*]

Parameter	G3139+DTIC (N = 386)	DTIC (N = 385)	Hazard ratio[b]	Log-rank P value
Number (%)[a] of subjects who died	266 (68.9)	269 (69.9)		
Median survival time (days)	274 [9.0 mo.]	238 [7.8 mo.]	0.89	0.18
95% Confidence interval	232 –306 [7.6 – 10.0 mo.]	219 – 271 [7.2 – 8.9 mo.]	0.75 – 1.06	

Cross-reference: Table 14.2.2.1; Listing 16.2.6.1
Survival time was defined as the time from the date of randomization to the date of death.
[a] Percentages were calculated by using N, the total number of subjects in the group, as the denominator.
[b] Hazard ratio and 95% confidence interval from the unadjusted Cox Proportional Hazards Model

Figure 3: Kaplan-Meier Survival Curves by Treatment Group (Intent-to-Treat Population)

[*Source: Sponsor's Figure 4 in Section 11.2.1*]



Cross-reference: Figure 14.2.1.1

The sponsor reported that the median follow-up time was 7 months (212 days) at the time of survival analysis.

A post-hoc analysis was performed by the sponsor that "included all patients who had been randomized on or before 01 August 2002 and thus had an opportunity for a minimum of 12 months of follow-up." A total of 520 patients were included in such an

analysis. The sponsor termed these 520 patients as 'intent-to-treat cohort.' The following table is the result of the sponsor's post-hoc analysis for this cohort. It did not suggest any statistically significant difference between the two treatment groups in this 'intent-to-treat cohort' population. Although the point estimate of the median survival time for the G3139+DTIC group was prolonged by 27 days while the point estimate for the DTIC group remained unchanged, the corresponding Kaplan-Meier curves (Figure 4) were visually similar to that performed on the intent-to-treat population.

Table 10: Survival time analysis: Intent-to-Treat Cohort randomized on or before 01 Aug 2002 (N = 520)

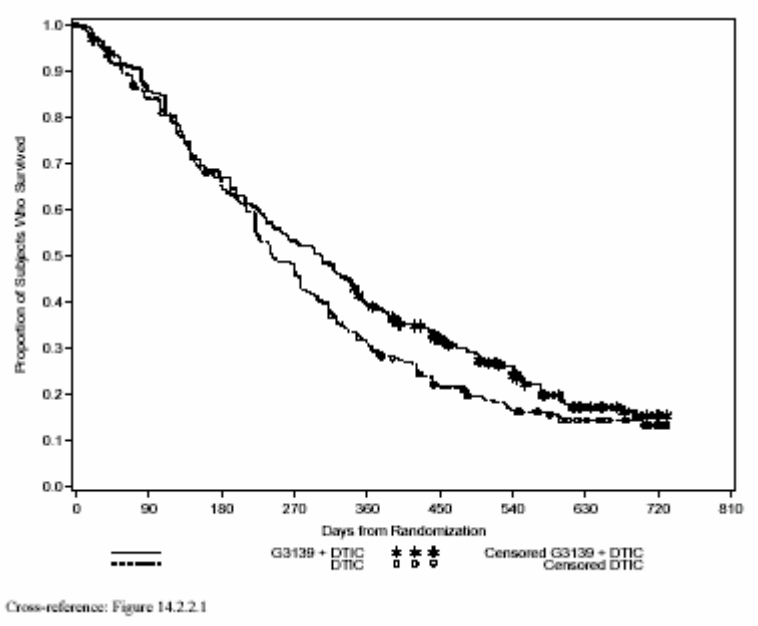
[*Source: Sponsor's Table 18 in Section 11.2.1*]

Parameter	**G3139+ DTIC (N = 260)**	**DTIC (N = 260)**	**Hazard ratio[b]**	**Log-rank P value**
Number (%)[a] of subjects who died	201 (77.3)	212 (81.5)		
Median survival time, days	302	238	0.84	0.066
95% Confidence interval	243 – 336	219 – 274	0.69-1.01	

Cross-reference: Table 14.2.2.2; Listing 16.2.6.1
Survival time was defined as the time from the date of randomization to the date of death.
[a] Percentages were calculated by using N, the total number of subjects in the group, as the denominator.
[b] Hazard ratio and 95% confidence interval from the unadjusted Cox Proportional Hazards Model

Figure 4: Kaplan-Meier Survival Curves by Treatment Group: Intent-to-Treat Cohort Randomized on or Before 01 Aug 2002 (N = 520)

[*Source: Sponsor's Figure 5 in Section 11.2.1*]



Cross-reference: Figure 14.2.2.1

The sponsor further performed an analysis on the "per-protocol cohort" of the "cohort randomized on or before August 1, 2002" and obtained a p-value of 0.035 and concluded that the survival time for the G3139+DTIC group was significantly longer than in the DTIC group in this subset.

REVIEWER COMMENTS:

[1] Efficacy evaluation should be solely based on the pre-specified primary analysis. Post-hoc analyses do not demonstrate efficacy and can only be considered exploratory.

[2] According to the sponsor's protocol, there was only one survival analysis, and it was to be performed when a total of 508 deaths had occurred. Survival analysis was actually performed when 537 deaths had occurred. At the completion of the study, the sponsor failed to demonstrate survival benefit in the G3139+DTIC group using the pre-specified primary analysis. This study failed its primary endpoint, to show a survival benefit.

[3] In addition, the sponsor's subgroup analysis of survival suggested that in this study the effect of G3139 + DTIC in overall survival seemed to be mainly driven by results from the non-US sites. For more details, please refer to the section entitled Other Special/Subgroup Populations.

5.4.2 Progression-Free Survival (1): Sponsor's Results and Reviewer's Comments

Table 11 is the sponsor's result of analysis of progression-free survival when missing data was imputed using the methodology of last-observation-carried-forward and

Figure 5 is the sponsor's Kaplan-Meier progression-free survival curves. Per the sponsor's analysis, a total of 523 patients (about 68% of the patients in each treatment group) had disease progression or death. Most patients had an event due to disease progression rather than death. A total of 248 (32%) of the patients were censored. Of those who were censored, 139 patients (18.0%; 65 from the G3139 + DTIC group and 74 from the DTIC group) had no lesion measurements or had lesion measurements at baseline only. Progression-free survival time for those patients was censored on Day 1. For the remaining 109 (14%) patients, progression-free survival time was censored at the last lesion assessment; of these patients, follow-up was continued on 58 (7.5%) patients at the time of analysis.

Based on the sponsor's analysis using the last-observation-carried-forward for missing data, the point estimate of the median progression-free survival was longer in the G3139+DTIC group, 74 days vs. 49 days (a difference of 25 days). The p-value from the logrank test comparing the entire curves between treatment groups was 0.0003. The sponsor also performed several other analyses (sensitivity or exploratory analyses) which resulted in very small p-values (≤ 0.001). The sponsor concluded that progression-free survival time was statistically significantly longer in the G3139+DTIC group than in the DTIC group.

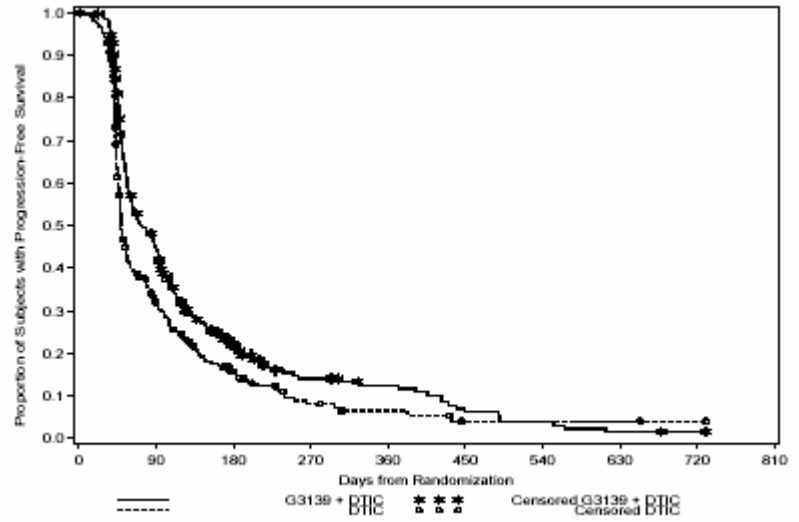
Table 11: Progression-Free Survival Analysis (Intent-to-Treat Population)

[*Source: Sponsor's Table 20 in Section 11.2.2.1*]

Parameter	G3139+DTIC (N = 386)	DTIC (N = 385)	Hazard ratio[b]	Log-rank P value
Number (%)[a] of subjects with disease progression or death	263 (68.1)	260 (67.5)		
Disease progression	257 (66.6)	255 (66.2)		
Death	6 (1.6)	5 (1.3)		
Median progression-free survival, days	74	49	0.73	0.0003
95% Confidence interval	62-91	48-55	0.62-0.87	

Cross-reference: Table 14.2.9; Listing 16.2.6.2.1
Time to progression was calculated from the date of randomization to the date progression was first documented or the subject died. The last observation carried forward was used for missing data.
[a] Percentages were calculated by using N, the total number of subjects in the group, as the denominator.
[b] Hazard ratio and 95% confidence interval from the unadjusted Cox Proportional Hazards Model

Figure 5: Kaplan-Meier Progression-Free Survival Curves by Treatment Group (Intent to-Treat Population) – (Sponsor's Figure)



Cross-reference: Figure 14.2.10

Determination of disease progression depended on assessment per RECIST criteria of multiple (target or nontarget) lesions. For target lesions, the protocol specified that the investigator was required to calculate the sum of longest diameters. If a target lesion was missing, then in general determination of disease progression could not be made without imputing the missing data unless a new lesion was observed or disease progression was observed in nontarget lesions. In order to explore how much influence missing data could have on the analysis result, the Agency requested the sponsor on 02 February 2004 to clarify how many observations were carried forward for the analysis and to reanalyze the data using a different approach: censoring progression data on the date of last complete observation when at least 50% of the target lesions had been evaluated. Table 12 and Table 13 were the sponsor's responses. Per the sponsor's analysis, the last-observation-carried-forward methodology was applied to only 26 patients (14 from the G3139 + DTIC group and 12 from the DTIC group). The median number of days that were carried forward for these patients was 142 in the G3139 + DTIC group and 127 in the DTIC group. As presented in Table 13, the difference in estimated medians was 13 days and the result of sponsor's reanalysis still showed a highly statistically significant difference between the two treatment groups.

Table 12: Summary of Missing Data in Analysis of Progression-Free Survival

[*Source: Sponsor's response in fax dated 18 February 2004*]

	G3139 + DTIC [N = 386]	DTIC [N = 385]	Data Handling Rule
No post baseline target lesion	65	74	Censor at Day 1 (no LOCF)
First missing target lesion(s) occurred at the same time PD was evidenced by other lesion(s)	29	26	PD at the same visit (no LOCF)
Missing one or more target lesion(s) prior to PD	14[†]	12[‡]	LOCF to the next measurement or the time when at least one target lesion was evaluated
Total	108	112	

[†] Mean [median] days carried forward in the TTP analysis for these 14 patients was 164 [142].
[‡] Mean [median] days carried forward in the TTP analysis for these 12 patients was 133 [127].

Table 13: Analysis of Progression-Free Survival with Data Censored at Date of Last Complete Observation

[*Source: Sponsor's response in fax dated 18 February 2004*]

	G3139 + DTIC [N = 386]	DTIC [N = 385]	Hazard Ratio	Log-rank P value
Median time to progression days)	61	48	0.75	0.0006

Note: P-value was from logrank test comparing the entire curves between treatment groups.

REVIEWER COMMENTS:

[1] Although the sponsor's analysis suggested a statistically significant difference in progression-free survival between the two treatment groups, it is not clear whether this is a true finding. The major issues are summarized in [2] – [4]. **For more detailed comments, please refer to Section 5.4.3 for this reviewer's evaluations and exploratory analyses.** Given the open-label nature of study, missing data, and differences in assessment intervals between the two treatment groups, without replication of the results in a second well-controlled study, the results from secondary endpoint analysis may be a false positive result.

[2] Lesions were measured periodically. The protocol (and the sponsor) defined progression as the date on which a scan or measurement was made, not the date of an office visit. Disease progression generally does not occur on the date of evaluation but rather prior to that date. If the intervals between two consecutive assessments (termed assessment intervals) are longer on one study arm, the documented date of

disease progression would potentially be delayed; hence, the observed progression-free survival time would tend to be prolonged. Similarly, if the first assessment date is delayed, the observed progression-free survival would also tend to be systematically prolonged even if the assessment intervals remain the same. In this study, although the assessment schedules were intended to be the same between the two treatment groups, because of the different nature of treatment schedules (G3139 plus DTIC requiring 6 days versus DTIC alone only 1-hour infusion), it was observed that the assessment schedule for patients in the G3139 + DTIC group was generally slightly delayed compared to that for patients in the DTIC group. The difference could be very subtle. However, in a large trial, statistical results of comparing the two schedules suggested that a difference was present. Even a subtle difference in assessment schedule between two treatment groups could very likely lead to a false positive conclusion when in fact there is no difference in progression-free survival between the two treatment groups. This is illustrated by simulation studies conducted by the statistical reviewer to evaluate the impact of different assessment schedules on the false positive rate. For detailed simulation results, please refer to the Appendix.

[3] For patients who had disease progression or death during the study, although missing target lesion measurements were observed in only some patients (6% of progressors in the G3139 + DTIC group vs. 5% of progressors in DTIC group), missing nontarget lesion measurements were observed in many patients prior to disease progression or death (34% of progressors in the G3139 + DTIC group vs. 23% of progressors in DTIC group). In some of the patients who had post baseline measurements but had no disease progression or death during the study, considerable missing nontarget lesion measurements were also observed. It was also observed that in some patients there were some cycles when no target lesions were measured, but nontarget lesions were measured. These cycles were ignored by the sponsor when determining presence or absence of disease progression because the determination was only based on cycles when there were target lesion measurements. These various problems of missing data could lead to a bias in estimating the treatment effect.

[4] Since disease progression was assessed periodically, it may not be appropriate to treat data solely as right-censored data especially when there was difference in assessment schedules and there were considerable missing data. This type of data for progression-free survival can be better presented in the form of interval-censored data in that if disease progression is observed in a given assessment, the date of occurrence can be on any day between the previous assessment and this assessment. Four approaches were used by this reviewer to obtain the interval for each patient. Using generalized log-rank test with interval censoring, it was observed that the statistical significance diminished after taking into account the uncertainty of missing values and different assessment schedules.

5.4.3 Progression-Free Survival (2): Reviewer's Evaluations and Exploratory Analyses

This section consists of three parts: evaluation of assessment schedules, evaluation of missing data, and exploratory analyses of interval-censored data.

5.4.3.1 Evaluation of The Impact of Assessment Schedules

Although the sponsor's analysis results suggested a statistically significant difference in progression-free survival between the two treatment groups, it is not clear whether this is a true finding. This statistical reviewer's further explorations of data are included in this section and in the Appendix.

Lesions were to be assessed and measured in every other cycle starting at Cycle 3, at study completion, and every two-month follow-up until disease progression or death. The following table summarizes the number of patients who had a total of 1, 2, 3, 4, or more assessments, respectively, by the date of documented disease progression or death based on the sponsor's data. It is to be noted that cycles when only nontarget lesions were measured were ignored by the sponsor in determining the presence or absence of disease progression; hence, not considered an assessment in the following summary. A total of 632 patients had post baseline lesion assessments. It appeared that more than 50% of the patients (57.9% in the G3139 + DTIC group and 65.9% in the DTIC group) only had one assessment by the date of disease progression or death. About 77% in each group had up to two assessments, and 85% up to 3 assessments, by the date of documented disease progression or death.

Table 14: Summary of Number of Patients by Number of Lesion Assessments

# of Assessments	G3139 + DTIC [N= 321]	DTIC [N = 311]	Combined [N = 632]
1	186 (57.9%)	205 (65.9%)	391 (61.9%)
2	60 (18.7%)	39 (12.5%)	99 (15.7%)
3	24 (7.5%)	26 (8.4%)	50 (7.9%)
4	25 (7.8%)	22 (7.1%)	47 (7.4%)
5 or more	26 (8.1%)	19 (6.1%)	45 (7.1%)

Note: Assessments after the date of documented disease progression or death were excluded.

Reviewer's Table

Since lesions were measured periodically, disease progression generally did not occur on the date of evaluation but rather prior to this date. If the intervals between two consecutive assessments (termed assessment intervals) are longer, the documented date of disease progression would tend to be delayed; hence, the observed progression-free survival time would tend to be prolonged. Similarly, if the first assessment date is

delayed, the observed progression-free survival would also tend to be inappropriately prolonged even if the assessment intervals remain the same.

In this study, although assessment schedules were intended to be the same between the two treatment groups, because of the nature of treatment schedules (G3139 via 5-day and DTIC only 1-hour infusion), it was observed that the actual assessment schedule for patients in the G3139 + DTIC group appeared generally slightly delayed compared to that for patients in the DTIC group as summarized in Table 15. This table summarizes the times from the date of randomization to the 1st, 2nd and 3rd lesion assessments for those assessments conducted by the date of documented disease progression or death. The first 3 assessments were chosen because most patients had documented disease progression or death by the third assessment (see Table 14). The median times to each assessment appeared slightly longer in the G3139 + DTIC group than in the DTIC group (48 vs. 43, 94 vs. 87, and 137 vs. 129, respectively).

The Kaplan-Meier curves for time to these three assessments are respectively depicted in Figure 6. The curves represented the proportion that the intervals (times from randomization to the assessment) were longer than a given number of days. For example, in the first plot, the number 0.5 in the vertical axis corresponded to 48 days in the horizontal axis for the G3139 + DTIC curve. This may be interpreted as that in 50% of the patients the time from the date of randomization to the 1st assessment was longer than 48 days for those patients in the G3139 + DTIC group who had at least 1st assessment. Since the G3139 + DTIC curve mostly stayed above the DTIC curve in each plot, it appeared that patients in the G3139 + DTIC had generally delayed assessment schedule compared to those in the DTIC group for those who had assessments by the date of documented disease progression or death. The nominal p-values from an exploratory logrank test comparing the two distributions tended to be very small as seen in Table 15. Even a slight difference in assessment schedule between treatment groups could potentially bias the estimation of treatment effect and likely lead to a false positive inference in a large study.

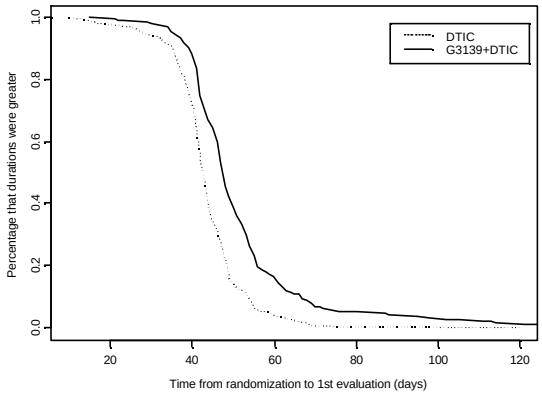
Table 15: Summary of Time from Randomization to the 1st, 2nd and 3rd Lesion Assessments[(1)] (Reviewer's Table)

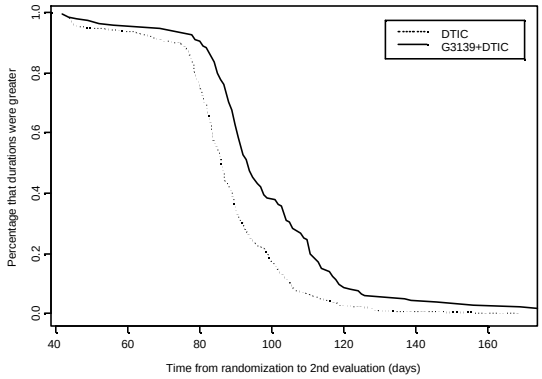
Time from randomization to the	Number of Patients		Median in days (95% C.I.)		**Logrank p-value[(2)]**
	G3139 + DTIC	DTIC	G3139 + DTIC	DTIC	
1st assessment	321	311	48 (47, 49)	43 (42, 44)	< 0.0001
2nd assessment	135	106	94 (92, 98)	87 (84, 89)	< 0.0001
3rd assessment	75	67	137 (134, 146)	129 (125, 133)	<0.0006

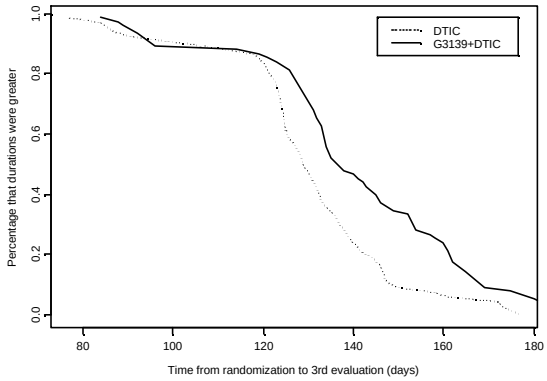
Notes: (1) Assessments after documented disease progression or death were excluded from the analysis.
(2) p-values were from logrank test comparing time to assessment between treatment groups. These reported p-values were nominal, not adjusting for multiplicity

Figure 6: Kaplan-Meier Curves of Time from Randomization to 1st, 2nd, and 3rd Lesion Assessments (Reviewer's Figures)

[**Note:** Assessments conducted after the date of documented disease progression or death were excluded.]







In order to investigate if the observed difference in time to lesion assessment between the two treatment groups was mainly affected by the difference in start day of the first treatment cycle, time to lesion assessment was recalculated from the date of treatment start (i.e., start of Cycle 1) instead of the date of randomization. The following table summarizes the result for the 1st, 2nd and 3rd lesion assessments that were conduced by the date of documented disease progression or death. The difference in median times reduced to 1 day for the first assessment and was about 5 days for the next two assessments. The Kaplan-Meier curves for time to these three assessments are respectively depicted in Figure 7. Again, the G3139 + DTIC curve mostly stayed above the DTIC curve in each plot. This suggested that even when the time was calculated from the treatment start date, patients in the G3139 + DTIC had generally delayed assessment timing compared to those in the DTIC group for those who had assessments by the date of documented disease progression or death. The nominal p-values were very small as seen Table 16.

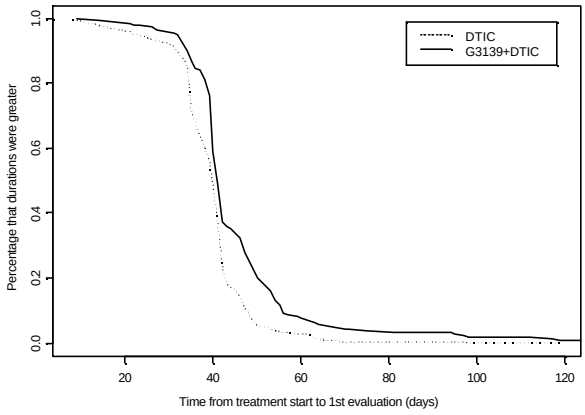
Table 16: Summary of Time from Treatment Start to the 1st, 2nd and 3rd Lesion Assessments[(1)] (Reviewer's Table)

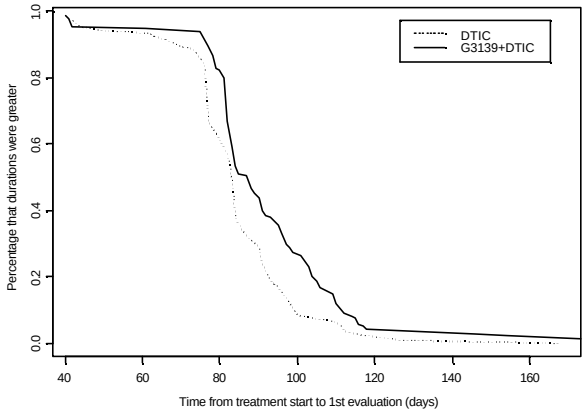
Time from treatment start to the	Number of Patients		Median in days (95% C.I.)		**Logrank p-value[(2)]**
	G3139+DTIC	DTIC	G3139+DTIC	DTIC	
1st assessment	321	311	41 (41, 42)	40 (40, 41)	< 0.0001
2nd assessment	135	106	88 (84, 91)	83.5 (82 84)	0.0009
3rd assessment	75	67	131 (127, 138)	126 (124, 130)	0.0065

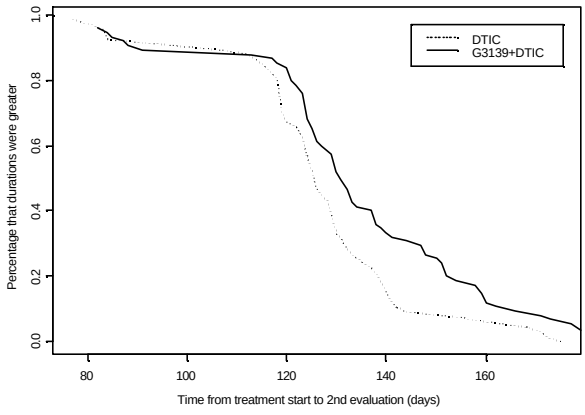
Notes: (1) Assessments after documented disease progression or death were excluded from the analysis.
(2) p-values were from logrank test comparing time to assessment between treatment groups. These reported p-values were nominal, not adjusting for multiplicity.

Figure 7: Kaplan-Meier Curves of Time from Treatment Start to 1st, 2nd, and 3rd Lesion Assessment (Reviewer's Figures)

[Note: Assessments after documented disease progression or death were excluded from the analysis.]







In order to investigate how lesion assessment schedule could influence the inference in comparing progression-free survival between treatment groups, simulation studies were performed in various assessment schedules **under the assumptions that the distributions of progression-free survival were equal between the two treatment groups**. Several scenarios were considered. In Scenario 1, patients in the control group were assessed every 6 weeks for up to 6 assessments, while in scenario 2, patients in the control group were assessed every 3 weeks for up to 12 assessments. Under each scenario, two different schedules for patients in the experimental group were applied: (a) patients in the experimental group were assessed 2 days later than those in the control group for each assessment; (b) the assessment interval for patients in the experimental group was 2 days longer than those in the control group. **The simulation results suggested that the chance of falsely inferring treatment difference in progression-free survival could be very large even for slightly different assessment schedules between the two treatment groups and the chance increased as the sample size increased and could reach to nearly 100%**.

Furthermore, in order to investigate how much bias could be introduced in estimating the difference in median progression-free survival between the two treatment groups, the statistical reviewer conducted further simulation studies where progression-free survival data was generated such that the actual difference in the medians between the two groups was only 4 days. It was, however, observed that the estimated difference in the medians could be more than 40 days under certain assessment schedules. For more details, please refer to the Appendix.

5.4.3.2 Evaluation of The Impact of Missing Data

The statistical reviewer noted from the sponsor's SAS program and the response to the Agency's 02 February 2004 fax that in determining the presence or absence of disease progression, only assessments in cycles when target lesions were measured was considered. This means that in a given cycle if nontarget lesions were measured but target lesions were not, then even if disease progression was present in nontarget lesions, the data were ignored. Based on this algorithm, the sponsor summarized the number of patients who had missing target lesions by the date of documented disease progression or death as in Table 12 of this review.

This reviewer summarized the information not only for target lesions but also for nontarget lesions as in Table 17 and Table 18. In these two tables, patients who had disease progression or death were considered and, to be consistent with the sponsor's algorithm, only assessments where there were target lesion measurements were considered. The number of patients who had missing lesion measurements **prior to** documented disease progression or death was tabulated in two ways in Table 17: (a) missing at least one lesion measurement, and (b) missing more than 50% of lesion measurements. Analogous to this, the number of patients who had missing lesion

measurements **at the same time as** the documented disease progression or death was also tabulated in two ways in Table 18. This reviewer's results for target lesions were similar to the sponsor's in that not many patients had missing measurements by the time of documented disease progression or death.

It was noted, however, that there were many patients with missing nontarget lesion measurements before disease progression or death was documented: 90 such patients from the G3139 + DTIC group and 61 from the DTIC group. This accounted for 34.2% (90/263) and 23.5% (61/260) of patients who had documented disease progression or death for the two treatment groups, respectively. Eighty-seven of the 90 patients from the G3139 + DTIC group and all of the 61 patients from the DTIC group had missing nontarget lesion measurements prior to documented disease progression or death. Of these patients who had missing nontarget lesion measurements, 43 (of 90) and 32 (of 61) patients respectively in the G3139 + DTIC group and the DTIC group had more than 50% of missing nontarget lesion measurements.

While the above two tables focus on patients who had disease progression or death, Table 19 focuses on those who did not have documented disease progression or death; i.e., those who were censored. Among those who were censored and had baseline and post baseline measurements, 12.1% [70.7%] of the patients from the G3139 + DTIC group and 13.7% [74.5%] of the patients from the DTIC group had missing target [nontarget] lesion measurements before they were censored. Also, 5.2% [32.8%] of those patients from the G3139 + DTIC group and 5.9% [35.3%] of those from the DTIC group had more than 50% of missing nontarget lesion measurements.

It should be noted that, in these three tables, only cycles when there were target lesion assessments were considered. This reviewer noted that for some patients there were cycles when nontarget lesions were assessed but target lesions were not assessed at all prior to the documented disease progression or death. There were 28 such patients from the G3139 + DTIC groups and 25 from the DTIC group. A total of 14 (12 from the G3139 + DTIC and 2 from the DTIC group) patients would have been considered progressed had these assessments also been considered. When there is a lot of missing data as in this endpoint measurement, a great amount of bias could be introduced in estimating the treatment effect and, thus, a bias in drawing statistical inference.

Table 17: Summary of Number of Patients with Missing Lesion Measurements <u>Prior to</u> Documented Disease Progression or Death Among Those Who Had Disease Progression or Death (Reviewer's Table)

Lesions	Missing **at Least One** Lesion Measurement		Missing **More Than 50%** of Lesion Measurements	
	G3139 + DTIC [N = 263][†]	DTIC [N = 260][†]	G3139 + DTIC [N = 263][†]	DTIC [N = 260][†]
Target	16 (6.0%)	13 (5.0%)	6 (2.3%)	2 (0.8%)
Non-target	87 (33.1%)	61 (23.5%)	43 (16.3%)	32 (12.3%)
Target or nontarget	90 (34.2%)	61 (23.5%)	45[‡] (17.1%)	32[‡] (12.3%)

[†] Number of patients who had disease progression or death in each treatment group.

[‡] Number of patients who had more than 50% of missing target lesion measurements or more than 50% of missing nontarget lesion measurements.

Table 18: Summary of Number of Patients with Missing Lesion Measurements <u>at The Same Time</u> of Documented Disease Progression or Death Among Those Who Had Disease Progression or Death (Reviewer's Table)

Lesions	Missing **at Least One** Lesion Measurement		Missing **More Than 50%** of Lesion Measurements	
	G3139 + DTIC [N = 263][†]	DTIC [N = 260][†]	G3139 + DTIC [N = 263][†]	DTIC [N = 260][†]
Target	30 (11.4%)	24 (9.2%)	6 (2.3%)	5 (1.9%)
Non-target	140 (53.2%)	145 (55.8%)	56 (21.3%)	50 (19.2%)
Target or nontarget	142 (54.0%)	148 (56.9%)	58[‡] (22.1%)	54[‡] (20.8%)

[†] Number of patients who had disease progression or death in each treatment group.

[‡] Number of patients who had more than 50% of missing target lesion measurements or more than 50% of missing nontarget lesion measurements.

Table 19: Summary of Number of Patients with Missing Lesion Measurements Among Those Who Were Censored and Had Baseline and Post Baseline Lesion Measurements (Reviewer's Table)

Lesions	Missing **at Least One** Lesion Measurement		Missing **More Than 50%** of Lesion Measurements	
	G3139 + DTIC [N = 58]†	DTIC [N = 51]†	G3139 + DTIC [N = 58]†	DTIC [N = 51]†
Target	7 (12.1%)	7 (13.7%)	3 (5.2%)	3 (5.9%)
Non-target	41 (70.7%)	38 (74.5%)	19 (32.8%)	18 (35.3%)
Target or nontarget	42 (72.4%)	40 (78.4%)	20‡ (34.5%)	18‡ (35.3%)

† Number of censored patients who had baseline and post baseline lesion measurements in each treatment group.
‡ Number of patients who had more than 50% of missing target lesion measurements or more than 50% of missing nontarget lesion measurements.

5.4.3.3 Exploratory Analyses of Interval-Censored Data

Since disease progression was assessed periodically, it may not be appropriate to treat data as right-censored data as in overall survival data, especially when there was difference in assessment schedules and there were considerable missing data. The type of data for progression-free survival can be better presented in the form of interval-censored data in that if disease progression is observed in a given assessment, the actual date of occurrence could be on any day between the previous assessment and this assessment. If disease progression is not observed by the end of the study, then the interval is from the last assessment date to an infinite date – corresponding to the right censored data. Therefore, for analysis of interval-censored data it is required to obtain the endpoints of the interval for each patient. This reviewer performed the generalized logrank test by taking these into account. Four approaches were used by this reviewer to obtain the interval for each patient. All these approaches were based on the sponsor's documented date of disease progression or death. These approaches are summarized as below:

- **Approach 1:** Only the cycles when <u>at least one</u> target lesions was measured were considered. The interval for each patient was purely from the previous assessment date to the assessment date when disease progression was observed regardless of missing lesion measurements. To be more specific, this approach did not take into account of missing lesion measurements, but took into account timing of assessments.
- **Approach 2:** Only the cycles when <u>all</u> target lesions were measured were considered. This approach did not take into account missing nontarget lesion measurements, but it did exclude assessment when there was at least one missing target lesion measurement.

- **Approach 3:** Only cycles when at most 50% of target and 50% of nontarget lesion measurements were missing were considered, i.e., this approach excluded assessments when there were too many missing lesion measurements.
- **Approach 4:** Only cycles when there were no missing target or nontarget lesion measurements were considered.

These four approaches are illustrated using the information in the following table. Suppose that a patient had documented disease progression or death in Visit 5 (Day 210) based on the sponsor's primary methodology of determining the presence or absence of disease progression or death. The interval where disease could possibly progress for this patient would be (168, 210] using Approach 1, (126, 210] using Approach 2, (168, 210] using Approach 3, and (42, 210] using Approach 4.

Table 20: Illustration of Approaches to Interval-Censoring (Reviewer's Table)

Post baseline Assessment	Assessment Day Since Randomization	% of missing target lesion measurements	% missing nontarget lesion measurements
1st	Day 42	0%	0%
2nd	Day 84	10%	40%
3rd	Day 126	0%	60%
4th	Day 168	30%	40%
5th	Day 210	60%	20%

The results[1] as summarized in the following table suggest that after taking into account the uncertainty of missing values and actual date of disease progression, the magnitude of statistical significance diminishes. Especially, when taking into account missing nontarget lesion measurements (Approaches 3 and 4) there is no statistically significant difference in progression-free survival between treatment groups at the significance level of 0.05.

[1] The algorithm for analysis of interval-censored data was based on Fay, M. P. provided on http://www.biostat.wustl.edu/archives/html/s-news/1999-04/msg00026.html.

Table 21: Results of Generalized Logrank Test with Interval-Censored Data (Reviewer's Table)

	Estimated Median (days)		P value from two-sided generalized logrank test
Approach	G3139 + DTIC	DTIC	
1	57	38	0.012
2	35	30	0.030
3	17	14	0.119
4	18	12	0.119

Note: P-value was from generalized logrank test comparing the entire curves between the two treatment groups.

5.4.4 Antitumor Response Rate: Sponsor's Analysis and Reviewer's Comments

If a response was confirmed by a repeat measurement performed no less than 3 weeks after the response was first noted, it was called a confirmed response. The sponsor's study report states that responses used for analysis purposes were calculated by computerized algorithm of lesion measurements based on RECIST criteria including confirmation measurements at least 3 weeks after the initial response. The following table is the sponsor's summary result of confirmed antitumor responses by treatment group provided in the NDA document. Based on the sponsor's analysis, a total of 71 patients had confirmed complete or partial response and the response rate was 11.7% in the G3139+DTIC group and 6.8% in the DTIC group. The point estimate of the odds ratios was 1.82. This means that the odds for a patient in the G3139 + DTIC group to have a confirmed (complete or partial) response was estimated to be 1.82 times the odds for a patient in the DTIC group. The sponsor also performed several other sensitivity or exploratory analyses and all p-values were less than 0.05, hence, concluded that there was a significantly greater antitumor response rate in the G3319 + DTIC group than in the DTIC group.

Table 22: Sponsor-Confirmed Antitumor Response Rates during The Treatment Phase

[*Source: Sponsor's Table 22 in Section 11.2.2.2*]

G3139 + DTIC (N = 386) n (%)[a]		DTIC (N = 385) n (%)[a]		Odds ratio[b] (95% CI)	Chi-square P value
Complete Response	**Partial response**	**Complete response**	**Partial response**		
3 (0.8)	42 (10.9)	2 (0.5)	24 (6.2)		
45 (11.7)		26 (6.8)		1.82 (1.10-3.02)	0.019

Cross-reference: Table 14.2.13
CI = confidence interval
[a] Percentages were calculated by using N, the total number of subjects in the group, as the denominator.
[b] 95% confidence interval based on the asymptotic method

REVIEWER COMMENTS:

[1] The review by RadPharm, an independent radiological review, showed that 11 of the 71 confirmed responders identified by the sponsor were not assessable and an additional 20 of the 71 could not be confirmed. RadPharm was able to verify a total of 40 of the sponsor's 71 responders: 26 (complete or partial) responders from the G3139 + DTIC group and 14 responders from the DTIC group. The following table summarizes the analysis result based on the RadPharm data. The response rate for the combination arm was 6.7% versus 3.6% for DTIC alone; the absolute difference in response rates was 3.10% and the point estimate of the odds ratio was 1.91. The p-values were close to the significance level of 0.05.

Table 23: RadPharm-Confirmed Antitumor Response Rates during The Treatment Phase (Reviewer's Table)

	G3139 + DTIC (N = 386)	DTIC (N = 385)	P value
# of responders (%)	26 (6.74%)	14 (3.64%)	
Odds ratio (95% CI)	1.91	(0.98, 3.72)	0.056
Difference in proportions (95% CI)	3.10%	(-0.03%, 6.43%)	0.52

Note: Based on the asymptotic method. Results based on the exact method were very similar to these. P-values not adjusted for multiplicity.

[2] Also, the sponsor's subgroup analysis by countries indicated that the statistical significance was mainly driven by the non-US sites (see Table 32).

5.4.5 Duration of Antitumor Response: Sponsor's Analysis and Reviewer's Comments

The following table is the sponsor's summary of duration of response for responders. The median duration of response was estimated to be 126 days for the G3139 + DTIC group and 128 days from the DTIC group.

Table 23: Duration of Sponsor-Confirmed Antitumor Response in Responding Subjects by Treatment Group: Intent-to-Treat Population
[*Source: Sponsor's Table 23 in Section 11.2.2.3*]

Days	**G3139 + DTIC (N = 386)**	**DTIC (N = 385)**
n[a]	45	26
Mean (SD)	175.7 (136.0)	153.5 (99.2)
Median	126	128
25^{th}, 75^{th} percentiles	86, 238	84, 142
Minimum, maximum	41, 565	42, 390

Cross-reference: Table 14.2.15
SD = standard deviation
[a] Calculated only for those subjects who responded (ie, those who had a complete response or a partial response) by RECIST criteria

REVIEWER COMMENTS:

[1] This reviewer noticed from the sponsor's SAS program that duration of response was calculated based on the assessment dates of target lesion measurements only. The estimated median durations of response were similar.

5.4.6 Findings in Special/Subgroup Populations

5.4.6.1 Gender, Race and Age

This section summarizes the sponsor's exploratory subgroup analyses by gender and age for overall survival, progression-free survival and antitumor response rate. Subgroup analyses by race were not performed because 746 of the 771 patients were Caucasians.

No unusual patterns were observed in the sponsor's results of subgroup analyses except that the strength of the G3139 + DTIC treatment effect relative to the DTIC alone treatment in antitumor response rate appeared more apparent in the male subgroup than in the female group and in the subgroup of below 65 than in the subgroup of above 65 years old (Table 29).

Table 24: Summary of Survival Analysis by Gender

[*Source: Sponsor's Table 14.2.35*]

	G3139 + DTIC (N=386)	DTIC Alone (N=385)	Hazard Ratio (95% CI)	Logrank P value
Male				
Proportion (%) of Subjects who Died	169/236 (71.6%)	184/253 (72.7%)	0.894 (0.73 – 1.10)	0.293
Median Survival Time (95%% CI)	243 (200 - 303)	226 (206 - 268)		
Female				
Proportion (%) of Subjects who Died	97/150 (64.7%)	85/132 (64.4%)	0.914 (0.68 - 1.22)	0.549
Median Survival Time (95%% CI)	306 (255 - 355)	268 (227 - 317)		

Table 25: Summary of Survival Analysis by Age

[*Source: Sponsor's Table 14.2.30*]

	G3139 + DTIC (N=386)	DTIC Alone (N=385)	Hazard Ratio (95% CI)	Logrank P value
Age < 65				
Proportion (%) of Subjects who Died	164/239 (68.6%)	167/241 (69.3%)	0.872 (0.70 - 1.08)	0.2135
Median Survival Time (95%% CI)	243 (211 - 322)	229 (213 - 274)		
Age ³ 65				
Proportion (%) of Subjects who Died	102/147 (69.4%)	102/144 (70.8%)	0.926 (0.70 - 1.22)–	0.5815
Median Survival Time (95%% CI)	281 (243 - 333)	248 (221 - 313)		

Table 26: Summary of Progression-Free Survival Analysis by Gender

[*Source: Sponsor's Table 14.2.36*]

	G3139 + DTIC [N=386]	DTIC Alone [N=385]	Hazard Ratio (95% CI)	Logrank P value
Male				
Number (%) of Patients with Progression of Disease	162/236 (68.6%)	169/253 (66.8%)	0.733 (0.59 – 0.91)	0.0042
Median Time (95%% CI)	69 (56 - 91)	49 (46 - 55)		
Female				
Number (%) of Patients with Progression of Disease	101/150 (67.3%)	91/132 (68.9%)	0.741 (0.56 – 0.98)	0.0364
Median Time (95%% CI)	87 (61 - 98)	53 (48 - 62)		

Table 27: Summary of Progression-Free Survival Analysis by Age

[*Source: Sponsor's Table 14.2.31*]

	G3139 + DTIC [N=386]	DTIC Alone [N=385]	Hazard Ratio (95% CI)	Logrank P value
Age < 65				
Number (%) of Patients with Progression of Disease	182/239 (76.2%)	159/241 (66.0%)	0.785 (0.63 - 0.97)	0.0257
Median Time (95%% CI)	64 (56 - 88)	50 (48 - 59)		
Age ³ 65				
Number (%) of Patients with Progression of Disease	81/147 (55.1%)	101/144 (70.1%)	0.617 (0.46 - 0.83)	0.0010
Median Time (95%% CI)	90 (65 - 105)	49 (46 - 56)		

Table 28: Summary of Sponsor-Confirmed Antitumor Response Rate by Gender

[*Source: Sponsor's Table 14.2.38*]

	G3139 + DTIC [N=386]	DTIC Alone [N=385]	Odds Ratio (95% CI)	Asymptotic chi-square P value
Male				
Confirmed CR/PR	27/236 (11.4%)	13/253 (5.1%)	2.385 (1.200 - 4.741)	0.0111
Female				
Confirmed CR/PR	18/150 (12.0%)	13/132 (9.8%)	1.248 (0.587 - 2.656)	0.5644

Table 29: Summary of Sponsor-Confirmed Antitumor Response Rate by Age

[*Source: Sponsor's Table 14.2.33*]

	G3139 + DTIC [N=386]	DTIC Alone [N=385]	Odds Ratio (95% CI)	Asymptotic chi-square P value
Age < 65				
Confirmed CR/PR	33/239 (13.8%)	16/241 (6.6%)	2.253 (1.204 - 4.214)	0.0095
Age ³ 65				
Confirmed CR/PR	12/147 (8.2%)	10/144 (6.9%)	1.191 (0.498 - 2.850)	0.6942

5.4.6.2 Other Special/Subgroup Populations

The sponsor also performed several other subgroup analyses. Of particular interest was subgroup analyses by country as extracted in this section. It was noted that the Australian data appeared statistically in favor of the G3139 + DTIC group in all three efficacy endpoints despite a small sample size. Even for the overall survival, the p-value from the Australia cohort was extremely small. More than half of the study patients came from the U.S., in which there was no statistically significant difference. A similar pattern was also observed in antitumor response rate with inconsistencies among countries. The results of subgroup analysis by country suggested that the strength of the treatment effect of G3139 + DTIC was driven by the non-U.S. sites, in particular Australia, and was not consistent among countries.

Table 30: Summary of Survival Analysis by Country

[*Source: Sponsor's Table 14.2.40*]

	G3139 + DTIC (N=386)	DTIC Alone (N=385)	Hazard Ratio (95% CI)	Logrank P value
US				
Proportion (%) of Subjects who Died	155/213 (72.8%)	150/217 (69.1%)	1.033 (0.82 - 1.29)	0.7782
Median Survival Time (95%% CI)	281 (215 – 333)	270 (235 - 315)		
Non-US				
Proportion (%) of Subjects who Died	111/173 (64.2%)	119/168 (70.8%)	0.726 (0.56 - 0.94)	0.0155
Median Survival Time (95%% CI)	262 (220 – 326)	123 (191 - 250)		
Australia				
Proportion (%) of Subjects who Died	27/ 42 (64.3%)	34/ 42 (81.0%)	0.483 (0.28 - 0.83)	0.0072
Median Survival Time (95%% CI)	351 (148 – 484)	208 (150 - 268)		

Table 31: Summary of Progression-Free Survival Analysis by Country

[*Source: Sponsor's Table 14.2.41*]

	G3139 + DTIC [N=386]	DTIC Alone [N=385]	Hazard Ratio (95% CI)	Logrank P value
US				
Number (%) of Patients with Progression of Disease	138/213 (64.8%)	145/217 (66.8%)	0.792 (0.63 - 1.00)	0.0467
Median Time (95%% CI)	64 (55 – 86)	49 (46 - 55)		
Non-US				
Number (%) of Patients with Progression of Disease	125/173 (72.3%)	115/168 (68.5%)	0.639 (0.49 - 0.83)	0.006
Median Time (95%% CI)	89 (64 – 105)	51 (48 - 60)		
Australia				
Number (%) of Patients with Progression of Disease	32/ 42 (76.2%)	33/ 42 (78.6%)	0.452 (0.27 - 0.76)	0.0018
Median Time (95%% CI)	94 (49 – 187)	51 (42 - 84)		

Table 32: Summary of Sponsor-Confirmed Antitumor Response Rate by Country

[*Source: Sponsor's Table 14.2.43*]

	G3139 + DTIC [N=386]	DTIC Alone [N=385]	Odds Ratio (95% CI)	Asymptotic chi-square P value
US				
Confirmed CR/PR	16/213 (7.5%)	12/217 (5.5%)	1.387 (0.640 - 3.007)	0.4050
Non-US				
Confirmed CR/PR	29/173 (16.8%)	14/168 (8.3%)	2.215 (1.126 - 4.360)	0.0191
Australia				
Confirmed CR/PR	10/ 42 (23.8%)	3/ 42 (7.1%)	4.063 (1.030 -16.024)	0.0347

5.4.7 Exploratory FDA Analysis of Performance Status

Patients were required to have ECOG performance status of 0, 1, or 2 for study eligibility. The majority of the patients were asymptomatic and were performance status (P.S.) = 0 at study entry. Details of this analysis will be presented at the ODAC meeting.

5.5 SAFETY

5.5.1 Dosing summary for GM301

For the DTIC alone arm (column on the right), the total median dose of DTIC delivered was 2009 mg/m2. The center two columns show the dosing in the combination arm. For the combination arm, the total median dose of DTIC was 2055 mg/m2 and the total median dose of G3139 was 77 mg/kg.

Table 33 Dosing During the Study

		G3139 + DTIC Combination Arm		**DTIC Alone**
		G3139 (mg/kg)	DTIC (mg/m2)	DTIC (mg/m2)
Cumulative doses received on study				
	n	371	365	360
	Mean (SD)	126.1 (85)	3417.9 (2191)	3371.9 (2301)
	Median	**76.9**	**2054.9**	**2009.1**
	IQ Range	(68, 177)	(1944, 4795)	(1957, 4862)
	Range	(9, 608)	(873, 9000)	(901, 9845)

Sponsor's table 14.1.20

Reviewer comments:
Since the G3139 dose per cycle was 7 mg/kg/day for 5 days, one cycle of therapy equals 35 mg/kg per patient. The DTIC dosing was similar in both arms. The median dose received indicates that the median number of cycles on each arm was also equal in the 2 arms. Dose delays were not common, occurring in 13% of the DTIC arm and in 16% of the combination arm. Each cycle of therapy was 3 weeks – 21 days. Following cycle 2 of treatment, 50% of patients receiving G3139+DTIC and 60% of patients receiving DTIC alone had discontinued treatment, primarily due to disease progression. Only 5 patients in the combination arm and 2 in the DTIC arm completed the planned 8 cycles. The mean number of days on therapy for the combination was 72 versus 57 for DTIC alone.

5.5.2 AEs AND SAEs

Table 34 Overview of adverse events by treatment group: Safety Population

Category	G3139 plus DTIC	DTIC
Number of patients randomized	**386**	**385**
Number of patients receiving any therapy	**371**	**360**
	n (%)	**n (%)**
At least 1 adverse event	368 (95.3)	346 (89.9)
At least 1 treatment-emergent adverse event	365 (98.4)	342 (95.0)
At least 1 Grade 3 or 4 treatment-emergent adverse event	249 (67.1)	154 (42.8)
At least 1 serious treatment-emergent adverse event	149 (40.2)	97 (26.9)
At least 1 treatment-emergent adverse event with	69 (18.6)	39 (10.8)
action taken - discontinued permanently (see below)		
Death while on protocol therapy or within 30 days	29 (7.8)	25 (6.9)
from date of last dose of study drug		
At least 1 treatment-emergent adverse event with	32 (8.6)	33 (9.2)
outcome of death		

Sponsor's table 24, clinical study report

Reviewer comment:
The frequency of grade 3-4 AEs, serious AEs, and TEAEs leading to discontinuation all were higher on the Genasense arm.

Table 35 Adverse Events In ³ 20% of Patients

Preferred term [a]	G3139 plus DTIC (N = 371) n (%)	DTIC (N = 360) n (%)	Total (N = 731) n (%)
Nausea	231 (62.3)	169 (46.9)	400 (54.7)
Fatigue	170 (45.8)	142 (39.4)	312 (42.7)
Pyrexia	197 (53.1)	63 (17.5)	260 (35.6)
Vomiting	139 (37.5)	75 (20.8)	214 (29.3)
Constipation	103 (27.8)	93 (25.8)	196 (26.8)
Anorexia	114 (30.7)	56 (15.6)	170 (23.3)
Diarrhea	100 (27.0)	63 (17.5)	163 (22.3)
Neutropenia	103 (27.8)	56 (15.6)	159 (21.8)
Thrombocytopenia	107 (28.8)	40 (11.1)	147 (20.1)
Anemia	86 (23.2)	61 (16.9)	147 (20.1)
Headache	97 (26.1)	47 (13.1)	144 (19.7)

Note: Results are presented for the Safety Population.
[a] If a subject had more than 1 occurrence of a specific event, that subject was counted only once for that preferred term.
Sponsor's summary table from supplemental submission received 2-24-04

Reviewer comment: *All toxicities were more frequent on the combination arm. Pyrexia was three times as frequent. Thrombocytopenia increased from 11% to almost 29%. Neutropenia and anorexia AEs were twice as frequent with G3139.*

Table 36 Reviewer's summary compilation of Treatment Emergent AEs (TEAEs) of clinical interest from the sponsor's tabulation Table 14.3.1.5: Safety population

System-Organ Class Preferred term: Number and percent of patients by each treatment arm:	**G3139 plus DTIC (N = 371) n (%)**	**DTIC (N = 360) n (%)**
Blood and Lymphatic	171 (46%)	113 (31%)
Neutropenia	103 (28%)	56 (16%)
Thrombocytopenia	107 (29%)	40 (11%)
Cardiac Disorders	43 (12%)	24 (7%)
GI Disorders	294 (79%)	257 (71%)
Nausea	231 (62%)	169 (47%)
Vomiting	139 (38%)	75 (21%)
General and Administration Site		
Fatigue	170 (46%)	142 (39%)
Pyrexia	197 (53%)	63 (18%)
Rigors	76 (21%)	16 (4%)
Influenza-like Illness	19 (5%)	3 (0.8%)
Injection-site Reactions	21 (6%)	5 (1%)
Drug Hypersensitivity NOS	8 (2%)	0 (0%)
Infections and Infestations	123 (33%)	65 (18%)
Bacteremia	8 (2%)	2 (0.6%)
Injection site infection	24 (7%)	4 (1%)
Prolonged aPTT	8 (2%)	0 (0%)
Anorexia	114 (31%)	56 (16%)
Headache	97 (26%)	47 (13%)
Depression	24 (7%)	13 (4%)
Alopecia	14 (4%)	8 (2%)
Pruritus and Rash	57 (15%)	18 (5%)
Upper extremity thrombosis *	18 (5%)	3 (0.8%)

* Axillary vein, Subclavian vein, Jugular vein, injection site thrombosis, and superior vena caval syndrome

Reviewer's table

Table 37 Serious treatment-emergent adverse events (SAEs) occurring in = 2% of subjects by preferred term and treatment group: Safety Population

SAEs Preferred term	G3139 plus DTIC (N = 371) n (%)[c]	DTIC (N = 360) n (%)[c]
Number with at least one SAE	149 (40%)	97 (27%)
Pyrexia	22 (5.9)	11 (3.1)
Thrombocytopenia	15 (4.0)	4 (1.1)
Nausea	14 (3.8)	5 (1.4)
Vomiting	13 (3.5)	6 (1.7)
Pulmonary embolism	11 (3.0)	3 (0.8)
Injection site infection	10 (2.7)	0 (0.0)
Anemia	9 (2.4)	8 (2.2)
Neutropenia	8 (2.2)	1 (0.3)
Dehydration	8 (2.2)	4 (1.1)
Deep venous thrombosis NOS	8 (2.2)	4 (1.1)

Sponsor's table 36, CSR page 145

Reviewer comments on safety:

There is increased toxicity with the combination of Genasense and DTIC. A total of 149 (40.2%) patients in the G3139 plus DTIC arm and 97 (26.9%) patients in the DTIC group had at least 1 serious AE (SAE). In the G3139 plus DTIC group, the most frequently reported SAEs included pyrexia (5.9%), thrombocytopenia, nausea, vomiting, pulmonary embolism (3%), and injection site infection (2.7%). Regarding all adverse events, the 2 treatment-emergent adverse events most clearly associated with G3139 were pyrexia (53%) and catheter-related events (infection (7%), upper extremity thrombosis (5%), and device malfunctions; the latter were due to the requirement for an indwelling central catheter for the 5-day continuous intravenous administration of G3139.

In the DTIC group, the most frequently reported serious treatment emergent adverse event also was pyrexia (3.1 %). Other types of treatment-emergent adverse events were similar in the 2 treatment groups but were more frequent and more severe in the combination arm.

6 SUMMARY AND CONCLUSIONS

6.1 STATISTICAL ISSUES AND COLLECTIVE EVIDENCE

The study GM301 was an open label study. It was designed to demonstrate a survival improvement for the use of G3139 in combination of DTIC in patients with advanced and metastatic malignant melanoma.. A very large phase 3 study, of 771 patients with metastatic melanoma was conducted. Sample size calculation was based on the primary endpoint of overall survival. Progression-free survival and antitumor response rate were secondary endpoints. The final analysis for the study was to be performed after a total of 508 deaths were observed. The sponsor performed an analysis of this rolling NDA when 537 deaths had occurred and is seeking marketing approval. The result of the final (and only) survival analysis did not show statistical significance in favor of the G3139 + DTIC treatment group (two-sided logrank p-value = 0.18, point estimate of the hazard ratio 0.89). However, based on the sponsor's analyses of progression-free survival and antitumor response rate, the p-values were all very small. Hence, the sponsor claims efficacy based on apparent statistically significant differences in progression-free survival and antitumor response rate.

There were several clinical and statistical issues involved in this study as listed below.

[1] Study G301 failed to provide sufficient evidence of efficacy in support of approval based on the primary endpoint overall survival (P-value = 0.18, Table 9). In addition, subgroup analysis of survival by country suggested that in this study the effect of the G3139 + DTIC treatment compared to DTIC alone was mainly driven by the non-US countries (Table 30). The sample size for this study was increased two times to ensure an adequately powered study for detecting a meaningful difference in overall survival.

[2] Efficacy results based on secondary endpoints such as progression-free survival and antitumor response rate were only to be supportive or exploratory, provided that there was a statistically significant finding in overall survival. From a statistical perspective, all the allocated type-I error rate was spent in conducting the primary analysis of overall survival.

[3] More importantly, there were several major issues in assessing the response status and date of progression-free survival, which could introduce bias in estimating treatment effects. These issues are summarized in the following bullets. For more details, please refer to Section 5.4.3.

- There appeared to be a difference in actual lesion assessment schedules between the two treatment groups. Simulation studies illustrated that even a subtle difference in assessment schedules could very well lead to a false positive inference in a substantially large study (Appendix) and could also lead to an

overestimated difference in median progression-free survival between treatment groups.

- Various sources of missing data were observed. Although missing target lesion measurements were observed in only some patients (6% of progressors in the G3139 + DTIC group vs. 5% of progressors in DTIC group), missing nontarget lesion measurements were observed in many patients before disease progression or death (34% of progressors in the G3139 + DTIC group vs. 23% of progressors in DTIC group). In addition, in some patients there were some cycles when no target lesions were measured, but nontarget lesions were measured. These cycles were ignored by the sponsor when determining presence or absence of disease progression. These various problems of missing data could introduce a bias in estimating the treatment effect with respect to progression-free survival and antitumor response rate.

[4] Since status of disease progression was assessed periodically, progression-free survival data could be better presented in the form of interval-censored data. This reviewer performed the generalized logrank test on interval-censored data. These exploratory analyses suggested that the highly statistical significance presented by the sponsor diminished after taking into account the uncertainty of missing values and different assessment schedules.

[5] Although there was statistical significance in antitumor response rate, subgroup analysis by country suggested that the amount of statistical evidence was mainly driven by the non-US countries (Section 5.4.6.2). Also, there was considerable discordance between the sponsor's computer program determined response and the RadPharm independent review.

6.2 CONCLUSIONS AND RECOMMENDATIONS

The single randomized open-label study presented in this application failed to demonstrate overall survival benefit of G3139 + DTIC over DTIC alone. Any claims of improved efficacy based on secondary endpoints, progression-free survival and antitumor response rate are questionable because of the open-label nature of study, missing data, and differences in assessment interval between the two treatment groups. The findings could be falsely positive, especially in view of the lack of confirmation by a second well-controlled and well-conducted trial.

APPENDIX

6.3 SIMULATION RESULTS OF DIFFERENT LESION ASSESSMENT SCHEDULES

In order to investigate how lesion assessment schedule could influence the inference in comparing progression-free survival between treatment groups, simulations using SPlus software were performed in various assessment schedules under the assumption that the distributions of progression-free survival was equal between the two treatment groups. Several scenarios were considered. In Scenario 1, patients in the control group were assessed every 6 weeks for up to 6 assessments, while in scenario 2, patients in the control group were assessed every 3 weeks for up to 12 assessments. Under each scenario, two different schedules for patients in the experimental group were applied: (a) patients in the experimental group were assessed 2 days later than those in the control group for each assessment; (b) the assessment interval for patients in the experimental group was 2 days longer than those in the control group. This resulted in a total of 4 different configurations of assessment schedules as listed in the following table:

Table 38: Assessment Schedules in Simulation Studies (Reviewer's Table)

Configuration	Control Group	Experimental Group
1A	Days 42, 84, 126, 168, 210, 252	(delayed by 2 days) Days 44, 86, 128, 170, 212, 254
1B	Same as above	(assessment interval 2 days longer) Days 44, 88, 132, 176, 220, 264
2A	Days 21, 42, 63, 84, 105, 126, 147, 168, 189, 210, 231, 252	(delayed by 2 days) Days 23, 44, 65, 86, 107, 128, 149, 170, 191, 212, 233, 254
2B	Same as above	(assessment interval 2 days longer) Days 23, 46, 69, 92, 115, 138, 161, 184, 207, 230, 252, 275

In all configurations, progression-free survival data was generated from *an exponential distribution with a median of 50 days. Sample size* for each treatment group was either *100 or 300*. For each configuration, *5000 replications* were performed. In each replication, p-value from the two-sided logrank test was obtained and whether the null hypothesis (equal progression-free-survival distributions) would be rejected or not was recorded. The proportion of replications where the null hypothesis was rejected was used to estimate the probability of inferring a statistically significant difference with respect to progression-free survival between treatment groups. It is to be noted that the probability

here was indeed the probability of **falsely** inferring difference in progression-free survival since data was generated under equal distributions. As observed in Table 39, the p-values and the probabilities of false inference generally depended on the sample size and the assessment schedule. When the assessment intervals were wider such as in Configuration 1A and 1B and/or the sample size increased, the p-values decreased and the probabilities of false inference increased. The probability of false inference could increase to nearly 1 when the sample size was 300. It is to be noted that the p-values (average of the 5000 replications) in fact should have been large since the two distributions were simulated as identical distributions. The results suggested that the chance of falsely inferring a difference in progression-free survival between treatment groups could be very high even for slightly different assessment schedules between treatment groups. It was also observed in this table that the Monte Carlo estimates of medians were generally larger than the actual median (50 days). As the assessment intervals became wider, the magnitude of the overestimation increased. Although in this table the difference in estimated medians did not appear too large, in some situations the difference in estimated medians could be very large. This is discussed in the next paragraph.

Table 39: Simulation Results under Equal Progression-Free Survival Distributions (Reviewer's Table)

(Monte Carlo Estimates Based on 5000 Replications)

Configuration	Sample size per treatment group	Median (days)		**Probability of false inference**[†]	Logrank P-value[‡]
		G3139 + DTIC	DTIC		
1A	100	86	84	0.65	0.100
	300	86	84	0.98	0.004
1B	100	88	84	0.60	0.114
	300	88	84	0.97	0.007
2A	100	65	63	0.08	0.457
	300	65	63	0.41	0.198
2B	100	69	63	0.10	0.427
	300	69	63	0.21	0.330

[†] The probability of false inference was estimated by the proportion of the 5000 replications where the null hypothesis was rejected. This represented the probability of falsely inferring a difference in progression-free survival between the two treatment groups.

[‡] This was the average of 5000 p-values. Each simulation produced a p-value. These p-values were from two-sided logrank test comparing progression-free survival between treatment groups.

In order to illustrate how serious the overestimation in difference between estimated medians could be, progression-free survival data was generated such that the medians were 45 days for the experimental group and 41 days for the control group. Assessment

schedules in Configurations 1A and 2A were revisited. Simulation results based on 5000 replications are summarized in Table Table 40. It was observed in this table that when the progression-free survival distributions were slightly different between treatment groups (only 4-day difference in actual medians), the difference in estimated medians could be overwhelmingly large. The power for rejecting the null hypothesis of equal progression-free survival distributions was large, and increased when the sample size increased. **This illustrated that a huge difference in estimated medians may not necessarily represent the true difference.**

Table 40: Simulation Results under Slightly Different Progression-Free Survival Distributions (Reviewer's Table)

(Monte Carlo Estimates Based on 5000 Replications)

Configuration	Sample size per treatment group	Median (days)		Power †	Logrank P-value‡
		G3139 + DTIC	DTIC		
1A	100	86	42	0.92	0.018
	300	86	42	≈1.00	$<10^{-4}$
2A	100	65	42	0.44	0.179
	300	65	42	0.88	0.027

† The power was estimated by the proportion of the 5000 replications where the null hypothesis was rejected. This represented the probability inferring a difference in progression-free survival between the two treatment groups.

‡ This was the average of 5000 p-values. Each simulation produced a p-value. These p-values were from two-sided logrank test comparing progression-free survival between treatment groups.

References

i Atkins MB, Lotze MT, Dutcher JP, Fisher RI, Weiss G, Margolin K, et al. High-dose recombinant interleukin 2 therapy for patients with metastatic melanoma: Analysis of 270 patients treated between 1985 and 1993. J Clin Oncol 1999;17:2105-2116.

ii Sun W, Schuchter LM. Metastatic Melanoma. Curr Treat Options Oncol 2001;2:193-204.

iii Balch CM, Soong SJ, Gershenwald JE, Thompson JF, Reintgen DS, Cascinelli N, et al. Prognostic factors analysis of 17,600 melanoma patients: Validation of the American Joint Committee on cancer melanoma staging system. J Clin Oncol 2001;19:3622-3634.

iv Flaherty, J Unger, PY Liu, K Albain, V Sondak Gender differences and independent predictors of survival in patients with metastatic malignant melanoma on SWOG trials ASCO 1996 vol 15: (Meeting abstract). #1344

v Manola J, Atkins M, Ibrahim J, Kirkwood J. Prognostic Factors in Metastatic Melanoma: A Pooled Analysis of Eastern Cooperative Oncology Group Trials. J Clin Oncol 2000;18:3782-3793.