

CENTER FOR DRUG EVALUATION AND RESEARCH

APPROVAL PACKAGE FOR:

APPLICATION NUMBER

BLA 125118/000

Statistical Review(s)



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

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/Serial Number: STN 125118/0
Drug Name: abatacept (Orencia)
Indication(s): Rheumatoid arthritis (RA)
Applicant: Bristol-Myers Squibb
Date(s): Received :12/17/2004, Action date:9/30/2005
Review Priority: Standard

Biometrics Division: Biologics Therapeutics Statistical Staff
Statistical Reviewer: Kyung Yul Lee, Ph.D.
Concurring Reviewers: Boguang Zhen, Ph.D.
Aloka Chakravarty, Ph.D.

Medical Division: DTBIMP
Clinical Team: Keith Hull, MD.
Project Manager: Erik Laughner

Keywords: Abatacept, Methotrexate (MTX), Rheumatoid Arthritis, Randomized trial

TABLE OF CONTENTS

1. EXECUTIVE SUMMARY	
1.1 Conclusions and Recommendations.....	7
1.2 Brief Overview of Clinical Studies.....	7
1.2.1 Study Design.....	7
1.2.2 Study Population.....	8
1.2.3 Dosage.....	8
1.2.4 Endpoints.....	9
1.3 Statistical Issues and Findings.....	9
1.3.1 Sponsor's Analysis.....	9
1.3.2 Reviewer's Analysis.....	13
1.3.3 Statistical Issues.....	15
2. INTRODUCTION	
2.1 Overview	15
2.1.1 History of Drug Development.....	15
2.1.2 Clinical Development Program.....	16
2.1.3 Objectives in Treatment of Rheumatoid Arthritis.....	17
2.2 Data Sources.....	17
3. STATISTICAL EVALUATION	
3.1 Evaluation of Efficacy	17
3.1.1 Study IM101100.....	17
3.1.1.1 Study Design.....	17
3.1.1.2 Primary and Secondary Efficacy Endpoints.....	18
3.1.1.3 Statistical Methods.....	18
3.1.1.4 Patient Disposition, Demographic and Baseline Characteristics.....	19
3.1.1.5 Results and Conclusions.....	21
3.1.2 Study IM101102.....	25
3.1.2.1 Study Design.....	25
3.1.2.2 Primary and Secondary Efficacy Endpoints.....	26
3.1.2.3 Statistical Methods.....	26
3.1.2.4 Patient Disposition, Demographic and Baseline Characteristics.....	28
3.1.2.5 Results and Conclusions.....	31
3.1.3 Study IM101029.....	37
3.1.3.1 Study Design.....	37
3.1.3.2 Primary and Secondary Efficacy Endpoints.....	37
3.1.3.3 Statistical Methods.....	38
3.1.3.4 Patient Disposition, Demographic and Baseline Characteristics.....	39
3.1.3.5 Results and Conclusions.....	43
3.1.4 Study IM101031.....	48
3.1.4.1 Study Design.....	48
3.1.4.2 Primary and Secondary Efficacy Endpoints.....	49
3.1.4.3 Statistical Methods.....	49
3.1.4.4 Patient Disposition, Demographic and Baseline Characteristics.....	49
3.1.4.5 Results and Conclusions.....	54

3.2 Evaluation of Safety.....56

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race and Age.....57

4.1.1 Study IM101102.....58

4.1.2 Study IM101029.....59

4.2 Other Special/Subgroup Populations60

4.2.1 Study IM101102.....61

4.2.2 Study IM101029.....63

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence64

5.2 Conclusions and Recommendations64

SIGNATURES/DISTRIBUTION LIST PAGE65

Appears This Way
On Original

LIST OF TABLES

Table 1. Number of Patients Randomized and Discontinued by Treatment.....	8
Table 2. Summary Efficacy Results of Improvement of Signs and Symptoms.....	10
Table 3. Summary of Efficacy Results of Improvement in Physical Function.....	11
Table 4. Summary of Efficacy Results of Change from Baseline Radiographic Scores.....	12
Table 5. Summary of Efficacy Results of Improvement in Health-related Quality of Life.....	13
Table 6. Summary of Efficacy Results of ACR20 Responses using GEE Methods.....	14
Table 7. Summary of Efficacy Results of HAQ_DI responses by GEE Methods.....	14
Table 8. Reasons for Discontinuation: Days 1- 360 (IM101100).....	19
Table 9. Baseline Demographic Characteristics (IM101100).....	20
Table 10. Baseline Clinical Rheumatoid Arthritis Characteristics (IM101100).....	21
Table 11. Baseline x-ray (Prespecified Genant-Modified Sharp) Scores (IM101100).....	21
Table 12. ACR 20 Responses at Day 180 (IM101100).....	21
Table 13. ACR Responses at Days 180 and 360 (IM101100).....	22
Table 14. Physical Function (mHAQ) at Days 180 and 360 (IM101100).....	23
Table 15. Mean Change from Baseline to Day 180 and Day 360 for the SF-36 (Physical and Mental Health Component Summary Scores) (IM101100).....	24
Table 16. Radiographic Results at Day 360 - Pre-Specified Genant-Modified Sharp Analysis (IM101100).....	25
Table 17. Reasons for Discontinuation: Days 1-169 and Days 170-365 (IM101102).....	28
Table 18. Baseline Demographic Characteristics (IM101102).....	29
Table 19. Baseline Clinical Rheumatoid Arthritis Characteristics (IM101102).....	30
Table 20. Baseline Radiographic Scores Genant-Modified Sharp Scoring System (IM101102).....	31
Table 21. ACR 20 Responses at Day 169 (IM101102).....	31
Table 22. Proportion of Subjects with Clinically Meaningful HAQ Responses at Day 365 (IM101102).....	31
Table 23. Radiographic Erosion Score Results at Day 365 Genant-Modified Sharp Scoring System (IM101102).....	32
Table 24. ACR 50 and ACR 70 Responses at Day 169 and at Day 365 (IM101102).....	33
Table 25. Major Clinical Response by Day 365 (IM101102).....	33
Table 26. Mean Change from Baseline Up to Days 169 and 365 for the HAQ Disability Index (LOCF Analysis)-IM101102.....	34
Table 27. Radiographic Joint Space Narrowing and Total Score Results at Day 365 Genant-Modified Sharp Scoring System (IM101102).....	35
Table 28. ACR20 Responses Analyzed by GEE (IM101102).....	35
Table 29. HAQ Responses (>0.3 units) Analyzed by GEE (IM101102).....	36
Table 30. Sensitivity Analyses Results for Radiographic Scores (IM101102).....	36
Table 31. Reasons for Discontinuation during Double-Blind Therapy (IM101029).....	36
Table 32. Baseline Demographic Characteristics (IM101029).....	39
Table 33. Baseline Clinical Rheumatoid Arthritis Characteristics (IM101029).....	40
Table 34. Etanercept and Infliximab User Status at Enrollment (IM101029).....	41
Table 35. Duration of Etanercept and Infliximab Use Prior to Randomization (IM101029).....	42
Table 36. ACR 20 Responses at Day 169 (IM101029).....	43
Table 37. ACR 20 Responses by Baseline Anti-TNF Use Current vs. Prior (IM101029).....	43
Table 38. Proportion of Subjects with Clinically Meaningful HAQ Responses at Day 169 (IM101029).....	44
Table 39. HAQ Response by Baseline Anti-TNF Use - Current vs Prior (IM101029).....	44
Table 40. Mean Change from Baseline Up to Day 169 for the HAQ Disability Index (LOCF Analysis)-IM101029.....	44

Table 41. ACR 50 and ACR 70 Responses at Day 169 (IM101029).....	45
Table 42. ACR Core Components - Median Percent Change from Baseline on Day 169 (LOCF Analysis)-IM101029.....	46
Table 43. Mean Change from Baseline in DAS 28(ESR) at Day 169 (LOCF Analysis)-IM101029	46
Table 44. ACR 20 Response Results Using GEE Approach (IM101029).....	47
Table 45. HAQ Response Results Using GEE Approach (IM101029).....	48
Table 46. Reasons for Discontinuation: Days 1 – 365 (IM101031).....	49
Table 47. Reasons for Discontinuation: Days 1 - 365: Background of Non-biologic vs. Biologic RA Therapy (IM101031).....	50
Table 48. Baseline Demographic Characteristics All Abatacept (IM101031).....	51
Table 49. Baseline Clinical Rheumatoid Arthritis Characteristics (IM101031).....	52
Table 50. Baseline Demographic Characteristics by Background of Non-Biologic or Biologic RA Therapy (IM101031).....	53
Table 51. Baseline Clinical Rheumatoid Arthritis Characteristics by Background of Non-Biologic or Biologic RA Therapy (IM101031).....	54
Table 52. Median Percent Improvement from Baseline in Select ACR Core Components on Day 365 (IM101031).....	55
Table 53. Mean Percent Improvement from Baseline in Select ACR Core Components on Day 365 (IM101031).....	55
Table 54. Mean Change from Baseline for the HAQ Disability Index by Subscale and Visit HAQ Disability Index (IM101031).....	56
Table 55. Subgroup Analysis of ACR 20 Response by baseline demographics (IM101102).....	58
Table 56. Subgroup Analysis of HAQ Response by baseline demographics (IM101102).....	58
Table 57. Subgroup Analysis of the change from baseline of erosion score baseline demographics (IM101102).....	59
Table 58. Subgroup Analysis of ACR 20 Response by baseline demographics (IM101029).....	59
Table 59. Subgroup Analysis of HAQ Response by baseline demographics (IM101029).....	60
Table 60. Subgroup Analysis of ACR20 Response by Baseline Disease Activity (IM101102).....	61
Table 61. Subgroup Analysis of HAQ Response by Baseline Disease Activity (IM101102).....	61
Table 62. Subgroup Analysis of Change from Baseline Erosion Score by Baseline Disease Activity (IM101102).....	62
Table 63. Subgroup Analysis of ACR20 Response by Baseline Disease Activity (IM101029).....	63
Table 64. Subgroup Analysis of HAQ Response by Baseline Disease Activity (IM101029).....	63

Appears This Way
On Original

LIST OF FIGURES

Figure 1. Proportion of Subjects Who Discontinued Through Day 360 for any reason (IM101100)
.....20

Figure 2. Summary of ACR 20 Response by Visit (IM101100).....22

Figure 3. Proportion of Subjects Who Discontinued Through Day 365 (All Reasons Combined)
(IM101102).....29

Figure 4: ACR 20 Responses Over Time (IM101102).....32

Figure 5. HAQ Responses Over Time (IM101102).....34

Figure 6. Proportion of Subjects who Discontinued During Double-Blind Therapy (All Reasons
Combined) Cumulative Proportion of Subjects (%) -IM101029.....40

Figure 7. ACR 20 Responses Over Time (IM101029).....45

Figure 8. HAQ Responses Over Time (IM101029).....47

Figure 9. Proportion of Subjects who Discontinued Through Day 365 (All Reasons Combined)
(IM101031).....50

Appears This Way
On Original

1. EXECUTIVE SUMMARY

The BLA 125118 is a submission on Abatacept for the treatment of moderately to severely active Rheumatoid Arthritis (RA) patients, who have had an inadequate response to one or more disease-modifying antirheumatic drugs (DMARDs).

1.1 Conclusions and Recommendations

Based on the results of the statistical analyses of the efficacy data of three principal studies and one safety study, the use of abatacept, given intravenously at a fixed dose approximating 10 mg/kg, shows in reducing the signs and symptoms of RA (See Table 2), improving physical function (See Table 3), inhibiting the progression of structural damage (See Table 4) in adults with moderately to severely active RA, who have had an inadequate response to one or more disease-modifying antirheumatic drugs (DMARDs), including TNF-blocking agents. These studies provide statistical support for the efficacy claim.

1.2 Brief Overview of Clinical Studies

Three efficacy studies (IM101100, IM101102, and IM101029) on abatacept are to evaluate the efficacy and safety of drug in treatment of moderately to severely active RA and one safety study (IM101031) is to assess the overall safety of abatacept.

1.2.1 Study Design

All three studies for efficacy and safety and one safety assessment study were randomized, double blind, multi-center, parallel dosing and placebo controlled. After the double-blind, placebo-controlled period of the studies (12 months for IM101100, IM101102 and IM101031 and 6 months for IM101029), subjects were eligible to receive treatment with a fixed dose abatacept in an open-label, uncontrolled setting.

Study IM101100 was conducted in total 66 sites: 32 sites USA, 23 sites in Europe, 7 sites in Canada, 2 sites in South America, and 2 sites in South Africa. Study IM101102 was conducted in total 116 sites: 31 sites in USA, 36 sites in Europe, 24 sites in Central and South America, and 13 sites in Canada. Study IM101029 was conducted in total 101 sites: 69 sites in USA, 24 sites in Europe, and 8 sites in Canada. Study IM101031 was conducted a total of 161 sites: 91 sites in the United States, 24 sites in Europe (Czech Republic, France, Hungary, Italy, Poland, Russia, Spain), 8 sites in Canada, 4 sites in Australia, 3 sites in Argentina, 10 sites in Brazil, 7 sites in Mexico, 3 sites in Peru, 2 sites in Thailand, 4 sites in Taiwan, and 5 sites in Turkey.

Two studies, Study IM101100 and Study IM101102, enrolled an RA population with an inadequate response to MTX treatment and Study IM101029 enrolled subjects with an inadequate efficacy response to TNF-blocking agents.

In Study IM101029, subjects were defined as having an inadequate efficacy response to a TNF-blocking agent if after ≥ 3 months of therapy with etanercept and/or infliximab, there was no clinical response or a clinically insignificant response.

In Study IM101031, subjects were eligible if they had active disease (defined as an average subject global assessment of disease activity of ≥ 20 mm [using a VAS] at screening and Day 1) despite treatment with 1 or more non-biologic DMARD and/or biologic drug approved for RA, or

their combination for at least 3 months; and were at a stable drug regimen for 28 days prior to study start. Subjects were maintained on their background RA therapy at the dose(s) they were receiving at the time of randomization. This study objective is for the safety assessment.

1.2.2 Study Population

In Study IM101100, 339 subjects received abatacept at 2 mg/kg (n =105) or 10 mg/kg (n = 115) or placebo (n = 119) IV for a total of 13 doses (1 year), in combination with MTX.

In Study IM101102, 652 subjects received fixed-dose abatacept (n = 433) or placebo (n = 219) IV for a total of 14 doses (one year).

In Study IM101029, a total of 391 subjects were randomized and treated with abatacept (n=258) or placebo (n = 133) for a total of 7 doses (six months).

In Study IM101031, a total of 1441 subjects were randomized and treated with either abatacept (n = 959) or placebo (n = 482) for a total of 14 doses (one year).

The intent-to-treat (ITT) analysis population which includes all subjects randomized into the study was used for efficacy analyses by excluding all subjects who were randomized but never received study medication. The following table presents the number of patients at randomization and discontinued in each group for the four studies.

Table 1. Number of Patients Randomized and Discontinued by Treatment

Study	Patient Population	Treatment			Total
		Abatacept 10mg n,(%)	Abatacept 2mg n,(%)	Placebo n,(%)	
IM101100	Randomized	115	105	119	339
	Discontinued at day 180	17(15)	25(24)	41(34)	83(24)
	Discontinued at day 360	8(7)	6(6)	7(6)	21(6)
IM101102	Randomized	433		219	652
	Discontinued at day 169	32(7)		45(21)	77(12)
	Discontinued at day 365	16(4)		12(5)	28(4)
IM101029	Randomized	258		133	391
	Discontinued at day 169	35(14)		34(26)	69(18)
IM101031	Randomized	959		482	1441
	Discontinued at day 365	123(13)		87(18)	210(15)

Table 1 shows that the discontinuation rates were higher in the placebo groups than the treated groups: Lack of efficacy and AEs were the most common reasons for discontinuation across four studies (See Tables 8, 17, 31, 46 and 47).

1.2.3 Dosage

Study IM101100 was a comparison of abatacept 10 mg/kg and 2 mg/kg vs. placebo. Studies IM101102 and IM101029 were comparisons of abatacept at a fixed dose by weight range that approximated 10 mg/kg and placebos. The fixed-dose regimen recommended 500 mg for subjects < 60 kg, 750 mg for subjects 60 to 100 kg, and 1 g for subjects > 100 kg. Study IM101031 was a comparison of abatacept 10 mg/kg vs. placebo.

1.2.4 Endpoints

The efficacy of abatacept was assessed as follows;

- improvement signs and symptoms of RA was evaluated using ACR20 responses
- improvement physical function was measured by the Health Assessment Questionnaire Disability Index (HAQ-DI) using HAQ or MHAQ (modified HAQ).
- inhibition of the progression of structural damage was assessed by the Genant-Modified Sharp scoring system.
- improvement in health-related quality of life, the validated Medical Outcomes Study Short Form (MOS SF-36 or SF-36) was used.

ACR 20 response rate at six months was a primary objective in all three efficacy studies. In Study IM101100, ACR 20 response at six months was the sole primary objective. In Study IM101102, ACR 20 response at six months, HAQ-DI response at 12 months and change in joint erosion score at 12 months were co-primary objectives. In Study IM101029, ACR 20 response and HAQ-DI response at six months were co-primary objectives. For the IM101102 and IM101029 with co-primary objectives, a hierarchical testing procedure was employed with sequential comparisons for ACR 20, followed by HAQ and then change in joint erosion score. The second and third co-primary analyses were only performed if the preceding objective(s) had been met.

The major secondary endpoints include ACR50, ACR70, major clinical response, change from baseline of SF-36 physical component and SF-36 mental component, change in joint space narrowing scores, and change in total Genant-modified Sharp scores (IM101102).

In Study IM101031, the primary objective is the safety assessment and the select ACR core components of the Subject Pain Assessment, the Subject Global Assessment of Disease Activity, the Physician's Global Assessment of Disease Activity, and the disability index of HAQ were observed as exploratory efficacy measures in this study.

1.3 Statistical Issues and Finding

1.3.1 Sponsor's Analysis

The ACR 20, ACR 50 and ACR 70 response rates at 6 and/or 12 months between abatacept and placebo, were analyzed using continuity corrected Chi-square tests. Cochran-Mantel Haenzel tests, stratified by baseline TNF-blocking agent status (current or prior user status), were used in IM101029 for a comparison of the ACR 20 response rates.

For the method of imputation of missing values in ACR scores in the primary analysis, non-responders (IM101102 and IM101029) and LOCF (IM101100, also for missing other than lack of efficacy) were used.

In Study IM101100, the modified Health Assessment Questionnaire (mHAQ) was used to assess improvement in physical function and the full HAQ (includes more items and is more sensitive) was used in the Phase III studies of Study IM101102, Study IM101029 and Study IM10103. The HAQ/mHAQ response was defined as at least a 0.3 unit reduction in HAQ/mHAQ score from

baseline. Similar methods to those stated above for the ACR 20 response rates were used for treatment comparisons of HAQ/mHAQ response rates.

Only when the comparison of the ACR 20 response rate was statistically significant, comparisons of ACR 50 and ACR 70 response rates were performed hierarchically. In the Phase III trials, subjects who discontinued for any reason prior to 6 months were prospectively defined as non-responders for both the signs and symptoms analyses and the HAQ analyses. In addition, several sensitivity analyses were used to demonstrate the robustness of the effect of abatacept, including the worst case scenario where the treatment subjects who discontinued were considered non-responders and the placebo subjects as responders.

Table 2. Summary Efficacy Results of Improvement of Signs and Symptoms

	ACR20 Responses/n(%) Diff* p-value	ACR50 Responses/n(%) Diff* p-value	ACR70 Responses/n(%) Diff* p-value
IM101100			
<u>Day 180</u>			
Abatacept 10mg	70/115 (61) <0.001	42/115 (37) <0.001	19/115 (17) <0.001
Abatacept 2 mg	44/105 (42) 0.31	24/105 (23) 0.027	11/105 (11) 0.005
Placebo	42/119 (35)	14/119 (12)	2/119 (2)
<u>Day 360</u>			
Abatacept 10mg	72/115 (63) <0.001	48/115 (42) <0.001	24/115 (21) 0.003
Abatacept 2 mg	44/105 (42) 0.377	24/105 (23) 0.625	13/105 (12) 0.227
Placebo	43/119 (36)	24/119 (20)	9/119 (8)
IM101102			
<u>Day 169</u>			
Abatacept 10mg	288/424(68) <0.001	169/424(40) <0.001	84/424(20) <0.001
Placebo	85/214(40)	36/214(17)	14/214(7)
<u>Day 365</u>			
Abatacept 10mg	310/424(73) <0.001	205/424(48) <0.001	122/424(29) <0.001
Placebo	85/214(40)	39/214(18)	13/214(6)
IM101029			
<u>Day 169</u>			
Abatacept 10mg	129/256(50) <0.001	52/256(20) <0.001	26/256(10) 0.003
Placebo	26/133(20)	5/133(4)	2/133(2)

*:Difference between abatacept and placebo.

All three studies showed that abatacept 10 mg had highly significant ACR 20, ACR 50, and ACR 70 responses (all $p < 0.003$) as compared to that of the placebo groups and the magnitude of the difference in response was similar in all 3 studies. In Study 101100, there was no difference in the ACR 20 response between abatacept 2mg and placebo, but ACR 50 and ACR 70 responses were significantly higher in the abatacept group as compared to the placebo group. The responses for subjects with an inadequate response to MTX and subjects with an inadequate response to TNF-blocking agents were consistent and statistically significant, regardless of the status of TNF-blocking agent use (current or prior users) and regardless of whether the subject had received etanercept or infliximab.

Table 3. Summary of Efficacy Results of Improvement in Physical Function

	HAQ Responses/n(%)	difference* p-value	mHAQ Responses/n(%)	difference* p-value
IM101100				
<u>Day 180</u>				
Abatacept 10mg			54/115 (47)	0.002
Abatacept 2 mg			40/105 (38)	0.099
Placebo			33/119 (28)	
<u>Day 360</u>				
Abatacept 10mg			44/115 (38)	0.002
Abatacept 2 mg			31/105 (30)	0.104
Placebo			24/119 (20)	
IM101102				
<u>Day 169</u>				
Abatacept 10mg	259/424(61)	<0.001		
Placebo	97/214(45)			
<u>Day 365</u>				
Abatacept 10mg	270/424(64)	<0.001		
Placebo	84/214(39)			
IM101029				
<u>Day 169</u>				
Abatacept 10mg	121/256(47)	<0.001		
Placebo	31/133(23)			
IM101031				
<u>Day 169</u>				
Abatacept 10mg	474/948(50)	<0.001		
Placebo	161/477(34)			
<u>Day 365</u>				
Abatacept 10mg	448/948(47)	<0.001		
Placebo	165/477(35)			

*Difference between the abatacept and the placebo group

All four studies showed that HAQ/mHAQ responses at 6 months or 12 months were significantly higher in the abatacept 10mg groups than that of placebo groups ($p < 0.002$), but not abatacept 2 mg group (IM101100).

Structural joint damage was assessed radiographically and expressed as change in Genant-modified Sharp Total Score and its components, the erosion score (ES) and joint space narrowing (JSN) score at Month 12 compared to baseline in Study IM101102. Structural damage measures were included as tertiary endpoints in Study IM101100 in order to collect preliminary data regarding the effects of abatacept on inhibition of structural damage that would allow the design of a study that was adequately powered to detect a difference in structural damage progression in the Phase III program.

Table 4. Summary of Efficacy Results of Change from Baseline Radiographic Scores

	Erosion		Joint Space Narrowing		Total Xray Score	
	Median	Mean (SD)	Median	Mean (SD)	Median	Mean (SD)
IM101100						
<u>Day 360</u>						
Abatacept 10mg (n=89)		0.50(1.8)		0.8(3.0)		1.3(4.3)
Abatacept 2 mg (n=73)		0.50(1.0)		0.5(1.2)		1.0(1.9)
Placebo (n=70)		0.85(1.7)		0.6(1.3)		1.5(2.5)
95% CI of Diff ¹		-0.4(-0.9, 0.2)		0.2(-0.5, 0.9)		-0.2(-1.3, 0.9)
95% CI of Diff ²		-0.4(-0.8, 0.1)		0.13(-0.5, 0.3)		-0.5(-1.2, 0.2)
IM101102						
<u>Day 365</u>						
Abatacept 10mg (n=391)	0.0	0.63(1.8)	0.0	0.58(1.5)	0.25	1.21(2.9)
Placebo (n=195)	0.27	1.14(2.8)	0.0	1.18(2.6)	0.53	2.32(5.0)
p-value	0.029		0.009		0.012	

Diff¹: Difference between abatacept 10 mg vs. Placebo

Diff²: Difference between abatacept 2 mg vs. Placebo

The baseline radiographic parameters were similar among and between treatment groups in IM101100 (See Table 12) and IM101102 (See Table 22). The mean and median changes from baseline in radiographic measurements at 1 year in IM101102 showed significant reduction in radiographic progression compared with placebo (See Tables 22 and 26). The mean change from baseline in radiographic measurements at 1 year in IM101100 was also similar to the study IM101102.

Health-related quality of life was assessed by the SF-36 questionnaire at 6 months in studies, IM101100, IM101102, and IM101029 and at 12 months in studies IM101100 and IM101102.

Appears This Way
On Original

Table 5. Summary of Efficacy Results of Improvement in Health-related Quality of Life

	Physical Component Summary		Mental Component Summary	
	Mean Change (SD)	p-value	Mean Change (SD)	p-value
IM101100				
<u>Day 180</u>				
Abatacept 10mg	8.2	0.0001	5.8	0.0092
Abatacept 2 mg	4.5	0.1987	2.7	0.8726
Placebo	3.1		2.5	
<u>Day 360</u>				
Abatacept 10mg	8.0	0.0001	6.4	0.0005
Abatacept 2 mg	5.2	0.0437	3.4	0.3463
Placebo	2.9		2.3	
IM101102				
<u>Day 169</u>				
Abatacept 10mg	8.8(0.4)	<0.001	6.2(0.5)	0.005
Placebo	4.8(0.6)		3.8(0.7)	
<u>Day 365</u>				
Abatacept 10mg	9.1(0.4)	<0.001	6.9(0.5)	0.011
Placebo	5.0(0.6)		4.7(0.7)	
IM101029				
<u>Day 169</u>				
Abatacept 10mg	6.6(0.6)	<0.001	5.2(0.6)	0.005
Placebo	1.1(0.8)		2.1(0.9)	

In these studies, clinically and statistically significant improvement was observed in the abatacept groups as compared with the placebo group in all 8 domains of the SF-36 as well as the Physical Component Summary (PCS) and the Mental Component Summary (MCS). Similar results were obtained in studies IM101100 (in subjects with inadequate response to MTX) and IM101029 (in subjects with inadequate response to TNF-blocking agents).

1.3.2 Reviewer's Analysis

The reviewer's analyses focused on the two efficacy of Phase 3 studies IM101102 and IM101029. The co-primary endpoints of ACR20 and HAQ-DI responses were analyzed using generalized estimating equation (GEE) methods for repeated measures as supportive analyses of the primary analyses.

Sensitivity analyses for change from baseline radiographic measures were performed because the co-primary endpoint of the change from baseline erosion score was performed using all observed data at baseline and at 12 months for Study IM101102.

The ACR20 response was analyzed using GEE methods for Phase III studies, IM101102 and IM101029 from visit day 1 to up to visit day 169. The models included age and weight as covariates. The Table 6 summarizes the ACR20 responses analyses results.

Table 6. Summary of Efficacy Results of ACR20 Responses using GEE Methods

	Placebo	Abatacept	Difference	p-value
<u>IM101102</u>				
Least Square Means	-1.2467	-0.7490	0.4977	<0.0001
Exp(Log OR)			1.6449	
CI			(1.43, 1.90)	
<u>IM101029</u>				
Least Square Means	-1.6866	-0.5348	1.1518	<0.0001
Exp(Log OR)			3.1639	
CI			(2.31, 4.34)	

The models included age and weight as covariates.

There was significant difference between the abatacept group and placebo group in the ACR20 responses overall time up to visit day 169 (IM101102 and IM101029) and these results were robust. The abatacept group had 1.6 times and 3.1 times higher in odds ratio for ACR 20 responses than the placebo group overall visits up to day 169 in studies, IM101102 and IM101029, respectively.

The repeated measures of HAQ-DI responses were analyzed using GEE method up to days 169 and 365 for both studies. The results are summarized in Table 7.

Table 7. Summary of Efficacy Results of HAQ_DI responses by GEE Methods

	Placebo	Abatacept	Difference	p-value
<u>IM101102</u>				
<u>Up to day 169</u>				
Least Square Means	-0.8028	-1.0615	0.2587	0.0005
Exp(Log OR)			1.2953	
CI			(1.12, 1.50)	
<u>Up to day 365</u>				
Least Square Means	-0.7001	-1.0063	0.3061	<0.0001
Exp(Log OR)			1.3582	
CI			(1.18, 1.56)	
<u>IM101029</u>				
<u>Up to day 169</u>				
Least Square Means	-1.0008	-1.6137	0.6128	<0.0001
Exp(Log OR)			1.8456	
CI			(1.44, 2.37)	

The models included age and weight as covariates

The abatacept group had statistically significantly higher response (at least 1.4 times in odds ratio) in HAQ-DI overall up to day 169 as compared to the placebo group for both studies.

For radiographic measures, the results of sensitivity analyses imputing 12 months missing values with the median of 12 months values calculated by the same baseline value category, the results were robust. However, after imputing baseline values for those who have only 12 months values but missing baseline values with a similar imputation method, the results show that there was no

statistically significant difference in the co-primary efficacy of the change from baseline radiographic scores. The results were not robust (See Table 29).

1.3.3 Statistical Issues

The probability of dependent drop-out differs between the placebo group and the treated group and this may lead to biased comparison of the effect if the missing data are ignored. More patients in the MTX group than in the abatacept group dropped out of the study due to lack of efficacy. The sponsor's sensitivity analyses for primary efficacy variables using several imputation methods including a worst case were robust. Reviewer's results for supportive analyses of GEE methods for primary endpoints ACR20 and HAQ responses using repeated measures for IM101102 and IM101029 were similar.

For a co-primary endpoint in IM101102, the change from baseline in joint erosion score at 12 months, all the subjects who had a baseline and a day 365 values were used for the primary analyses. The reviewer performed sensitivity analyses by imputing first missing day 365 values, then additionally, imputing missing baseline values. The result of sensitivity analysis imputing missing 12 months values was robust, but after imputing the baseline values, the sensitivity analysis result was not robust. Although the sensitivity analysis result of reviewer's radiographic measures was not robust, the result showed that there was numerically less worsening from baseline of erosion score as well as joint space narrowing score in the abatacept group as compared to the placebo group.

Overall, the results of analyses performed by the sponsor and by the reviewer were consistently in favor of abatacept group compare to the placebo group in the effectiveness for the treatment of moderately to severely active Rheumatoid Arthritis (RA) patient.

2. INTRODUCTION

2.1 Overview

RA is a chronic, systemic, inflammatory and destructive autoimmune disease with clinical manifestations that primarily involves the synovial joints. Rheumatoid arthritis (RA) affects approximately 1% of the population worldwide and commonly leads to severe, chronic functional disability and a reduced quality of life.

2.1.1 History of Drug Development

Abatacept (BMS-188667, CTLA4Ig) is the first agent in a new class of drugs, the selective T-cell costimulation modulators, for the treatment of RA. Abatacept is a fully human, recombinant, soluble fusion protein consisting of the extracellular domain of human CTLA-4 and a fragment (hinge-CH2-CH3 domains) of the Fc domain of human IgG1.

Abatacept works upstream, directly on T-cell activation, may lead to broader utility, including use in patients not responding adequately to TNF-blocking agents in contrast to anti-cytokine RA therapies such as the TNF blocking agents that work downstream following T-cell activation. Furthermore, since abatacept inhibits, but does not completely block T-cell activation, important aspects of host defense, especially innate immunity, remain functional. Based upon the unique mechanism of action, abatacept may provide an important new therapeutic option for RA patients.

2.1.2 Clinical Development Program

The abatacept clinical development program in RA consists of 6 Phase II and Phase III studies. One study was a preliminary dose-finding, Phase II study of abatacept monotherapy (IM103002) and the other remaining 5 RA studies included 3 principal studies for demonstrating efficacy: Study IM101100, Study IM101102, and Study IM101029, a large Phase III safety study: Study IM101031, and a Phase II study of abatacept in combination with etanercept (IM101101). These 5 studies, IM101100, IM101102, IM101029, IM101031, and IM101101 have completed, double-blind, placebo-controlled treatment periods during which 1797 subjects were treated with 10 mg/kg or a fixed dose regimen approximating 10 mg/kg of abatacept are integrated for safety analyses. The double-blind periods in these studies were followed by ongoing, open label, uncontrolled treatment periods during which 2285 subjects have been treated with a fixed dose regimen approximating 10 mg/kg through 28-Jul-2004.

Two of the studies, IM101100 and IM101102 enrolled an RA population with an inadequate response to MTX treatment and IM101029 study enrolled subjects with an inadequate efficacy response to TNF-blocking agents after at least 3 months of therapy with either etanercept and/or infliximab. The MTX-failure population was defined as having active disease in spite of MTX treatment with a minimum of 15 mg/week for at least 3 months in IM101102, or 10 to 30 mg/week for at least 6 months in Study IM101100. Subjects continued to receive MTX during the studies.

In IM103002, abatacept monotherapy at doses of 10 and 2 mg/kg was generally well tolerated and effective at reducing the signs and symptoms of RA. Based on the Study IM103002 results and the standard practice of adding other therapies to methotrexate (MTX), this study IM101101 was designed to test the hypothesis that abatacept plus etanercept may have greater clinical efficacy when compared with etanercept plus placebo in RA subjects with active disease despite treatment with etanercept.

All 3 principal studies, IM101100 (n=339), IM101102 (n=652), and IM101029 (n=393) were all randomized, double-blind, and placebo controlled. The proposed dose regimen (or a similar dose regimen) was used in all 3 studies. Study IM101100 was a comparison of abatacept 10 mg/kg (n=115) and 2 mg/kg (n=105) vs. placebo (n=109); Studies IM101102 and IM101029 were comparisons of abatacept at a fixed dose by weight range that approximated 10 mg/kg (n=433 for IM101102 and n=258 for IM101029) and placebo (n=219 for IM101102 and n=133 for IM101029). The fixed-dose regimen recommended 500 mg for subjects < 60 kg, 750 mg for subjects 60 to 100 kg, and 1 g for subjects > 100 kg.

Background medication was closely controlled in all 3 principal studies to avoid confounding the interpretation of efficacy due to an imbalance in the contribution of other active RA medications between treatment groups. In IM101100 and IM101102, low-dose corticosteroids and/or NSAIDs were allowed on a background of MTX. In IM101029, low-dose corticosteroids and/or NSAIDs were allowed on a background of non-biologic DMARDs or anakinra. Change of background therapy, including change of dose (except a decrease of MTX dose because of toxicity), was not allowed during the first 6 months of the studies. In contrast, subjects enrolled in the large general study (IM101031) were allowed to take a variety of background DMARDs (both biologic and non-biologic) and adaptations to their RA medication were allowed after 3 months in the study.

After the double-blind, placebo-controlled period of the studies (12 months for IM101100 and IM101102 and 6 months for IM101029), subjects were eligible to receive treatment with fixed

dose abatacept in an open-label, uncontrolled setting. Open-label treatment was also offered to subjects in IM101101 and IM101031. Long-term, open-label efficacy data up to 3 years are available from IM100100 and IM100101. Subjects from the other studies have only been treated in the open-label periods for a small number of months (approximately 3 months) and efficacy data are not yet available.

2.1.3 Objectives in Treatment of Rheumatoid Arthritis

The RA studies utilized specific, validated measures to assess efficacy as summarized in the FDA and EMEA guidances for industry for the development of products for the treatment of RA.

- Signs and symptoms of RA using the ACR core data set and response criteria
- Physical disability using the disability index (DI) of the Health Assessment Questionnaire (HAQ)
- Inhibition of structural damage progression using the Genant-modified Sharp scoring System
- Quality of life with SF-36

The ACR 20 response rate at 6 months was a primary endpoint in all 3 principal studies. In IM101100, ACR 20 response was the sole primary endpoint. In IM101102, ACR 20 response at 6 months, HAQ-DI response at 12 months and change in joint erosion score at 12 months were co-primary endpoints. In IM101029, ACR 20 response and HAQ-DI response at 6 months were co-primary endpoints. In the 2 studies with co-primary endpoints, a hierarchical testing procedure was employed with sequential comparisons for ACR 20, followed by HAQ and then change in joint erosion score. The second and third co-primary analyses were only performed if the preceding endpoint(s) had been met.

2.2 Data Sources

<\\CBS5042329\MEDR Submissions\2004BLA\DCC60000241\roadmap.pdf>

3. STATISTICAL EVALUATION

This review focuses on the efficacy results of the Phase IIb of IM101100, and Phase III studies, IM101102, IM101029 and IM101031.

3.1 Evaluation of Efficacy

3.1.1 Study IM101100

3.1.1.1 Study Design

Study IM101100 was a randomized, double-blind, placebo-controlled parallel-group with a treatment period of 12 months (2 phases: Days 1 to 180 and Days 181 to 360).

During Days 1 to 180; subjects were maintained on a stable dose of MTX (10-30 mg/wk). Low-dose stable systemic (oral or injectable) corticosteroids (10 mg/day or less) and/or stable non-steroidal anti-inflammatory drugs (NSAIDs) including aspirin (ASA) were allowed. Decreases in NSAID dose were permitted only for toxicity and no additional DMARDs were allowed.

During Days 181 to 360; adjustments in corticosteroids (maximum of 10 mg/day) and MTX (maximum of 30 mg/wk) were permitted, as was addition of either hydroxychloroquine, sulfasalazine, gold or azathioprine. Adjustments in NSAID dose or addition of NSAIDs including ASA were also permitted. No other changes in anti-rheumatic therapy were allowed. After completing the 12-month treatment period, subjects could have continued in a long-term extension. Results of the long-term extension will be reported separately.

Prior to study start, subjects must have been treated with MTX (10-30 mg/wk) for at least 6 months, and at a stable dose for 28 days prior to Day 1.

Subjects of 339 were randomized 1:1:1 to one of 3 treatment groups:

- 1) Abatacept 10 mg/kg by intravenous infusion (n=115)
- 2) Abatacept 2 mg/kg by intravenous infusion (n=105)
- 3) Abatacept placebo by intravenous infusion (n=119)

3.1.1.2 Primary and Secondary Efficacy Endpoints

The primary efficacy endpoint is the ACR 20 at 6 months of 2 different doses of abatacept (10 and 2 mg/kg) combined with MTX with MTX plus placebo in subjects with active RA.

There are 8 secondary endpoints: ACR 50 and ACR 70 at Day 169 (Month 6), ACR 20, ACR 50, ACR 70 and ACR-N over time, major clinical response achievement (continuous ACR 70 for 6 months), onset of action and time to maximal response, safety of chronic use, lack of immunogenicity, changes in surrogate markers (CRP, soluble IL-2r [sIL-2r]) in subjects receiving abatacept in combination with MTX, and the improvement in physical function as assessed by the Modified Health Assessment Questionnaire (mHAQ) and the physical component summary scale of the SF-36 questionnaire in subjects receiving abatacept in combination with MTX.

3.1.1.3 Sponsor's Statistical Methods

All efficacy analyses were performed on the intent-to-treat population (ITT), which consisted of all randomized subjects who received at least one dose of study medication (primary efficacy data set). Other data sets were derived from the primary efficacy data set, including the last observation carried forward (LOCF) data set and worst observation carried forward (WOCF) data set for sensitivity analyses.

The primary efficacy analysis tested for differences in ACR 20 response between the two abatacept treatment groups and the placebo group at 6 months (Day 180) using the Cochran Mantel-Haenszel test. A sequential testing procedure was used by comparing data for the 10 mg/kg group with the data for the placebo group at the 0.05 level of significance first. If this was significant, the data for the 2 mg/kg group was compared with the placebo group at the 0.05 level to preserve the overall alpha level of 5%. Similar analyses were carried out for the ACR 50 and ACR 70 responses at 6 months.

Subjects who discontinued the study due to lack of efficacy (ie, worsening RA) were considered ACR non-responders at all subsequent time points. For all subjects who discontinued for other reasons, their last ACR response was carried forward. All statistical tests and confidence intervals were 2-sided.

ACR 20, ACR 50, and ACR 70 response rates on days 180 and 360 were compared between each

abatacept treatment group and the placebo group at the Dunnett-adjusted 0.027 (two-tailed) level of significance and the differences in ACR responses between the abatacept groups and the placebo group were summarized using 95% confidence intervals.

The modified Health Assessment Questionnaire (mHAQ) was used to assess disability or physical function which evaluates in 8 domains: dressing, arising, eating, walking, hygiene, reach, grip and common activities. The mean of the 8 scores (scaled to be between 0 and 3) and mean percent change for baseline (with 95% confidence intervals) were used to assess physical function. A subject was considered to be a mHAQ responder if there was a reduction from baseline in their mHAQ score of at least 0.30 at Days 180 and 360. For the mHAQ responder analysis, all subjects who discontinued were considered as mHAQ non-responders subsequent to their discontinuation.

3.1.1.4 Patient Disposition, Demographic and Baseline Characteristics

Table 8 summarizes the reason for discontinuation for Study IM101100 from day 1 through day 360.

Table 8. Reasons for Discontinuation: Days 1- 360 (Study IM101100)

	Number (%) of Subjects			
	Abatacept 10 mg/kg + MTX (N=115)	Abatacept 2 mg/kg + MTX (N=105)	Placebo + MTX (N=119)	Total (N=339)
No. Discontinued	17 (14.8)	25 (23.8)	41 (34.5)	83 (24.5)
Lack of Efficacy	12 (10.4)	16 (15.2)	28 (23.5)	56 (16.5)
Adverse Event	3 (2.6)	7 (6.7)	9 (7.6)	19 (5.6)
Withdrawal of Consent	2 (1.7)	2 (1.9)	4 (3.4)	8 (2.4)
Completed 180 Days of Therapy	98 (85.2)	80 (76.2)	78 (65.5)	256 (75.5)
No. Discontinued	25 (21.7)	31 (29.5)	48 (40.3)	104 (30.7)
Death	0	1 (1.0)	0	1 (0.3)
Adverse Event	5 (4.3) ^{ai}	9 (8.6)	11 (9.2)	25 (7.4)
Lack of Efficacy	13 (11.3)	17 (16.2)	30 (25.2)	60 (17.7)
Lost to Follow-up	1 (0.9)	2 (1.9)	0	3 (0.9)
Withdrawal of Consent	5 (4.3)	2 (1.9)	6 (5.0)	13 (3.8)
Other	1 (0.9)	0	1 (0.8)	2 (0.6)
Completed 360 Days of Therapy	90 (78.3)	74 (70.5)	71 (59.7)	235 (69.3)

Lack of efficacy and AEs were the most common reasons for discontinuation. The proportion of subjects who discontinued for these reasons was more than 2 times higher in the placebo group than in the 10 mg/kg group. More subjects in the abatacept groups completed 180 days of treatment compared with the placebo group: 98 subjects (85%) in the 10 mg/kg group, 80 subjects (76%) in the 2 mg/kg group and 78 subjects (66%) in the placebo group.

Figure 1 summarizes discontinuation rates for any reason and graphically illustrate the lower discontinuation rates for abatacept compared with placebo, especially at a dose of 10 mg/kg.

Figure 1. Proportion of Subjects Who Discontinued Through Day 360 for any reason (IM101100)

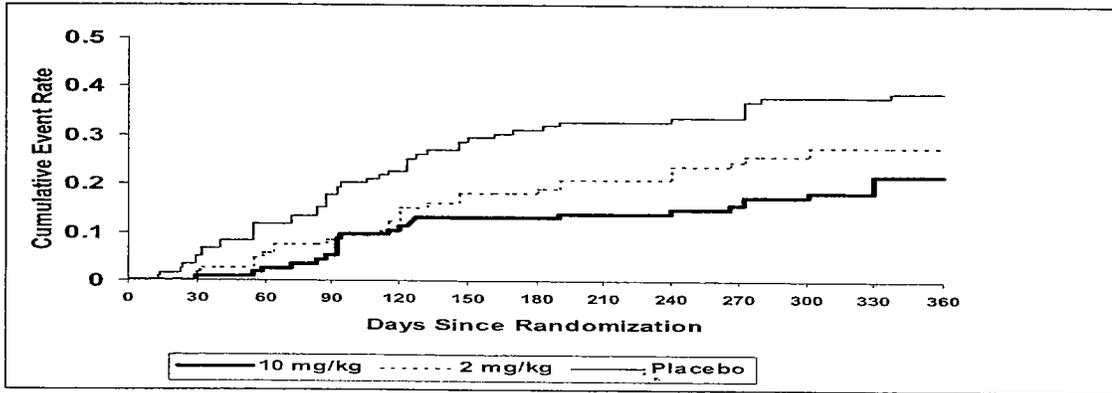


Table 9. Baseline Demographic Characteristics (IM101100)

	Abatacept 10 mg/kg + MTX (N=115)	Abatacept 2 mg/kg + MTX (N=105)	Placebo + MTX (N=119)
Age (yrs)			
Mean ± SD	55.8 ± 12.5	54.4 ± 11.3	54.7 ± 12.0
(Range)	(17-83)	(23-80)	(23-80)
Weight (kg)			
Mean ± SD	77.8 ± 18.6	78.7 ± 21.4	79.9 ± 17.6
(Range)	(40.1-144.0)	(48.4-186.8)	(44.0-140.0)
Gender			
Males, n (%)	29 (25)	39 (37)	40 (34)
Females, n (%)	86 (75)	66 (63)	79 (66)
Race			
White, n (%)	100 (87)	91 (87)	104 (87)
Black, n (%)	6 (5)	0	3 (3)
Other, n (%)	9 (8)	14 (13)	12 (10)

The demographic characteristics were similar among the treatment groups. The majority of subjects were white females, approximately 55 years old.

Appears This Way
On Original

Table 10. Baseline Clinical Rheumatoid Arthritis Characteristics (IM101100)

Characteristic	Abatacept 10 mg/kg + MTX (N=115)	Abatacept 2 mg/kg + MTX (N=105)	Placebo + MTX (N=119)
Duration of RA (yrs)	n=114	n=105	n=117
Mean ± SD (Range)	9.7±9.8 (0-38)	9.7±8.1 (0-36)	8.9±8.3 (0-41)
Tender Joints	n=115	n=105	n=119
Mean ± SD	30.8 ± 12.2	28.2 ± 12.0	29.2 ± 13.0
Range	11.0-66.0	3.0-62.0	4.0-68.0
Swollen Joints	n=115	n=105	n=119
Mean ± SD	21.3 ± 8.4	20.2 ± 8.9	21.8 ± 8.8
Range	9.0-54.0	4.0-48.0	8.0-64.0
Pain (VAS 100 mm)	n=113	n=104	n=119
Mean ± SD	62.1 ± 21.4	64.3 ± 22.3	65.2 ± 22.1
Range	0.0-99.0	8.0-100.0	3.0-95.0
Physical Function (mHAQ 0-3)	n=115	n=105	n=119
Mean ± SD	1.0 ± 0.5	1.0 ± 0.5	1.0 ± 0.6
Range	0.0-2.5	0.0-2.5	0.0-2.3
Subject Global Assessment (VAS 100 mm)	n=113	n=105	n=119
Mean ± SD	60.1 ± 20.7	59.4 ± 23.7	62.8 ± 21.6
Range	10.0-100.0	8.0-99.0	4.0-94.0
Physician Global Assessment (VAS 100 mm)	n=113	n=105	n=119
Mean ± SD	62.1 ± 14.8	61.0 ± 16.7	63.3 ± 15.5
Range	20.0-98.0	8.0-95.0	18.0-93.0
CRP (mg/dL)	n=112	n=99	n=115
Mean ± SD	2.9 ± 2.8	3.2 ± 2.5	3.2 ± 3.2
Range	0.2-19.9	0.2-10.8	0.2-20.9
Morning Stiffness (in minutes)	n=115	n=103	n=119
Mean ± SD	97.9 ± 63.1	104.1 ± 63.9	106.0 ± 64.2
Range	0.0-180.0	0.0-180.0	0.0-180.0
MTX Dose (mg/wk) ^a	n=113	n=103	n=117
Mean ± SD	15.0 ± 4.4	15.8 ± 4.5	15.9 ± 4.1
Range	7.5-25.0	10.0-30.0	7.5-25.0
Rheumatoid Factor (IU/mL)	n=99	n=90	n=90
% Positive	86%	86%	76%

Table 11. Baseline x-ray (Prespecified Genant-Modified Sharp) Scores (IM101100)

Parameter (Baseline Mean ± SD)	Abatacept 10 mg/kg + MTX (N=115)	Abatacept 2 mg/kg + MTX (N=105)	Placebo + MTX (N=119)
	n=115	n=103	n=117
Joint Erosion	21.8 ± 17.8	20.0 ± 14.7	18.5 ± 16.2
Joint Space Narrowing	29.1 ± 27.0	24.3 ± 21.1	25.7 ± 24.5
Total Score	51.0 ± 43.2	44.3 ± 34.5	44.2 ± 39.6

Baseline clinical characteristics and baseline x-ray scores were similar among the treatment groups. Despite concurrent treatment with MTX (mean dose 15-16 mg/week), subjects had a high degree of baseline disease activity on the basis of the number of tender and swollen joints (mean of approximately 29 tender and 21 swollen joints). Subjects had a mean duration of RA of approximately 9 to 10 years and a mean duration of morning stiffness of 98 to 106 minutes.

3.1.1.5 Results and Conclusions

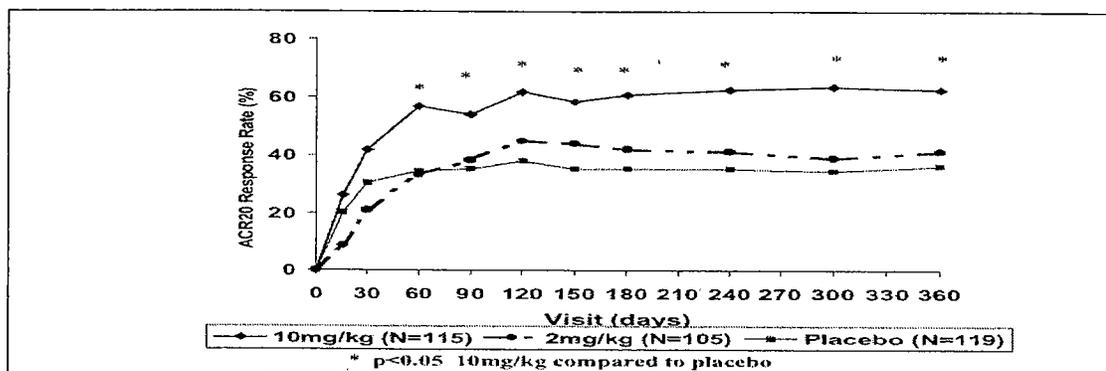
The primary efficacy variable of ACR 20 response at Day 180 was summarized in Table 12.

Table 12. ACR 20 Responses at Day 180 (IM101100)

Parameter	Abatacept 10 mg/kg + MTX (N=115)	Abatacept 2 mg/kg + MTX (N=105)	Placebo + MTX (N=119)
n (%)	70 (60.9)	44 (41.9)	42 (35.3)
Estimate of the difference with respect to placebo (95% CI)	25.6 (12.8, 38.4)	6.6 (-6.2, 19.4)	N/A
p-value ^a	< 0.001	0.31	N/A

Analysis of the primary efficacy variable, the percentage of subjects who had an ACR 20 response at Day 180 was significantly higher in the 10 mg/kg group than in the placebo group (61% vs. 35%, $p < 0.001$). There was no significant difference in the rate of ACR 20 responses at Day 180 between the 2 mg/kg group and the placebo group.

Figure 2. Summary of ACR 20 Response by Visit (IM101100)



Statistically significant improvements in ACR 20 responses (compared with placebo) were seen as early as Day 60 with 10 mg/kg ($p = 0.04$) and the improvements remained statistically significant through Day 360 ($p < 0.05$).

Table 13. ACR Responses at Days 180 and 360 (IM101100)

Parameter	Abatacept 10 mg/kg + MTX (N=115)	Abatacept 2 mg/kg + MTX (N=105)	Placebo + MTX (N=119)
<u>At day 180</u>			
ACR 50			
n (%)	42 (36.5)	24 (22.9)	14 (11.8)
Estimate of the difference with respect to placebo (95% CI)	24.8 (13.8, 35.7)	11.1 (1.2, 20.9)	N/A
p-value ^a	< 0.001	0.027	N/A
ACR 70			
n (%)	19 (16.5)	11 (10.5)	2 (1.7)
Estimate of the difference with respect to placebo (95% CI)	14.8 (7.5, 22.2)	8.8 (2.7, 14.9)	N/A
p-value ^a	< 0.001	0.005	N/A
<u>At day 360</u>			
ACR 20			
n (%)	72 (62.6)	44 (41.9)	43 (36.1)
Estimate of the difference with respect to placebo (95% CI)	26.5 (13.7, 39.3)	5.8 (-7.0, 18.6)	N/A
p-value ^a	< 0.001	0.377	N/A
ACR 50			
n (%)	48 (41.7)	24 (22.9)	24 (20.2)
Estimate of the difference with respect to placebo (95% CI)	21.6 (9.7, 33.4)	2.7 (-8.1, 13.5)	N/A
p-value ^a	< 0.001	0.625	N/A
ACR 70			
n (%)	24 (20.9)	13 (12.4)	9 (7.6)
Estimate of the difference with respect to placebo (95% CI)	13.3 (4.4, 22.2)	4.8 (-3.0, 12.6)	N/A
p-value ^a	0.003	0.227	N/A

The rates of ACR 50 and ACR 70 responses at Day 180 and Day 360 and ACR 20 at Day 360 were significantly higher in both the 10 mg/kg and 2 mg/kg groups ($p < 0.05$) than in the placebo group.

Table 14. Physical Function (mHAQ) at Days 180 and 360 (IM101100)

	Abatacept 10 mg/kg + MTX (N=115)	Abatacept 2 mg/kg + MTX (N=105)	Placebo.+ MTX (N=119)
<u>At day 180</u>			
	n=107 ^a	n=98	n=110
Baseline Mean	1.0	1.1	1.1
Mean % Change	41.2 ^b	21.6	13.7
<u>At day 360</u>			
	n=109 ^a	n=100	n=111
Baseline Mean	1.0	1.1	1.1
Mean % Change	42.3 ^b	22.9	10.3

a Number of subjects with both a baseline mHAQ score and Day 180 or Day 360 mHAQ score.

b Indicates $p < 0.05$ in comparison with placebo.

Subjects treated with 10 mg/kg had a statistically significant improvement in physical function (mHAQ) at Day 180 as compared to the placebo group. The improvement in mHAQ at Day 180 with 2 mg/kg as compared to the placebo group was not statistically significant.

Similar to the observations at Day 180, statistically significant improvements in mHAQ were seen in subjects treated with 10 mg/kg compared to the placebo group at Day 360. Improvement in physical function (mHAQ) at Day 360 with 2 mg/kg was not statistically significant.

Appears This Way
On Original

Table 15. Mean Change from Baseline to Day 180 and Day 360 for the SF-36 (Physical and Mental Health Component Summary Scores)-IM101100

Summary Score ^a	Abatacept 10 mg/kg + MTX (N=115)	Abatacept 2 mg/kg + MTX (N=105)	Placebo + MTX (N=119)
At day 180			
Physical Component Summary	n=115	n=104	n=119
Baseline Mean	30.7	30.7	32.2
Postbaseline Mean	39.0	35.2	35.2
Mean Change from Baseline	8.3	4.6	3.0
Adjusted Difference from placebo ^b	5.0	1.4	
95% CI compared with placebo	(2.5, 7.5)	(-0.7, 3.4)	
p-value	0.0001	0.1987	
Mental Component Summary	n=115	n=104	n=119
Baseline Mean	45.6	43.7	42.2
Postbaseline Mean	50.6	46.4	45.4
Mean Change from Baseline	5.0	2.7	3.2
Adjusted Difference from placebo ^b	3.3	0.2	
95% CI compared with placebo	(0.8, 5.8)	(-2.3, 2.8)	
p-value	0.0092	0.8726	
At day 360			
Physical Component Summary	n=115	n=104	n=119
Baseline Mean	30.7	30.7	32.2
Postbaseline Mean	38.8	36.0	35.0
Mean Change from Baseline	8.1	5.3	2.8
Adjusted Difference from placebo ^b	5.1	2.3	
95% CI compared with placebo	(2.5, 7.6)	(0.1, 4.5)	
p-value	0.0001	0.0437	
Mental Component Summary	n=115	n=104	n=119
Baseline Mean	45.6	43.7	42.2
Postbaseline Mean	51.2	47.2	45.2
Mean Change from Baseline	5.7	3.5	3.0
Adjusted Difference from placebo ^b	4.1	1.1	
95% CI compared with placebo	(1.8, 6.3)	(-1.2, 3.4)	
p-value	0.0005	0.3463	

a Includes all subjects who received at least one dose of study drug and who had SF-36 values at baseline and at Day 180 or Day 360. LOCF used for the analysis.

b Adjustment based on ANCOVA model with treatment and baseline values as covariates.

Statistically significant improvements from baseline in the physical component summary and the mental health component summary scores were observed in the 10 mg/kg group compared with the placebo group at Day 180 and Day 360, but not abatacept 2 mg/kg group. There was no statistically significant difference between abatacept 2mg and placebo in mental health component summary score at days 180 and 360 and in the physical component summary at day 360.

Table 16. Radiographic Results at Day 360 - Pre-Specified Genant-Modified Sharp Analysis (IM101100)

Parameter	Abatacept 10 mg/kg + MTX (N=115)	Abatacept 2 mg/kg + MTX (N=105)	Placebo + MTX (N=119)
No. Subjects with paired radiographs ^a	n=89	n=73	n=70
Joint Erosion			
Baseline Mean ± SD	24.4 (17.7)	20.8 (15.3)	19.5 (15.3)
Postbaseline Mean ± SD	24.9 (17.9)	21.2 (15.4)	20.35 (15.2)
Mean Change ± SD	0.5 (1.8)	0.5 (1.0)	0.85 (1.7)
Difference from Placebo (95% CI)	-0.4 (-0.9, 0.2)	-0.4 (-0.8, 0.1)	
% with progression	32%	32%	39%
Joint Space Narrowing			
Baseline Mean ± SD	33.1 (28.0)	25.1 (21.4)	25.8 (23.8)
Postbaseline Mean ± SD	33.9 (28.0)	25.6 (21.6)	26.4 (23.7)
Mean Change ± SD	0.8 (3.0)	0.5 (1.2)	0.6 (1.3)
Difference from Placebo (95% CI)	0.2 (-0.5, 0.9)	-0.13 (-0.5, 0.3)	
% with progression	21%	32%	30%
Total Score			
Baseline Mean ± SD	57.6 (43.9)	45.9 (35.5)	45.3 (37.7)
Postbaseline Mean ± SD	58.8 (44.2)	46.8 (35.8)	46.7 (37.5)
Mean Change ± SD	1.3 (4.3)	1.0 (1.9)	1.5 (2.5)
Difference from Placebo (95% CI)	-0.2 (-1.3, 0.9)	-0.5 (-1.2, 0.2)	
% with progression	35%	40%	46%

a n= subset of N, subjects who had an x-ray at both baseline and at Day 360.

In the pre-specified analysis, there was a lower proportion of subjects with progression in the abatacept 10 mg as compared to the placebo. A greater number of subjects in the placebo group than in the abatacept groups discontinued due to a lack of clinical efficacy and did not have a termination radiograph.

When compared with placebo, abatacept 10 mg/kg significantly improved the signs and symptoms of disease, physical function and quality of life in subjects who had active RA despite MTX treatment. There were also improvements in many of the efficacy parameters in the abatacept 2 mg/kg group compared with the placebo group, but the majority of these improvements were not statistically significant.

3.1.2 Study IM101102

3.1.2.1 Study Design

Study IM101102 was a randomized, double-blind, placebo-controlled, parallel-dosing design study with a treatment period of 12 months (2 periods: Days 1 to 169 and Days 170 to 365). Subjects of 652 were randomized 2:1 ratio (433 for abatacept and 219 for placebo) of the following treatments on a background of MTX therapy: abatacept or placebo. Subjects must have been treated with MTX (minimum 15 mg/wk; a lower dose of 10 mg/wk was allowed if necessitated due to toxicity) for at least 3 months, and at a stable dose for 28 days prior to treatment (Day 1).

Subjects received a fixed dose of abatacept based upon their screening visit weight: <60 kg received 500 mg, 60 kg to 100 kg received 750 mg, and >100 kg received 1 gram. Study medication was administered each dose intravenously over approximately 30 minutes on Days 1, 15, 29, and every 28 days thereafter for a total of 14 doses.

Background MTX was continued during this study at the dose level and regimen administered at the time of randomization. To minimize potential MTX toxicity, all subjects received either folic acid or leucovorin according to manufacturer recommendations.

Baseline and 1 year radiographs (hands/wrists and feet) were required for all subjects.

Radiographs of the hands/wrists and feet were required at the early termination visit. In addition, all subjects who discontinued early were required to return for radiographs of the hands/wrists and feet and for collection of concomitant medication information at the Day 365 time point.

Subjects who did not complete the study had a visit 28 days, 56 days, and 85 days after the last dose of study medication. Changes in DMARD, steroid or NSAID therapy were not permitted until after assessments were performed on the day of discontinuation. Subjects who completed the study had 2 additional follow-up visits 28 days and 56 days after the Day 365 visit.

3.1.2.2 Primary and Secondary Efficacy Endpoints

The co-primary primary endpoints are as follows:

- 1) Symptomatic relief as measured by ACR 20 response following 6 months of treatment (Day 169).
- 2) Physical function as measured by the disability index of the HAQ at 12 months (Day 365).
- 3) Structural damage as assessed by erosion score using a Genant-modified Sharp scoring method at 12 months of treatment (Day 365).

There are 13 secondary endpoints: total score and joint space narrowing score using a Genant-modified Sharp method at 12 months of treatment (Day 365), ACR 50 and ACR 70 at Day 169 (Month 6), ACR 20, ACR 50, and ACR 70 over time, major clinical response achievement (continuous ACR 70 for 6 months), disease activity measured by Disease Activity Score-28 (DAS 28), at baseline, Day 169, and Day 365, safety of chronic use of abatacept in combination with MTX, the discontinuation rate in subjects receiving abatacept in combination with MTX, the incidence of new tender and swollen joints and 100% reduction in joint counts at Day 169.

3.1.2.3 Statistical Methods

Sponsor's Statistical methods

The intent-to-treat (ITT) analysis population which includes all subjects randomized into the study was used for efficacy analyses by excluding all subjects who were randomized but never received study medication except for the radiographic endpoint analysis.

The ACR 20, ACR 50 and ACR 70 response rates at 6 months between abatacept and placebo were compared with Chi-square test (a continuity correction). Only when the comparison of the ACR 20 response rate was statistically significant, comparison of ACR 50 and ACR 70 response rates was performed sequentially.

Any subject who prematurely discontinued the study after receiving study medication was considered an ACR non-responder. In particular, all subjects who discontinued prior to Month 6 (Day 169) were considered non-responders at Month 6, regardless of the reason of discontinuation.

Any subject for whose data was missing at a given visit not due to premature discontinuation, had ACR response imputed for this visit (current) after observing data from the previous scheduled visit (before) and the next scheduled visit (after). If there are observed positive responses at both visits, a positive response was imputed for the current visit, otherwise the imputed response was negative. If the current visit was the last scheduled efficacy visit (Day 365), then imputation depended on the observed responses at the previous 2 consecutive scheduled efficacy visits. If both responses were positive, the imputed value was positive, otherwise the imputed response was negative.

Mean changes from baseline in the disability index of the HAQ were compared between treatment groups using analysis of covariance (ANCOVA) models with 95% confidence intervals. This analysis was based on the last observation carried forward (LOCF) dataset. For the primary analysis of HAQ response, missing HAQ response was handled similar to the ACR response.

The primary radiological measure was the change from baseline in erosion score using the Genant-modified Sharp method at 12 months. A nonparametric ANCOVA model was used to compare the changes from baseline in Genant-modified Sharp scores (erosion score, joint space narrowing score and total score) at 12 months (Day 365) regardless of any study drug discontinuation.

Secondary measures such as joint space narrowing scores and total scores were to be statistically tested only if the treatment comparison for joint erosion scores was significant. Changes from baseline and proportions of subjects with worsening (i.e. increases from baseline) in the Genant-modified Sharp scores on Day 365 were summarized by treatment with mean, standard deviation, median, and the inter-quartile range.

The radiographic data set for the primary radiographic analyses included all observed data at baseline and at 12 months. Missing annual radiographic data was imputed with linear extrapolation for discontinued subjects based on the baseline value and the on-treatment assessment at the time of discontinuation, provided both assessments were available. In linear extrapolation, the missing Genant-modified Sharp erosion score and joint space narrowing score at Day 365 were imputed first, and the sum of the imputed erosion score and joint space narrowing score were taken as the imputed value for the total score. Subjects with only 1 radiographic film either at baseline, early termination, or 12 months did not have their scores imputed at other time points. These subjects were excluded from the primary analyses.

Sensitivity analyses of ACR and HAQ response rate included the modified worst case sensitivity analysis and the worst case sensitivity analysis. In the modified worst case sensitivity analysis, placebo subjects who discontinued due to lack of efficacy were still counted as ACR or HAQ non-responders at all scheduled protocol visits subsequent to the time of discontinuation. Placebo subjects who did not complete 6 months (Day 169) or 12 months (Day 365) for any reason other than lack of efficacy were classified as to their ACR 20 or HAQ status for the 6 month or 12 month analysis based on the last available data observed at or prior to their discontinuation. If the last observed ACR or HAQ response at or prior to their discontinuation was positive, this subject was counted as an ACR responder for the Day 169 and HAQ responder for the day 365 in the sensitivity analysis. If the last observed ACR or HAQ response at or prior to their discontinuation was negative, the subject was classified as an ACR non-responder at Day 169 and a HAQ non-responder at Day 365. In the worst case sensitivity analysis, placebo subjects discontinued for

any reason were counted as ACR or HAQ responders at all scheduled protocol visits subsequent to the time of discontinuation.

Two secondary analyses were performed for changes from baseline in Genant-modified Sharp erosion score, joint space narrowing score and total score at Day 365. Subjects without baseline data were excluded.

Reviewer's Statistical Methods

For supportive analyses for the primary analyses, the ACR20 and HAQ responses which are repeated binary outcomes over all visits were analyzed using a repeated measures Generalized Estimating Equation (GEE) approach.

For a co-primary efficacy of the change from baseline radiographic erosion score at day 365, the sponsor used all the subjects who had a baseline and a day 365 values. For sensitivity analysis, the reviewer first imputed day 365 values for subjects who had baseline values but missing day 365 values. After imputing missing day 365 values, baseline values those who had day 365 values but missing baseline values were imputed. The baseline values are categorized to 1%, 5%, 10%, 25%, 50%, 75%, 90% and 95% based on total baseline values. The median value of day 365 values were calculated by baseline value category and imputed for missing day 365 values. The missing baseline values were imputed as the same way of the missing 365 day values. A nonparametric ANCOVA model was used for sensitive analyses after imputing missing data.

3.1.2.4 Patient Disposition, Demographic and Baseline Characteristics

Among total of 1250 enrolled subjects, 652 were randomized and treated. Among subjects randomized and treated 433 were randomized to abatacept and 219 to placebo.

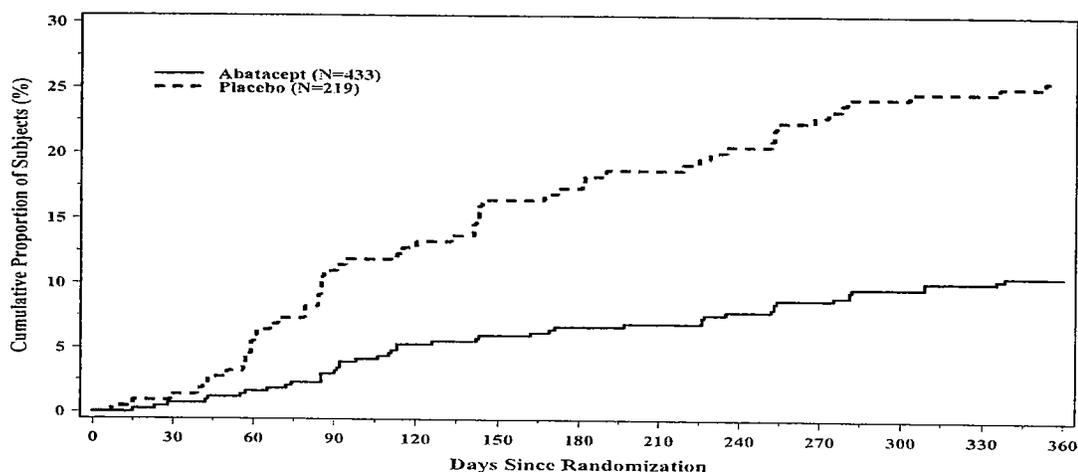
Table 17. Reasons for Discontinuation: Days 1-169 and Days 170-365 (IM101102)

	-----Number (%) of Subjects-----		
	Abatacept (N=433)	Placebo (N=219)	Total (N=652)
Number Discontinued	32 (7.4)	45 (20.5)	77 (11.8)
Death	0	0	0
Adverse Event	11 (2.5)	3 (1.4)	14 (2.1)
Lack of Efficacy	11 (2.5)	33 (15.1)	44 (6.7)
Lost to Follow-up	1 (0.2)	1 (0.5)	2 (0.3)
Withdrawal of Consent	7 (1.6)	4 (1.8)	11 (1.7)
Other	2 (0.5)	4 (1.8)	6 (0.9)
Completed 169 Days of Therapy	401 (92.6)	174 (79.5)	575 (88.2)
Number Discontinued (> Day 169)	16 (3.7)	12 (5.5)	28 (4.3)
Death	1 (0.2)	1 (0.5)	2 (0.3)
Adverse Event	7 (1.6)	1 (0.5)	8 (1.2)
Lack of Efficacy	2 (0.5)	7 (3.2)	9 (1.4)
Lost to Follow-up	0	0	0
Withdrawal of Consent	3 (0.7)	1 (0.5)	4 (0.6)
Other	3 (0.7)	2 (0.9)	5 (0.8)
Completed 365 Days of Therapy	385 (88.9)	162 (74.0)	547 (83.9)

A greater proportion of subjects in the abatacept group (93% and 89%) completed 169 days and 365 days of treatment compared with the placebo group (79% and 74%). Lack of efficacy (15% and 3%) was the most common reason for discontinuation in the placebo group. AEs and lack of efficacy (3%, each) were the most common reasons for discontinuation in the abatacept group up to days 169. Discontinuation due to AEs was the most common reason in the abatacept group (2%) from days 170-365.

Best Possible Copy

Figure 3. Proportion of Subjects Who Discontinued Through Day 365 (All Reasons Combined) (IM101102)



Estimation of discontinuation rates were obtained from Kaplan-Meier analysis.

Table 18. Baseline Demographic Characteristics (IM101102)

		Abatacept N = 433	Placebo N = 219	Total N = 652
Age (years)	N	433	219	652
	Mean	51.5	50.4	51.1
	SD	12.9	12.4	12.7
	Median	53.0	52.0	52.0
	Min	18.0	23.0	18.0
	Max	84.0	87.0	87.0
Weight (Kg)	N	433	219	652
	Mean	72.3	70.2	71.6
	SD	17.5	16.1	17.1
	Median	69.0	67.1	69.0
	Min	37.5	37.5	37.5
	Max	138.6	134.7	138.6
Gender	Male	96 (22.2%)	40 (18.3%)	136 (20.9%)
	Female	337 (77.8%)	179 (81.7%)	516 (79.1%)
Race	White	379 (87.5%)	193 (88.1%)	572 (87.7%)
	Black	10 (2.3%)	4 (1.8%)	14 (2.1%)
	American Indian or Alaska Native	3 (0.7%)	1 (0.5%)	4 (0.6%)
	Asian	18 (4.2%)	10 (4.6%)	28 (4.3%)
	Other	23 (5.3%)	11 (5.0%)	34 (5.2%)
	Geographic Region	North America	93 (21.5%)	46 (21.0%)
	South America	173 (40.0%)	93 (42.5%)	266 (40.8%)
	Europe	143 (33.0%)	67 (30.6%)	210 (32.2%)
	ROW	24 (5.5%)	13 (5.9%)	37 (5.7%)

The majority of subjects were white females, approximately 52 years old.

The baseline clinical RA characteristics and baseline radiographic scores were summarized Tables 19 and 20 for both treatment groups.

Table 19. Baseline Clinical Rheumatoid Arthritis Characteristics (IM101102)

		Abatacept N = 433	Placebo N = 219
Duration of RA (yrs)	N	433	219
	Mean	8.5	8.9
	SD	7.3	7.1
	Median	6.0	7.0
	Min	0.0	1.0
	Max	44.0	35.0
Duration of RA Disease	<= 2 Years	99 (22.9%)	45 (20.5%)
	> 2 to <= 5 Years	93 (21.5%)	46 (21.0%)
	> 5 to <= 10 Years	106 (24.5%)	54 (24.7%)
	> 10 Years	135 (31.2%)	74 (33.8%)
Tender Joints	N	433	219
	Mean	31.0	32.3
	SD	13.2	13.6
	Median	28.0	31.0
	Min	3.0	6.0
	Max	66.0	68.0
Swollen Joints	N	433	219
	Mean	21.4	22.1
	SD	8.8	8.8
	Median	19.0	20.0
	Min	9.0	9.0
	Max	56.0	48.0
Subject Pain Assessment (VAS 100 mm)	N	433	219
	Mean	63.3	65.9
	SD	21.1	20.6
	Median	67.0	70.0
	Min	4.0	3.0
	Max	99.0	99.0
Physical Function (HAQ Disability Index)	N	431	219
	Mean	1.7	1.7
	SD	0.7	0.6
	Median	1.8	1.8
	Min	0.0	0.0
	Max	3.0	3.0
Subject Global Assessment (VAS 100 mm)	N	433	219
	Mean	62.7	62.8
	SD	21.2	21.6
	Median	64.4	64.0
	Min	5.0	3.0
	Max	99.0	97.0
Physician Global Assessment (VAS 100 mm)	N	433	219
	Mean	68.0	67.4
	SD	16.0	17.0
	Median	69.0	68.0
	Min	5.0	12.0
	Max	99.0	99.0
CRP (mg/dL)	N	433	219
	Mean	3.3	2.8
	SD	3.1	2.8
	Median	2.2	2.0
	Min	0.1	0.1
	Max	21.1	12.5
Rheumatoid Factor (IU/mL)	Negative	52 (12.0%)	31 (14.2%)
	Positive	354 (81.8%)	172 (78.5%)
Morning Stiffness (in minutes)	N	432	219
	Mean	97.8	89.6
	SD	60.9	61.2
	Median	90.0	60.0
	Min	0.0	0.0
	Max	180.0	180.0
DAS-28	N	394	198
	Mean	6.8	6.8
	SD	0.9	0.8
	Median	6.8	6.8
	Min	3.7	4.4
	Max	8.7	8.7
MTX (Oral/Parenteral) Dose (mg/wk)	N	433	219
	Mean	16.1	15.7
	SD	3.6	3.5
	Median	15.0	15.0
	Min	7.5	7.5
	Max	30.0	25.0

The demographic and baseline clinical characteristics were similar for both treatment groups. Despite concurrent treatment with MTX with mean doses of approximately 16 mg/week in both treatment groups, subjects in both treatment groups had a high degree of baseline disease activity on the basis of the mean number of tender (31 to 32) and swollen (21 to 22) joints and DAS score (6.8 for both treatment groups). The mean duration of RA was approximately 9 years in both treatment groups. The mean duration of morning stiffness was 98 minutes for the abatacept group compared with 90 minutes for the placebo group.

Table 20. Baseline Radiographic Scores Genant-Modified Sharp Scoring System (IM101102)

		Abatacept N = 433	Placebo N = 219
Erosion Score	n	396	198
	Baseline Median (Range)	16.58 (0.00, 112.2)	16.58 (0.26, 95.82)
Joint Space Narrowing Score	n	396	198
	Baseline Median (Range)	15.86 (0.00, 108.8)	16.53 (0.00, 94.28)
Total Score	n	396	198
	Baseline Median (Range)	31.65 (0.54, 221.0)	33.35 (1.69, 190.1)

No difference was observed in the baseline radiographic scores between the two groups.

3.1.2.5 Results and Conclusions

There were 3 co-primary efficacy endpoints in this study. The first one was symptomatic relief measured by ACR 20 response rate at Day 169, the second one was physical function measured by HAQ response at Day 365, and third one was structural damage assessed change from baseline in erosion score using the Genant-modified Sharp method at Day 365.

Table 21. ACR 20 Responses at Day 169 (IM101102)

		Abatacept N = 424	Placebo N = 214
ACR 20	Number of responders (%)	288 (67.9%)	85 (39.7%)
	Estimate of difference (95% CI)	28.2 (19.8, 36.7)	N/A
	p-value	<0.001 **	N/A

Analysis of the primary efficacy variable for an ACR 20 response at Day 169 was significantly higher in the abatacept group compared with the placebo group (68% vs. 40%, $p < 0.001$).

Table 22. Proportion of Subjects with Clinically Meaningful HAQ Responses at Day 365 (IM101102)

		Abatacept N = 424	Placebo N = 214
HAQ	Number of responders (%)	270 (63.7%)	84 (39.3%)
	Estimate of difference (95% CI)	24.4 (15.9, 32.9)	N/A
	p-value	<0.001 **	N/A

At Day 365, significantly more subjects in the abatacept group compared with the placebo group (64% vs 39%, $p < 0.001$), achieved a HAQ response that was clinically meaningful (defined as an improvement of at least 0.3 units in the HAQ disability index). Also, a statistically significant ($p < 0.001$) treatment difference was also observed at Day 169 for the abatacept group (61%) compared with the placebo group (45%).

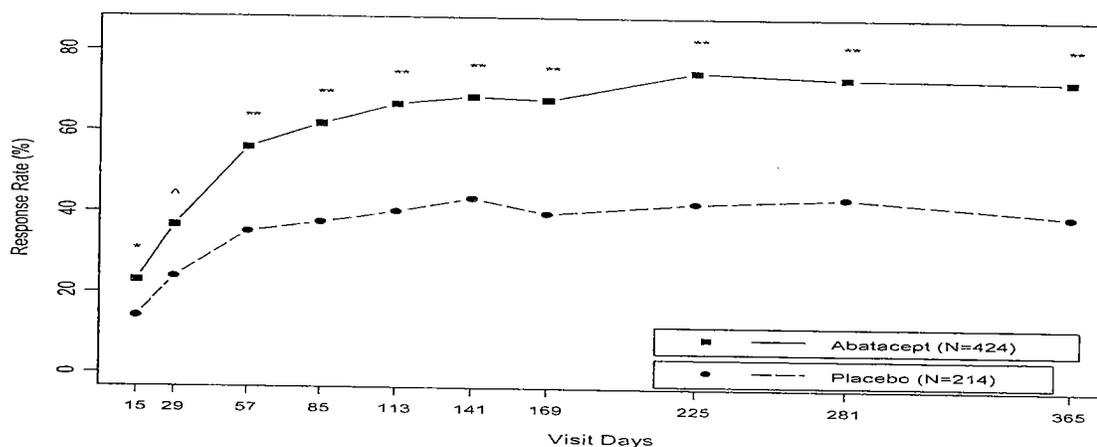
Table 23. Radiographic Erosion Score Results at Day 365 Genant-Modified Sharp Scoring System (IM101102)

		Abatacept N = 424	Placebo N = 214
Erosion Score	n	391	195
	Baseline Mean (SD)	21.68 (18.07)	21.83 (18.63)
	Baseline Median (Range)	16.60 (0.00, 112.2)	16.71 (0.26, 95.82)
	Mean Change from Baseline (SD)	0.63 (1.77)	1.14 (2.81)
	Median Change from Baseline (25%, 75%)	0.00 (0.00, 1.02)	0.27 (0.00, 1.27)
	p-value	0.029 *	N/A

Subjects in the abatacept group had significantly less progression of structural damage compared with the placebo group in the change from baseline in erosion score ($p = 0.029$) as demonstrated by both median and mean changes from baseline.

Sponsor's Secondary Efficacy Analysis

Figure 4: ACR 20 Responses Over Time (IM101102)



** - significant at the 0.001 level, ^ - significant at the 0.01 level; * - significant at the 0.05 level;
One additional non-biologic DMARD could be added on or after Day 169.

Figure 4 shows improvements in ACR 20 responses at Day 15 and continued through Day 365 for the abatacept group compared with the placebo group ($p < 0.001$ at Day 365).

Table 24. ACR 50 and ACR 70 Responses at Day 169 and at Day 365 (IM101102)

ACR 50 and ACR 70 Responses at Day 169			
		Abatacept N = 424	Placebo N = 214
ACR 50	Number of responders (%)	169 (39.9%)	36 (16.8%)
	Estimate of difference (95% CI)	23.0 (15.0, 31.1)	N/A
	p-value	<0.001 **	N/A
ACR 70	Number of responders (%)	84 (19.8%)	14 (6.5%)
	Estimate of difference (95% CI)	13.3 (7.0, 19.5)	N/A
	p-value	<0.001 **	N/A
ACR Responses at Day 365			
		Abatacept N = 424	Placebo N = 214
ACR 20	Number of responders (%)	310 (73.1%)	85 (39.7%)
	Estimate of difference (95% CI)	33.4 (25.1, 41.7)	N/A
	p-value	<0.001 **	N/A
ACR 50	Number of responders (%)	205 (48.3%)	39 (18.2%)
	Estimate of difference (95% CI)	30.1 (21.8, 38.5)	N/A
	p-value	<0.001 **	N/A
ACR 70	Number of responders (%)	122 (28.8%)	13 (6.1%)
	Estimate of difference (95% CI)	22.7 (15.6, 29.8)	N/A
	p-value	<0.001 **	N/A

At Day 365, the abatacept group had a statistically significant (73% vs. 40%, $p < 0.001$) difference in ACR 20 response compared to the placebo group. Additionally, statistically significant differences in ACR 50 and ACR 70 responses were observed at Days 169 and 365 for the abatacept group compared to the placebo group.

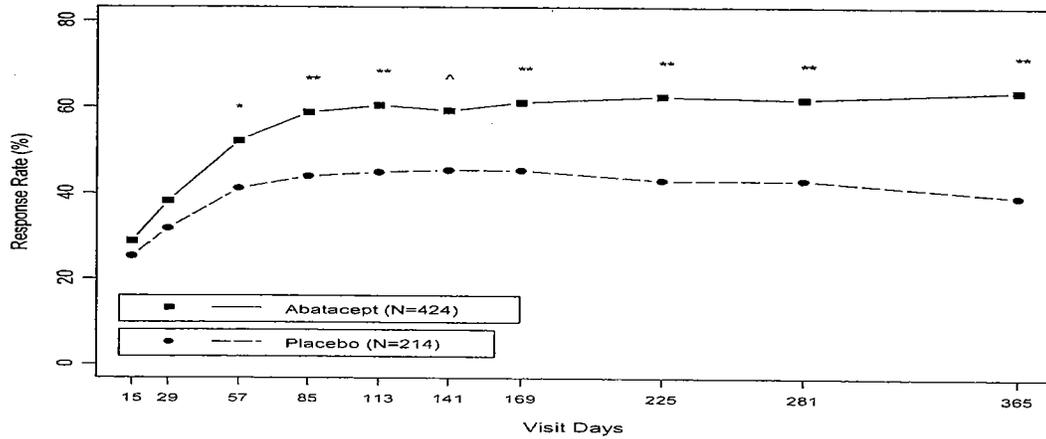
Table 25. Major Clinical Response by Day 365 (IM101102)

		Abatacept N = 424	Placebo N = 214
Major Clinical Response	Number of responders (%)	60 (14.2%)	4 (1.9%)
	Estimate of difference (95% CI)	12.3 (7.3, 17.2)	N/A
	p-value	<0.001 **	N/A

** $p < 0.001$, ^ $p < 0.01$, * $p < 0.05$: probability for testing the difference in major clinical response between abatacept and placebo. One additional non-biologic DMARD could be added on or after Day 169.

A major clinical response (MCR) defined as maintenance of an ACR 70 response over a continuous 6-month period was significantly ($p < 0.001$) more for subjects treated with abatacept (14%) than for subjects treated with placebo (2%) by Day 365.

Figure 5. HAQ Responses Over Time (IM101102)



** - significant at the 0.001 level, ^ - significant at the 0.01 level; * - significant at the 0.05 level;
 One additional non-biologic DMARD could be added on or after Day 169.

Figure 5 shows that there was improvement in the abatacept group (61%) compared to the placebo group (45%) from Day 57 through Day 365.

The mean change from baseline in the HAQ subscales presented at Days 169 and 365 in Table 26.

Table 26. Mean Change from Baseline Up to Days 169 and 365 for the HAQ Disability Index (LOCF Analysis)-IM101102

		Abatacept N = 424	Placebo N = 214
Day 169	n	420	211
	Baseline Mean (SD)	1.69 (0.65)	1.69 (0.58)
	Adjusted Mean Change from Baseline (SE)	-0.59 (0.03)	-0.40 (0.04)
	Comparison with Placebo Diff (95% CI)	-0.19 (-0.29, -0.10)	N/A
	Comparison with Placebo p-value	<0.001 **	N/A
Day 365	n	422	212
	Baseline Mean (SD)	1.69 (0.65)	1.69 (0.58)
	Adjusted Mean Change from Baseline (SE)	-0.66 (0.03)	-0.37 (0.04)
	Comparison with Placebo Diff (95% CI)	-0.29 (-0.38, -0.19)	N/A
	Comparison with Placebo p-value	<0.001 **	N/A

Greater mean reductions from baseline were observed for the HAQ disability index for the abatacept group compared with the placebo group both days 169 and 365 (p <0.001).

Table 27. Radiographic Joint Space Narrowing and Total Score Results at Day 365 Genant-Modified Sharp Scoring System-IM101102

		Abatacept N = 424	Placebo N = 214
Joint Space Narrowing Score	n	391	195
	Baseline Mean (SD)	22.79 (20.16)	23.02 (20.36)
	Baseline Median (Range)	16.23 (0.00, 108.8)	16.58 (0.00, 94.28)
	Mean Change from Baseline (SD)	0.58 (1.54)	1.18 (2.58)
	Median Change from Baseline (25%, 75%)	0.00 (0.00, 0.49)	0.00 (0.00, 0.97)
	p-value	0.009 ^	N/A
Total Score	n	391	195
	Baseline Mean (SD)	44.47 (37.33)	44.85 (37.72)
	Baseline Median (Range)	31.85 (0.54, 221.0)	33.44 (2.26, 190.1)
	Mean Change from Baseline (SD)	1.21 (2.94)	2.32 (5.04)
	Median Change from Baseline (25%, 75%)	0.25 (0.00, 1.78)	0.53 (0.00, 2.54)
	p-value	0.012 *	N/A

One additional non-biologic DMARD could be added on or after Day 169.

**p<0.001, ^p<0.01, *p<0.05: probability for comparison of change in joint space narrowing score between abatacept and placebo.

There were statistically significant treatment differences in joint space narrowing score (p = 0.009) and total score (p = 0.012) for the abatacept group compared with the placebo group.

Reviewer's Results

The generalized estimating equations approach for the analysis of repeated measurements of ACR20 and HAQ_DI responses was used for the supportive analysis of the primary analysis. The GEE model of ACR 20 and HAQ responses included treatment effect and age and weight as covariates with working correlation structure. The result was summarized in Tables 27 and 28.

Table 28. ACR20 Responses Analyzed by GEE (IM101102)

	Estimate	Standard Error	95% CL	Chi-square	p-value
Up to day 169					
Age	-0.0041	0.0020	(-0.0081, -0.0001)	3.92	0.0479
Weight	0.0002	0.0016	(-0.0029, 0.0033)	0.02	0.8940
Abatacept	0.4977	0.0726	(0.3554, 0.6400)	54.1	<.0001
Exp(logOR)*	1.6449	0.1194	(1.4267, 1.8964)	46.99	<.0001
Up to day 365					
Age	-0.0038	0.0018	(-0.0073, -0.0004)	4.48	0.0343
Weight	-0.0002	0.0014	(-0.0030, 0.0026)	0.02	0.8911
Abatacept	0.5215	0.0671	(0.3901, 0.6530)	69.54	<.0001
Exp(LogOR)*	1.6846	0.1130	(1.4771, 1.9213)	60.46	<.0001

*: Exp (log (OR)) compared abatacept to placebo.

The ACR20 response of abatacept group was 1.6 times higher in odds ratio as compared to the placebo group (p<0001) overall time periods up to day 169 and up to day 365.

Table 29. HAQ Responses (>0.3 units) Analyzed by GEE (IM101102)

	Estimate	Standard Error	95% CL	Chi-square	p-value
Up to day 169					
Age	-0.0050	0.0024	(-0.0097, -0.0003)	4.32	0.0377
Weight	0.0007	0.0019	(-0.0030, 0.0044)	0.19	0.6589
Abatacept	0.2587	0.0742	(0.1132, 0.4043)	13.33	0.0003
Exp(logOR)*	1.2953	0.0962	(1.1199, 1.4982)	12.14	0.0005
Up to day 365					
Age	-0.0052	0.0023	(-0.0096 -0.0007)	5.29	0.0214
Weight	0.0004	0.0017	(-0.0030, 0.0039)	0.08	0.7717
Abatacept	0.3061	0.0714	(0.1662, 0.4461)	20.39	<.0001
Exp(LogOR)*	1.3582	0.0970	(1.1808, 1.5622)	18.38	<.0001

*: Exp (log (OR)) compared abatacept to placebo.

The HAQ response of the abatacept group was 1.3 times higher in odds ratio than that of the placebo group (p<.0001) overall time period up to day 169 and up to day 365.

The sensitivity analyses of the change from baseline radiographic erosion score at day 365 was performed by imputing day 365 values for subjects who had only baseline values and imputing missing baseline values for subjects those who had only day 365 values. Table 30 summarizes the sensitivity analyses results of using nonparametric ANCOVA analyses for imputed baseline and day 365 values.

Table 30. Sensitivity Analyses Results for Radiographic Scores (IM101102)

	Abatacept n=424	Placebo n=214	p-value
Imputing day 365 values			
Erosion score			
N	406	204	
Baseline Mean (SD)	21.96 (18.13)	21.62 (18.50)	
Baseline Median (Range)	16.68 (112.2)	16.58 (95.57)	
Mean Change from Baseline (SD)	0.65 (1.93)	1.14 (2.78)	
Median Change from Baseline (25%, 75%)	0.0 (0.0, 1.05)	0.27 (0.0, 1.28)	0.025
JSNS Score			
N	406	204	
Baseline Mean (SD)	22.85 (20.09)	22.66 (20.17)	
Baseline Median (Range)	16.43 (108.8)	16.46 (94.28)	
Mean Change from Baseline (SD)	0.56 (1.61)	1.20 (2.57)	
Median Change from Baseline (25%, 75%)	0.0 (0.0, 0.71)	0.0 (0.0, 1.18)	
Imputing baseline values+day 365 value			
Erosion score			
N	417	210	
Baseline Mean (SD)	21.86 (18.03)	21.57 (18.38)	
Baseline Median (Range)	16.66 (112.2)	16.58 (95.57)	
Mean Change from Baseline (SD)	0.66 (1.95)	1.07 (2.81)	
Median Change from Baseline (25%, 75%)	0.0 (0.0, 1.05)	0.27 (0.0, 1.28)	0.06

JSNS Score		
N	417	210
Baseline Mean (SD)	22.76 (20.07)	22.63 (20.05)
Baseline Median (Range)	16.23 (108.8)	16.46 (94.28)
Mean Change from Baseline (SD)	0.59 (1.70)	1.19 (2.55)
Median Change from Baseline (25%, 75%)	0.0 (0.0, 0.72)	0.0 (0.0, 1.20)
Total score		
N	417	210
Baseline Mean (SD)	44.62 (37.18)	44.20 (37.17)
Baseline Median (Range)	32.12 (220.5)	33.24 (188.6)
Mean Change from Baseline (SD)	1.14 (2.86)	2.16 (4.90)
Median Change from Baseline (25%, 75%)	0.0 (0.0, 1.55)	0.29 (0.0, 2.23)

The number of subjects who had baseline values but missing for day 365 values were 15 and 9 for abatacept and placebo, respectively. After imputing these 24 subjects' day 365 values with the median of the day 365 values calculated by the category of baseline values, the results were robust. However, after imputing baseline values for those who had only day 365 values using the same imputation method with the above, the results show that there was no statistically significant difference in the co-primary efficacy of the change from baseline radiographic scores and the results were not robust.

3.1.3 Study IM101029

3.1.3.1 Study Design

Study IM101029 was a randomized, double-blind, placebo-controlled, parallel-dosing design study, with a treatment period of 6 months. Subjects with active RA who met the inclusion/exclusion criteria for this study were randomized 2:1 ratio of the following treatments on a background of DMARDs or anakinra therapy: abatacept or placebo.

Subjects must have been treated with background DMARDs or anakinra for at least 3 months and at a stable dose for 28 days prior to first study treatment (Day 1). Subjects continued to receive background DMARDs or anakinra during this study at the dose level and regimen administered at the time of randomization. To minimize potential methotrexate (MTX) toxicity, all subjects being treated with MTX received either folic acid or leucovorin according to manufacturer recommendations.

Subjects received a dose of abatacept based on their screening visit weight. Subjects weighing <60 kg received 500 mg, subjects weighing 60 kg to 100 kg received 750 mg and subjects weighing >100 kg received 1 gram. Study medication was administered on Days 1, 15, 29, and every 28 days thereafter, for a total of 7 doses.

Only subjects with active RA who had failed etanercept or infliximab for inadequate efficacy were eligible for this study. Subjects who failed anti-TNF therapy for other reasons, such as toxicity, intolerability, or lack of availability were not eligible.

3.1.3.2 Primary and Secondary Efficacy Endpoints

The co-primary primary endpoints are as follows:

- 1) Symptomatic relief as measured by ACR 20 response following 6 months of treatment

(Day 169).

2) Physical function as measured by the disability index of the HAQ at 6 months (Day 169).

There are 13 secondary endpoints: ACR 50 and ACR 70 at Day 169 (Month 6), ACR 20, ACR 50, and ACR 70 over time, major clinical response achievement (continuous ACR 70 for 6 months), disease activity measured by Disease Activity Score-28 (DAS 28) at Day 169, safety of chronic use of abatacept in combination with MTX, the discontinuation rate in subjects receiving abatacept in combination with MTX, the incidence of new tender and swollen joints and 100% reduction in joint counts at Day 169.

3.1.3.3 Statistical methods

Sponsor's Statistical Analysis

All efficacy analyses were performed on the intent-to-treat population (ITT), which consisted of all randomized subjects who received at least one dose of study medication (primary efficacy dataset). This dataset was categorized based on the randomized assignment to study medication.

Other datasets were derived from the primary efficacy dataset, including the last observation carried forward (LOCF) dataset and worst case sensitivity data set.

The protocol called for a secondary efficacy dataset to be used for analyses (per protocol analysis) only if more than 10% of subjects within any treatment group had significant protocol violations at the primary analysis time point (Day 169). This dataset was not required since fewer than 10% of the subjects had significant protocol violations in any treatment group.

The imputation methods of the missing data for ACR and HAQ responses were the similar with the Study IM101102. Any subject who prematurely discontinued the study after receiving study medication was considered an ACR non-responder at all scheduled protocol visits subsequent to the point of discontinuation. In particular, all subjects who discontinued prior to Month 6 (Day 169) were considered non-responders at Month 6, regardless of the reason of discontinuation.

For the primary endpoints, the designation of baseline anti-TNF status (current or prior) for use as covariates in the Cochran-Mantel Haenszel (CMH) tests was based on the stratified randomization schedule of the IVRS.

A sequential testing procedure was employed to test for differences in the ACR 20 response rates and HAQ responses between the abatacept and the placebo groups at Day 169 as the similar way of Study IM101102.

In order to assess the true impact of being either a current or a prior anti-TNF user, a sensitivity analysis of the primary efficacy endpoints, ACR 20 and HAQ Disability Index, was also carried out using the baseline anti-TNF status (current or prior) from the CRF data. Since subject randomization was stratified on the anti-TNF user status (current or prior use), differences in the ACR 20 and HAQ responses between the abatacept group and the placebo group were assessed within each stratum. In addition to summarizing the treatment effect among the prior and current users, although not pre-specified, post-hoc statistical comparisons between treatment groups were also made. These comparisons were done to assess whether the overall observed treatment benefit was similar among the prior and current users. For the analysis within the strata (current or prior anti-TNF use), data categorizations obtained from the CRF data were used.

For the co-primary analysis of HAQ response, the proportion of subjects achieving HAQ responses (a reduction from baseline in their HAQ score of at least 0.3 unit) at Day 169 were summarized by treatment group. A two-sided Cochran-Mantel-Haenszel (CMH) Chi square test (with stratification based on baseline anti-TNF use (current or prior)) was used to compare the HAQ response data between the abatacept group and the placebo group at the 0.05 level of significance. For the HAQ responder analysis, all subjects who discontinued were considered as HAQ non-responders subsequent to their discontinuation with the similar method of Study IM101102.

Mean changes from baseline in HAQ disability index are compared between treatment groups using analysis of covariance (ANCOVA) models and 95% confidence intervals are presented. This analysis was based on the LOCF data set.

Reviewer's Statistical Methods

For supportive analyses for the primary analyses, the ACR20 and HAQ responses which are repeated binary outcomes over all visits were analyzed using a repeated measures Generalized Estimating Equation (GEE) approach with the stratification of TNF use.

3.1.3.4 Patient Disposition, Demographic and Baseline Characteristics

A total of 738 subjects were enrolled and 391 subjects were randomized and treated.

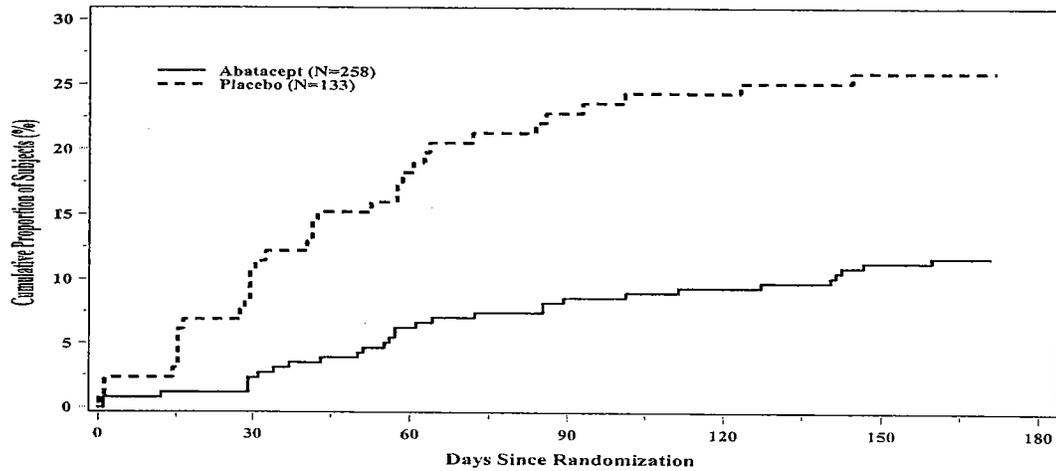
Table 31. Reasons for Discontinuation during Double-Blind Therapy (IM101029)

	-----Number (%) of Subjects-----		
	Abatacept (N=258)	Placebo (N=133)	Total (N=391)
Number Discontinued	35 (13.6)	34 (25.6)	69 (17.6)
Death	0	0	0
Adverse Event	9 (3.5)	5 (3.8)	14 (3.6)
Lack of Efficacy	14 (5.4)	27 (20.3)	41 (10.5)
Lost to Follow-up	5 (1.9)	0	5 (1.3)
Withdrawal of Consent	5 (1.9)	2 (1.5)	7 (1.8)
Other	2 (0.8)	0	2 (0.5)
Completed 169 Days of Therapy	223 (86.4)	99 (74.4)	322 (82.4)

The proportion of subjects completed 169 days of treatment was 86% in the abatacept group as compared with 74% of the placebo group. Lack of efficacy (20%) and AEs (4%) were the most common reasons for discontinuation in the placebo group. Lack of efficacy (5%) and AEs (4%) were the most common reasons for discontinuation in the abatacept group. Figure 6 summarizes discontinuation rates for any reason.

*Appears This Way
On Original*

Figure 6. Proportion of Subjects who Discontinued During Double-Blind Therapy (All Reasons Combined) Cumulative Proportion of Subjects (%) -IM101029



Estimation of discontinuation rates were obtained from Kaplan-Meier analysis.

The demographic and baseline clinical characteristics were similar between the two treatment groups as shown Tables 32 and 33.

Table 32. Baseline Demographic Characteristics (IM101029)

		Abatacept N = 258	Placebo N = 133	Total N = 391
Age (years)	N	258	133	391
	Mean	53.4	52.7	53.2
	SD	12.4	11.3	12.1
	Median	55.0	53.0	54.0
	Min	21.0	22.0	21.0
	Max	81.0	77.0	81.0
Weight (kg)	N	257	132	389
	Mean	78.2	78.2	78.2
	SD	19.0	21.0	19.7
	Median	77.0	74.4	76.2
	Min	33.0	43.5	33.0
	Max	145.6	158.8	158.8
Gender	Male	59 (22.9%)	27 (20.3%)	86 (22.0%)
	Female	199 (77.1%)	106 (79.7%)	305 (78.0%)
Race	White	248 (96.1%)	124 (93.2%)	372 (95.1%)
	Black	9 (3.5%)	5 (3.8%)	14 (3.6%)
	American Indian or Alaska Native	1 (0.4%)	1 (0.8%)	2 (0.5%)
	Asian	0	2 (1.5%)	2 (0.5%)
	Other	0	1 (0.8%)	1 (0.3%)
Geographic Region	North America	189 (73.3%)	99 (74.4%)	288 (73.7%)
	South America	0	0	0
	Europe	69 (26.7%)	34 (25.6%)	103 (26.3%)

Table 32 shows that the majority of subjects were Caucasian, females, and approximately 53 years of age, and >70% of the subjects in this study were enrolled from North America.

Table 33. Baseline Clinical Rheumatoid Arthritis Characteristics (IM101029)

		Abatacept N = 258	Placebo N = 133
Duration of RA (yrs)	N	258	133
	Mean	12.2	11.4
	SD	8.5	8.9
	Median	11.0	10.0
	Min	1.0	0.0
	Max	43.0	44.0
Duration of RA Disease	<= 2 Years	32 (12.4%)	16 (12.0%)
	> 2 to <= 5 Years	31 (12.0%)	26 (19.5%)
	> 5 to <= 10 Years	59 (22.9%)	25 (18.8%)
	> 10 Years	136 (52.7%)	66 (49.6%)
Tender Joints	N	258	133
	Mean	31.2	32.8
	SD	13.0	13.4
	Median	30.0	32.0
	Min	4.0	12.0
	Max	67.0	67.0
Swollen Joints	N	258	133
	Mean	22.3	22.0
	SD	10.2	10.0
	Median	21.0	20.0
	Min	6.0	7.0
	Max	62.0	61.0
Subject Pain Assessment (VAS 100 mm)	N	255	133
	Mean	70.8	69.9
	SD	19.8	19.0
	Median	73.0	74.0
	Min	5.0	9.0
	Max	100.0	100.0
Physical Function (HAQ Disability Index)	N	258	133
	Mean	1.8	1.8
	SD	0.6	0.6
	Median	1.9	2.0
	Min	0.0	0.0
	Max	3.0	2.9
Subject Global Assessment (VAS 100 mm)	N	255	133
	Mean	69.2	69.7
	SD	19.7	20.3
	Median	71.0	74.0
	Min	8.0	6.0
	Max	100.0	100.0
Physician Global Assessment (VAS 100 mm)	N	258	132
	Mean	68.8	67.3
	SD	17.7	16.8
	Median	70.5	69.5
	Min	13.0	16.0
	Max	100.0	100.0
CRP (mg/dL)	N	258	133
	Mean	4.6	4.0
	SD	4.0	3.6
	Median	3.3	2.9
	Min	0.1	0.3
	Max	24.9	15.1
Rheumatoid Factor (IU/mL)	Negative	52 (20.2%)	30 (22.6%)
	Positive	189 (73.3%)	97 (72.9%)
Morning Stiffness (in minutes)	N	258	133
	Mean	121.1	115.3
	SD	61.5	60.8
	Median	120.0	120.0
	Min	0.0	5.0
	Max	180.0	180.0
DAS-28	N	222	109
	Mean	6.9	6.9
	SD	1.0	1.0
	Median	7.0	7.0
	Min	3.9	2.6
	Max	8.9	8.7
MTX (Oral/Parenteral) Dose (mg/wk)	N	194	109
	Mean	15.2	14.4
	SD	5.3	6.1
	Median	15.0	15.0
	Min	0.5	1.0
	Max	35.0	28.0
Current/Prior Anti-TNF Status	Current	98 (38.0%)	55 (41.4%)
	Prior	160 (62.0%)	78 (58.6%)

Baseline clinical RA characteristics were similar for both treatment groups. Despite treatment with DMARDs, disease activity was high in both treatment groups, as determined by the mean number of tender joints (≥ 31), mean number of swollen joints (22), mean physical score (HAQ Disability Index) measurements of 1.8, mean levels of CRP of ≥ 4.0 mg/dL, and mean DAS score of 6.9. The mean duration of RA was approximately 12 years. The mean duration of morning stiffness was 121 minutes for the abatacept group compared with 115 minutes for the placebo group.

Table 34. Etanercept and Infliximab User Status at Enrollment (IM101029)

	Abatacept (N=258)		Placebo (N=133)	
	Etanercept	Infliximab	Etanercept	Infliximab
Current Anti-TNF Users	32 (12.4%)	66 (25.6%)	23 (17.3%)	32 (24.1%)
Prior Anti-TNF Users	51 (19.8%)	109 (42.2%)	30 (22.6%)	48 (36.1%)
Total	83 (32.2%)	175 (67.8%)	53 (39.8%)	80 (60.2%)

The percentage of subjects who were current anti-TNF users was similar between the abatacept and placebo groups (38% vs. 41%, respectively). Also, the percentage of subjects who were prior anti-TNF users was similar between the abatacept and placebo groups (62% vs. 59% respectively).

Within each treatment group, there were more subjects who had received infliximab: abatacept group, 175 subjects (68%); placebo group, 80 subjects (60%) compared with those who had received etanercept: abatacept group, 83 subjects (32%); placebo group, 53 subjects (40%).

Table 35. Duration of Etanercept and Infliximab Use Prior to Randomization (IM101029)

	Abatacept (N=258)		Placebo (N=133)	
	Etanercept (N= 32)	Infliximab (N= 66)	Etanercept (N= 23)	Infliximab (N= 32)
Current Anti-TNF Users				
<3 Months	3 (9.4%)	3 (4.5%)	1 (4.3%)	3 (9.4%)
3-8 Months	10 (31.3%)	14 (21.2%)	9 (39.1%)	3 (9.4%)
>8 Months	19 (59.4%)	49 (74.2%)	13 (56.5%)	26 (81.3%)
Prior Anti-TNF Users				
<3 Months	1 (2.0%)	6 (5.5%)	0	0
3-8 Months	20 (39.2%)	36 (33.0%)	8 (26.7%)	22 (45.8%)
>8 Months	30 (58.8%)	67 (61.5%)	22 (73.3%)	26 (54.2%)
Median Days since Discontinuation	213.0	182.0	163.0	196.5

The majority of subjects in this study had received etanercept or infliximab for more than 8 months prior to discontinuation. Few subjects (17 subjects; 4%) had received etanercept or infliximab for less than 3 months. The baseline demographic and disease characteristics of each stratified group were similar to the total population.

3.1.3.5 Results and Conclusions

Sponsor's Primary Efficacy Analyses Results

There were two sequential co-primary efficacy objectives in this study. The first one was improvement of signs and symptoms of RA, as measured by ACR 20 response rate at Day 169 and the second one was improvement of physical function, as measured by HAQ response rate at Day 169.

First Co-Primary Efficacy Objective: ACR 20 Response at Day 169

Table 36. ACR 20 Responses at Day 169 (IM101029)

		Abatacept N = 256	Placebo N = 133
ACR 20	Number of responders (%)	129 (50.4%)	26 (19.5%)
	Est of Weighted Diff. (95% CI)	30.8 (20.6, 41.1)	N/A
	p-value	<0.001 **	N/A

Differences represent the weighted average of the individual stratum differences between the two groups
 **p<0.001, ^p<0.01, *p<0.05: probability for testing the difference in ACR response between abatacept and placebo.

The proportion of subjects with an ACR 20 response at Day 169 was significantly higher in the abatacept group compared with the placebo group (50% vs. 20%, p < 0.001).

Table 37. ACR 20 Responses by Baseline Anti-TNF Use Current vs. Prior (IM101029)

Baseline Anti-TNF Use-Current		Abatacept (N = 97)	Placebo (N = 55)
Day 169	Number of responders (%)	44 (45.4%)	8 (14.5%)
	Estimate of difference (95% CI)	30.8 (13.7, 47.9)	N/A
	p-value	<0.001 **	N/A
Baseline Anti-TNF Use-Prior		(N = 159)	(N = 78)
Day 169	Number of responders (%)	85 (53.5%)	18 (23.1%)
	Estimate of difference (95% CI)	30.4 (16.0, 44.8)	N/A
	p-value	<0.001 **	N/A

**p<0.001, ^p<0.01, *p<0.05: probability for testing the difference in ACR response between abatacept and placebo.

Subjects were stratified according to whether they had inadequate efficacy while receiving anti-TNF therapy at the time of enrollment (current anti-TNF user status) or whether they had previously failed anti-TNF therapy for inadequate efficacy (prior anti-TNF user status). There was statistically higher ACR 20 response in the abatacept group as compared with the placebo group in both strata (45% vs. 15%, p < 0.001 for current anti-TNF users; 54% vs. 23%, p < 0.001 for prior anti-TNF users). The results were consistent in both subjects who had received etanercept and subjects who had received infliximab.

**Second Co-Primary Efficacy Objective: Physical Function (HAQ Disability Index)
Response at Day 169**

Table 38. Proportion of Subjects with Clinically Meaningful HAQ Responses at Day 169 (IM101029)

		Abatacept N = 256	Placebo N = 133
HAQ	Number of responders (%)	121 (47.3%)	31 (23.3%)
	Est of Weighted Diff. (95% CI)	24.0 (13.8, 34.2)	N/A
	p-value	<0.001 **	N/A

Differences represent the weighted average of the individual stratum differences between the two groups
 **p<0.001, ^p<0.01, *p<0.05: probability for testing the difference in HAQ response between abatacept and placebo using CMH Test.

Significantly (p < 0.001) more subjects treated with abatacept (47%) achieved clinically meaningful improvement in physical function (defined by an improvement of 0.3 units from baseline in the HAQ Disability Index score) at Day 169 than subjects in the placebo group (23%).

Table 39. HAQ Response by Baseline Anti-TNF Use - Current vs Prior (IM101029)

Baseline Anti-TNF Use - Current		Abatacept (N = 97)	Placebo (N = 55)
Day 169	Number of responders (%)	42 (43.3%)	12 (21.8%)
	Estimate of difference (95% CI)	21.5 (4.2, 38.7)	N/A
	p-value	0.013 *	N/A
Baseline Anti-TNF Use - Prior		(N = 159)	(N = 78)
Day 169	Number of responders (%)	79 (49.7%)	19 (24.4%)
	Estimate of difference (95% CI)	25.3 (11.0, 39.6)	N/A
	p-value	<0.001 **	N/A

There was a statistically higher HAQ response in the abatacept group compared with the placebo group in both strata (43% vs. 22%, p = 0.013 for current anti-TNF users; 50% vs. 24%, p < 0.001 for prior anti-TNF users). The results were consistent for both subjects who had received etanercept and for subjects who had received infliximab.

Table 40. Mean Change from Baseline Up to Day 169 for the HAQ Disability Index (LOCF Analysis)-IM101029

		Abatacept N = 256	Placebo N = 133
Day 169	n	249	130
	Baseline Mean (SD)	1.83 (0.58)	1.82 (0.60)
	Adjusted Mean Change from Baseline (SE)	-0.45 (0.03)	-0.11 (0.04)
	Comparison with Placebo Diff (95% CI)	-0.34 (-0.44, -0.23)	N/A
	Comparison with Placebo p-value	<0.001 **	N/A

Population: All randomized and treated subjects.
 **p<0.001, ^p<0.01, *p<0.05: probability for testing the difference in mean change in HAQ scores between abatacept and placebo.

Greater mean changes from baseline up to Day 169 were observed for the HAQ disability index and all its 8 subscales for the abatacept group compared with the placebo group.

Sponsor's Secondary Efficacy Analysis

Table 41. ACR 50 and ACR 70 Responses at Day 169 (IM101029)

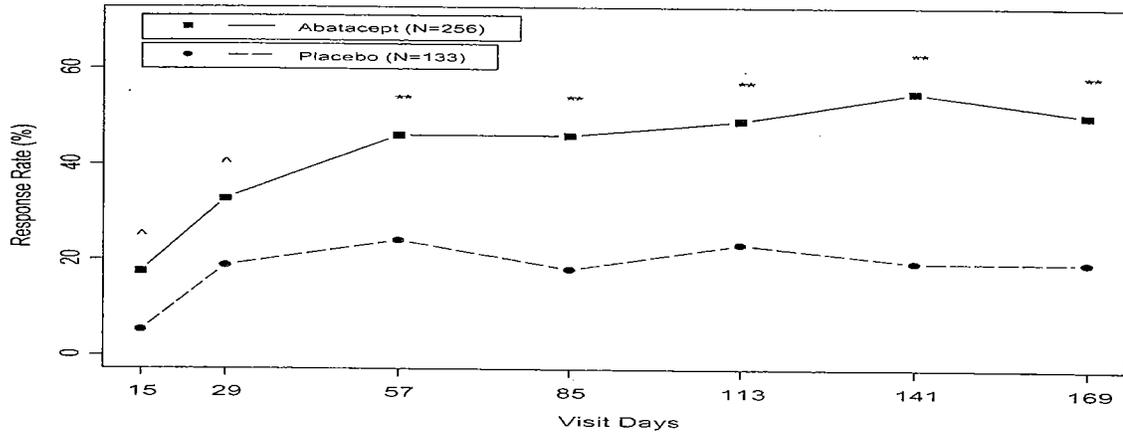
		Abatacept N = 256	Placebo N = 133
ACR 50	Number of responders (%)	52 (20.3%)	5 (3.8%)
	Estimate of difference (95% CI)	16.6 (8.6, 24.5)	N/A
	p-value	<0.001 **	N/A
ACR 70	Number of responders (%)	26 (10.2%)	2 (1.5%)
	Estimate of difference (95% CI)	8.7 (2.7, 14.6)	N/A
	p-value	0.003 ^	N/A

Population: All randomized and treated subjects.

**p<0.001, ^p<0.01, *p<0.05: probability for testing the difference in ACR response between abatacept and placebo.

At Day 169, an ACR 50 and ACR70 responses were achieved by significantly more subjects in the abatacept group compared with the placebo group (20% vs. 4%, p < 0.001, 10% vs. 2%, p = 0.003, respectively).

Figure 7. ACR 20 Responses Over Time (IM101029)



*-significant at the 0.05 level; ^-significant at the 0.01 level; **-significant at the 0.001 level

Greater proportions of subjects with ACR 20 responses in abatacept group were seen as early as the first measurement from Day 15 as compared with the placebo group. The difference between two groups was also statistically significant at Day 169 (p < 0.001).

Based on median percent change at Day 169, subjects treated with abatacept had greater improvements in all individual ACR components than subjects treated with placebo.

Table 42. ACR Core Components - Median Percent Change from Baseline on Day 169 (LOCF Analysis)-IM101029

		Abatacept N = 256	Placebo N = 133
Tender Joints	n	254	130
	Baseline Median	30.00	31.00
	Post-Baseline Median	13.00	23.50
	Median % Improvement from Baseline	56.67	22.22
	% Improvement from Baseline Percentile (25th, 75th)	(25.49, 80.00)	(-7.69, 52.38)
Swollen Joints	n	254	130
	Baseline Median	21.00	20.00
	Post-Baseline Median	10.00	14.00
	Median % Improvement from Baseline	52.19	30.15
	% Improvement from Baseline Percentile (25th, 75th)	(20.83, 76.67)	(0.00, 60.00)
Subject Pain Assessment (VAS 100 mm)	n	251	130
	Baseline Median	73.00	73.50
	Post-Baseline Median	43.00	64.00
	Median % Improvement from Baseline	37.84	5.75
	% Improvement from Baseline Percentile (25th, 75th)	(10.96, 69.89)	(-18.9, 28.87)
Physical Function (HAQ Disability Index)	n	248	128
	Baseline Median	1.88	2.00
	Post-Baseline Median	1.38	1.75
	Median % Improvement from Baseline	20.53	6.25
	% Improvement from Baseline Percentile (25th, 75th)	(0.00, 41.67)	(-7.69, 18.18)
Subject Global Assessment (VAS 100 mm)	n	251	130
	Baseline Median	71.00	72.50
	Post-Baseline Median	44.00	62.50
	Median % Improvement from Baseline	36.67	11.03
	% Improvement from Baseline Percentile (25th, 75th)	(5.26, 69.12)	(-7.46, 32.50)
Physician Global Assessment (VAS 100 mm)	n	254	129
	Baseline Median	70.50	69.00
	Post-Baseline Median	32.00	54.00
	Median % Improvement from Baseline	51.46	19.18
	% Improvement from Baseline Percentile (25th, 75th)	(20.45, 72.22)	(-1.06, 53.33)
CRP	n	254	131
	Baseline Median	3.35	2.80
	Post-Baseline Median	1.30	2.30
	Median % Improvement from Baseline	56.00	15.87
	% Improvement from Baseline Percentile (25th, 75th)	(12.50, 80.65)	(-44.4, 42.50)

Improvements in the DAS 28 score at Day 169 were consistent with the improvements assessed by the ACR variables.

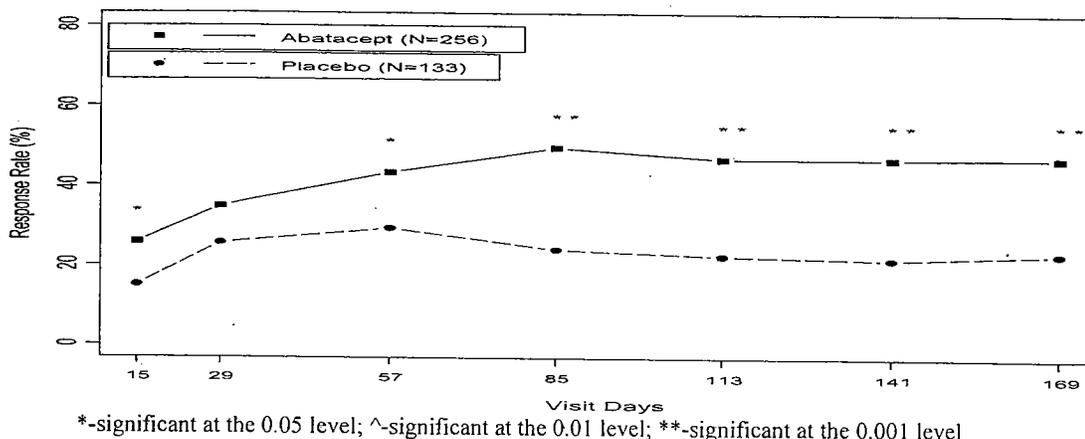
Table 43. Mean Change from Baseline in DAS 28(ESR) at Day 169 (LOCF Analysis)-IM101029

		Abatacept N = 256	Placebo N = 133
Day 169	n	182	98
	Baseline Mean (SD)	6.88 (0.99)	6.88 (0.92)
	Post-Baseline Mean (SD)	4.90 (1.55)	6.17 (1.34)
	Adjusted Mean Change from Baseline (SE)	-1.98 (0.10)	-0.71 (0.14)
	Difference from Placebo (95% CI)	-1.27 (-1.62, -0.93)	N/A
	Comparison with Placebo p-value	<0.001 **	N/A
	Subjects with Improvement (DAS 28 change >= 1.2)	129 (70.9)	31 (31.6)
	Subjects with Low Disease Activity (DAS 28 <= 3.2)	30 (16.5)	4 (4.1)
	Subjects in Remission (DAS 28 < 2.6)	19 (10.4)	1 (1.0)

The DAS 28 score was calculated two ways using ESR and using CRP. Significant ($p < 0.001$) improvements were observed in disease activity as measured by the DAS 28 (ESR) at Day 169 for the abatacept group (adjusted mean change from baseline (SE) of -1.98 [0.10]) compared with the placebo group (adjusted mean change from baseline (SE) of -0.71 [0.14]). Clinically meaningful

improvement (defined as change ≥ 1.2 in the DAS 28 score by EULAR response criteria) was observed in a greater proportion of the abatacept group (71%) compared with the placebo group.

Figure 8. HAQ Responses Over Time (IM101029)



Greater clinically meaningful improvement was achieved as early as the first measurement on Day 15 for the abatacept group compared with the placebo group.

In the modified worst case scenario sensitivity analysis, all subjects who discontinued in the group for any reason other than lack of efficacy were assigned the same ACR response at Day 169 and HAQ response at Day 365 as their last ACR 20 and HAQ responses just prior to discontinuation. All placebo subjects who discontinued due to lack of efficacy were considered ACR20 or HAQ non-responders subsequent to discontinuation. All subjects treated with abatacept who discontinued for any reason were considered ACR 20 or HAQ non-responders subsequent to discontinuation. These results were consistent with the ACR 20 and HAQ response rates in the primary analysis.

Reviewer's Analysis

The ACR20 was analyzed using generalized estimating equation approach by stratified TNF current of prior user. The GEE model included age, weight, TNF user and interaction between treatment and TNF user. The result was summarized in Table 44.

Table 44. ACR 20 Response Results Using GEE Approach (IM101029)

	Estimate	Standard Error	95% CL	Chi-square	p-value
Up to day 169					
Age	0.0047	0.0057	(-0.0064, 0.0158)	0.69	0.4046
Weight	-0.0006	0.0035	(-0.0075, 0.0063)	0.03	0.8655
TNF Current	0.1606	0.3146	(-0.4560, 0.7771)	0.41	0.5198
Interaction#	-0.0828	0.3605	(-0.7892, 0.6237)	0.05	0.8196
Abatacept	1.1770	0.1893	(0.8059, 1.5481)	38.05	<.0001
Exp(logOR)*	3.1132	0.5622	(2.1852, 4.4353)	39.55	<.0001

#: interaction between treatment group and TNF use

*:Exp(log(OR))

The ACR20 response of abatacept group was three times higher than that of the placebo group overall time period (p<.0001) after adjusting for stratification of TNF use and interaction between treatment group and TNF use including age and weight as covariates.

Table 45. HAQ Response Results Using GEE Approach (IM101029)

	Estimate	Standard Error	95% CL	Chi-square	p-value
Up to day 169					
Age	-0.0029	0.0042	(-0.0112, 0.0053)	0.45	0.5010
Weight	0.0014	0.0024	(-0.0033, 0.0061)	0.43	0.5126
TNF Current	-0.0017	0.2440	(-0.4798, 0.4765)	0.00	0.9912
Interaction#	0.0003	0.2750	(-0.5387, 0.5392)	0.00	0.9992
Abatacept	0.6127	0.1526	(0.3136, 0.9118)	21.09	<.0001
Exp(logOR)*	1.8457	0.2535	(1.4102, 2.4158)	19.52	<.0001

#: interaction between treatment group and TNF use

*:Exp(log(OR))

The HAQ response of abatacept group was 1.8 times higher in odds ratio than that of the placebo group overall time period (p<.0001) after adjusting for stratification of TNF use and interaction between treatment group and TNF use including age and weight as covariates.

3.1.4 Study IM101031

3.1.4.1 Study Design

Study IM101031 was a randomized, double-blind, placebo-controlled, parallel-dosing design study with a treatment period of 12 months. Male or female subjects at least 18 years old with active RA, with or without co-morbidities and currently treated with background non-biologic DMARDs and/or biologic therapy(ies) approved for RA, were eligible for participation.

Background RA therapy refers to the regimen of non-biologic or biologic agents used by the subject for treatment of RA at enrollment.

Prior to study participation, subjects must have been treated with non-biologic RA therapy, biologic RA therapy, or their combination for 3 months and at a stable regimen for 28 days prior to Day 1. Subjects were maintained on their background RA therapy at the dose(s) they were receiving at the time of randomization. Stable, low-dose oral corticosteroids (10 mg/day or less) and/or stable doses of non-steroidal anti-inflammatory drugs (NSAIDs) including aspirin (ASA) were allowed. In subjects receiving oral corticosteroids, a stable dose (maximum of 10 mg prednisone equivalent daily) must have been maintained for at least 25 of the 28 days prior to Day 1.

Subjects were randomized 2:1 to either of the following treatment groups:

1) All Abatacept: Includes all subjects randomized to abatacept fixed dose regardless of background RA therapy (any non-biologic RA therapy, biologic RA therapy, or their combination)

2) All Placebo: Includes all subjects randomized to placebo (dextrose 5% in water [D5W] or Normal Saline [NS]) regardless of background RA therapy (any nonbiologic RA therapy, biologic RA therapy, or their combination)

3.1.4.2 Primary and Secondary Efficacy Endpoints

The primary objective was to summarize the incidence of adverse events (AEs), serious adverse events (SAEs), and discontinuations due to AEs during 1 year of combined treatment with abatacept and 1 or more of the DMARDs and/or biologics approved for RA in subjects with active RA with or without co-morbid medical conditions.

Exploratory efficacy measures in this study consisted of select ACR core components of the Subject Pain Assessment, the Subject Global Assessment of Disease Activity, the Physician's Global Assessment of Disease Activity, and the disability index of HAQ.

3.1.4.3 Statistical Methods

Descriptive statistics were provided for each of these measures. For the assessments of changes from baseline at a given time point, data for any subject who had a baseline evaluation and an on-therapy evaluation at the specified time point were included in the analyses for the ITT population (except for subjects from Site 151, which were excluded from all efficacy analyses). If a subject discontinued prematurely, only assessments occurring within 42 days of the last infusion dose were included in the efficacy analyses.

3.1.4.4 Patient Deposition and Baseline Characteristics

Of 1795 subjects who were enrolled in this study, 1441 subjects were randomized and treated: 959 to the abatacept group and 482 to the placebo group.

Table 46. Reasons for Discontinuation: Days 1 – 365 (IM101031)

	-----Number (%) of Subjects-----		
	All Abatacept (N=959)	All Placebo (N=482)	Total (N=1441)
Number Discontinued	123 (12.8)	87 (18.0)	210 (14.6)
Death	5 (0.5)	3 (0.6)	8 (0.6)
Adverse Event	51 (5.3)	19 (3.9)	70 (4.9)
Lack of Efficacy	26 (2.7)	44 (9.1)	70 (4.9)
Lost to Follow-up	3 (0.3)	4 (0.8)	7 (0.5)
Withdrawal of Consent	24 (2.5)	10 (2.1)	34 (2.4)
Other	14 (1.5)	7 (1.5)	21 (1.5)
Completed 365 Days of Therapy	836 (87.2)	395 (82.0)	1231 (85.4)

Population: All randomized and treated subjects.

A greater proportion of subjects in the abatacept group (87%) completed 365 days of treatment compared with the placebo group (82%). Subjects completing the double-blind period could continue in the open-label period, which is ongoing.

From Days 1 to 365, a greater proportion of subjects in the placebo group discontinued the study compared with the abatacept group (18% and 13%, respectively). Lack of efficacy and AEs were the most common reasons for discontinuation (Table 45). More subjects in the placebo group (9%) discontinued for lack of efficacy compared with the abatacept group (3%). Adverse events led to discontinuation for 5% of subjects in the abatacept group and 4% of subjects in the placebo group.

Reasons for discontinuation in subjects treated on a background of non-biologic vs biologic RA therapy are summarized in Table 47.

Table 47. Reasons for Discontinuation: Days 1 - 365: Background of Non-biologic vs. Biologic RA Therapy (IM101031)

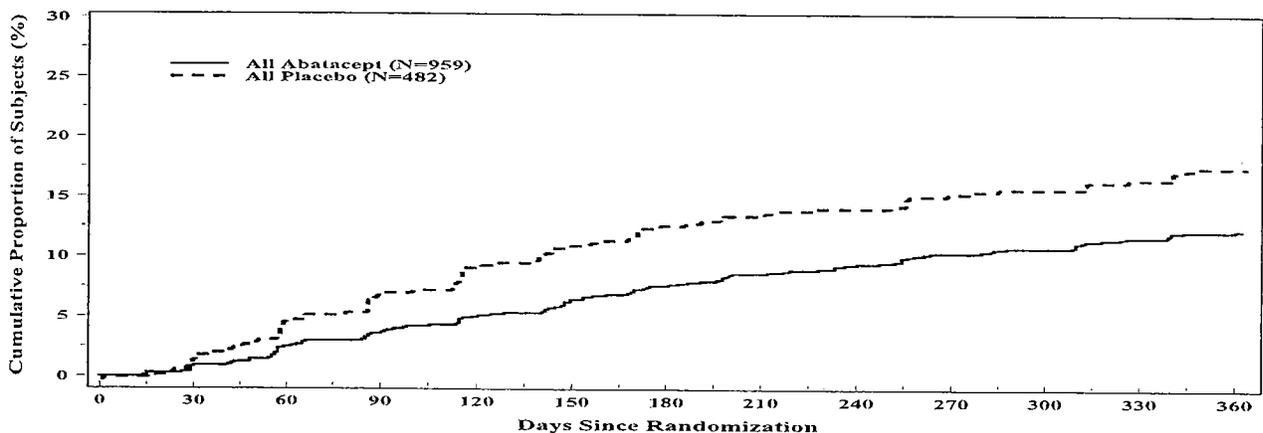
	-----Number (%) of Subjects-----				Total (N=1441)
	Abatacept + Non-Biologic RA Therapy (N=856)	Placebo + Non-Biologic RA Therapy (N=418)	Abatacept Biologic RA Therapy (N=103)	Placebo + Biologic RA Therapy (N=64)	
Number Discontinued	102 (11.9)	67 (16.0)	21 (20.4)	20 (31.3)	210 (14.6)
Death	5 (0.6)	3 (0.7)	0	0	8 (0.6)
Adverse Event	42 (4.9)	17 (4.1)	9 (8.7)	2 (3.1)	70 (4.9)
Lack of Efficacy	18 (2.1)	30 (7.2)	8 (7.8)	14 (21.9)	70 (4.9)
Lost to Follow-up	3 (0.4)	4 (1.0)	0	0	7 (0.5)
Withdrawal of Consent	21 (2.5)	8 (1.9)	3 (2.9)	2 (3.1)	34 (2.4)
Other	13 (1.5)	5 (1.2)	1 (1.0)	2 (3.1)	21 (1.5)
Completed 365 Days of Therapy	754 (88.1)	351 (84.0)	82 (79.6)	44 (68.8)	1231 (85.4)

Of the 959 subjects assigned to the abatacept treatment arm, 856 (89%) were receiving background non-biologic RA therapy and 103 (11%) were receiving background biologic RA therapy. Corresponding numbers of subjects assigned to the placebo group who were receiving background RA therapy with non-biologic or biologic drugs were 418 (87%) and 64 (13%), respectively.

The relative discontinuation rates for abatacept- and placebo-treated subjects in the Non-Biologics RA Therapy subgroups were similar to those for the overall population since the majority (>85%) of the overall study population were receiving background non-biologic RA therapy. For the Biologic RA Therapy subgroups, more subjects receiving placebo discontinued for lack of efficacy (22%) compared with those receiving abatacept (8%). The discontinuation rate for AEs was 9% in the abatacept + Biologic RA Therapy subgroup and 3% in the placebo + Biologic RA Therapy subgroup.

Figure 9 summarizes the cumulative proportion of subjects that discontinued for any reason from Day 1 to 365.

Figure 9. Proportion of Subjects who Discontinued Through Day 365 (All Reasons Combined) (IM101031)



Demographic characteristics were similar for the abatacept and placebo treatment groups.

Table 48. Baseline Demographic Characteristics All Abatacept (IM101031)

		All Abatacept N = 959	All Placebo N = 482	Total N = 1441
Age (years)	N	959	482	1441
	Mean	52.4	52.1	52.3
	SD	11.7	12.0	11.8
	Median	52.0	52.0	52.0
	Min	19.0	18.0	18.0
	Max	79.0	87.0	87.0
Weight (Kg)	N	959	482	1441
	Mean	71.3	72.9	71.9
	SD	18.9	19.8	19.2
	Median	68.2	69.1	68.6
	Min	36.0	36.0	36.0
	Max	159.5	153.8	159.5
Gender	Male	170 (17.7%)	84 (17.4%)	254 (17.6%)
	Female	789 (82.3%)	398 (82.6%)	1187 (82.4%)
Race	White	818 (85.3%)	407 (84.4%)	1225 (85.0%)
	Black	49 (5.1%)	29 (6.0%)	78 (5.4%)
	American Indian or Alaska Native	1 (0.1%)	1 (0.2%)	2 (0.1%)
	Asian	76 (7.9%)	41 (8.5%)	117 (8.1%)
	Native Hawaiian or Other Pacific Islander	1 (0.1%)	0	1 (0.1%)
	Other	14 (1.5%)	4 (0.8%)	18 (1.2%)
Geographic Region	North America	411 (42.9%)	204 (42.3%)	615 (42.7%)
	South America	262 (27.3%)	133 (27.6%)	395 (27.4%)
	Europe	187 (19.5%)	94 (19.5%)	281 (19.5%)
	ROW	99 (10.3%)	51 (10.6%)	150 (10.4%)

The majority of subjects were Caucasian, female, approximately 52 years of age. The geographic distribution of subjects was comparable in the 2 treatment groups; 43% of subjects were enrolled at sites in North America (mainly in US), 27% at sites in South America, 20% at sites in Europe, and 10% at sites in the rest of the world.

Baseline clinical characteristics also were similar for the abatacept and placebo treatment group in Table 49.

Appears This Way
On Original

Table 49. Baseline Clinical Rheumatoid Arthritis Characteristics (IM101031)

		All Abatacept N = 959	All Placebo N = 482	Total N = 1441
Duration of RA (yrs)	N	959	481	1440
	Mean	9.7	9.8	9.7
	SD	8.7	9.2	8.9
	Median	7.0	7.0	7.0
	Min	0.0	0.0	0.0
	Max	50.0	43.0	50.0
Duration of RA Disease	<= 2 Years	214 (22.3%)	103 (21.4%)	317 (22.0%)
	> 2 to <= 5 Years	182 (19.0%)	103 (21.4%)	285 (19.8%)
	> 5 to <= 10 Years	198 (20.6%)	115 (23.9%)	313 (21.7%)
	> 10 Years	365 (38.1%)	160 (33.2%)	525 (36.4%)
Subject Pain (VAS 100 mm)	N	955	481	1436
	Mean	61.3	61.3	61.3
	SD	20.4	20.7	20.5
	Median	61.0	65.0	62.0
	Min	2.0	2.0	2.0
	Max	100.0	100.0	100.0
Physical Function (HAQ)	N	955	480	1435
	Mean	1.5	1.5	1.5
	SD	0.6	0.6	0.6
	Median	1.5	1.6	1.5
	Min	0.0	0.0	0.0
	Max	3.0	3.0	3.0
Subject Global (VAS 100 mm)	N	955	481	1436
	Mean	60.6	60.9	60.7
	SD	19.7	19.7	19.7
	Median	60.0	61.0	61.0
	Min	0.0	6.0	0.0
	Max	100.0	100.0	100.0
Physician Global (VAS 100 mm)	N	949	480	1429
	Mean	57.8	58.2	57.9
	SD	17.4	17.5	17.5
	Median	58.0	60.0	58.0
	Min	5.0	11.0	5.0
	Max	98.0	98.0	98.0
CRP (mg/dL)	N	950	477	1427
	Mean	1.8	2.0	1.9
	SD	2.4	2.5	2.4
	Median	0.9	1.1	0.9
	Min	0.0	0.1	0.0
	Max	17.8	21.0	21.0

Subjects had a mean duration of RA of approximately 10 years. Mean subject pain, subject global, and physician global VAS scores for the entire study population were 61.3, 60.7, and 57.9, respectively, and the mean HAQ disability index was 1.5.

Baseline demographic and clinical characteristics are summarized by background RA therapy (non-biologic vs. biologic) in Table 50.

Appears This Way
On Original

Table 50. Baseline Demographic Characteristics by Background of Non-Biologic or Biologic RA Therapy (IM101031)

		Aba+NonBio N = 859	Placebo+NonBio N = 423	Aba+Bio N = 100	Placebo+Bio N = 59	Total N = 1441
Age (years)	N	859	423	100	59	1441
	Mean	52.2	51.9	54.5	52.9	52.3
	SD	11.7	12.0	11.3	11.6	11.8
	Median	52.0	52.0	54.0	56.0	52.0
	Min	19.0	22.0	25.0	18.0	18.0
	Max	79.0	87.0	78.0	69.0	87.0
Weight (Kg)	N	859	423	100	59	1441
	Mean	70.3	72.0	80.5	79.0	71.9
	SD	18.4	19.5	20.0	20.6	19.2
	Median	68.0	68.0	75.3	77.0	68.6
	Min	36.0	36.0	45.4	47.8	36.0
	Max	159.5	153.8	141.5	137.0	159.5
Gender	Male	148 (17.2%)	70 (16.5%)	22 (22.0%)	14 (23.7%)	254 (17.6%)
	Female	711 (82.8%)	353 (83.5%)	78 (78.0%)	45 (76.3%)	1187 (82.4%)
Race	White	721 (83.9%)	353 (83.5%)	97 (97.0%)	54 (91.5%)	1225 (85.0%)
	Black	46 (5.4%)	26 (6.1%)	3 (3.0%)	3 (5.1%)	78 (5.4%)
	Amer Indian or AK Native	1 (0.1%)	1 (0.2%)	0	0	2 (0.1%)
	Asian	76 (8.8%)	40 (9.5%)	0	1 (1.7%)	117 (8.1%)
	Native HI/Pacific Isl	1 (0.1%)	0	0	0	1 (0.1%)
	Other	14 (1.6%)	3 (0.7%)	0	1 (1.7%)	18 (1.2%)
	Geo Region	North America	327 (38.1%)	155 (36.6%)	84 (84.0%)	49 (83.1%)
	South America	261 (30.4%)	132 (31.2%)	1 (1.0%)	1 (1.7%)	395 (27.4%)
	Europe	173 (20.1%)	86 (20.3%)	14 (14.0%)	8 (13.6%)	281 (19.5%)
	ROW	98 (11.4%)	50 (11.8%)	1 (1.0%)	1 (1.7%)	150 (10.4%)

Abbreviations: Amer=American, AK=Alaskan, HI=Hawaiian, Isl=Islander, Geo=Geographic.

Aba+NonBio represents the treatment group: Abatacept + Non-Biologic RA Therapy. Placebo+NonBio represents the treatment group: Placebo + Non-Biologic RA Therapy. Aba+Bio represents the treatment group: Abatacept + Biologic RA Therapy. Placebo+Bio represents the treatment group: Placebo + Biologic RA Therapy.

Appears This Way
On Original

Table 51. Baseline Clinical Rheumatoid Arthritis Characteristics by Background of Non-Biologic or Biologic RA Therapy (IM101031)

		Aba+NonBio N = 859	Placebo+NonBio N = 423	Aba+Bio N = 100	Placebo+Bio N = 59	Total N = 1441
Dur RA (yrs)	N	859	422	100	59	1440
	Mean	9.5	9.5	11.5	11.3	9.7
	SD	8.7	9.1	8.9	9.7	8.9
	Median	7.0	6.0	9.5	8.0	7.0
	Min	0.0	0.0	0.0	0.0	0.0
	Max	50.0	42.0	45.0	43.0	50.0
Dur RA Disease	<= 2 Years	205 (23.9%)	97 (22.9%)	9 (9.0%)	6 (10.2%)	317 (22.0%)
	> 2 to <= 5 Years	159 (18.5%)	90 (21.3%)	23 (23.0%)	13 (22.0%)	285 (19.8%)
	> 5 to <= 10 Years	176 (20.5%)	99 (23.4%)	22 (22.0%)	16 (27.1%)	313 (21.7%)
	> 10 Years	319 (37.1%)	136 (32.2%)	46 (46.0%)	24 (40.7%)	525 (36.4%)
Subject Pain	N	856	422	99	59	1436
	Mean	61.2	61.4	61.9	61.0	61.3
	SD	20.4	20.7	20.4	20.5	20.5
	Median	61.0	65.0	62.0	61.0	62.0
	Min	2.0	8.0	5.0	2.0	2.0
	Max	100.0	100.0	100.0	99.0	100.0
Phys Func (HAQ)	N	857	421	98	59	1435
	Mean	1.5	1.5	1.5	1.5	1.5
	SD	0.6	0.7	0.6	0.6	0.6
	Median	1.5	1.5	1.5	1.8	1.5
	Min	0.0	0.0	0.0	0.3	0.0
	Max	2.9	3.0	3.0	2.8	3.0
Subject Global	N	857	422	98	59	1436
	Mean	60.6	61.2	60.0	58.8	60.7
	SD	19.7	20.1	19.8	17.1	19.7
	Median	60.0	62.0	58.5	60.0	61.0
	Min	6.0	6.0	0.0	17.0	6.0
	Max	100.0	100.0	100.0	86.0	100.0
Physician Global	N	850	421	99	59	1429
	Mean	57.8	58.2	57.4	57.9	57.9
	SD	17.4	17.5	18.1	18.2	17.5
	Median	58.0	60.0	58.0	58.0	58.0
	Min	5.0	11.0	16.0	19.0	5.0
	Max	98.0	98.0	92.0	90.0	98.0
CRP (mg/dL)	N	850	419	100	58	1427
	Mean	1.9	2.1	1.4	1.4	1.9
	SD	2.4	2.6	1.9	1.7	2.4
	Median	0.9	1.2	0.7	0.9	0.9
	Min	0.0	0.1	0.0	0.1	0.0
	Max	17.8	21.0	13.3	9.5	21.0

Aba+NonBio represents the treatment group: Abatacept + Non-Biologic RA Therapy. Placebo+NonBio represents the treatment group: Placebo + Non-Biologic RA Therapy. Aba+Bio represents the treatment group: Abatacept + Biologic RA Therapy. Placebo+Bio represents the treatment group: Placebo + Biologic RA Therapy. Subject Pain, Subject Global and Physician Global are measured by VAS (100 mm).

Relative to the Non-biologics RA Therapy subgroup, subjects in the Biologics RA Therapy subgroup tended to have a longer mean duration of RA (approximately 10 years vs 11 to 12 years), were somewhat older (mean age of approximately 52 years vs 53 to 55 years), and had a higher mean body weight (approximately 70 to 72 kg vs 79 to 81 kg). The mean baseline CRP level was lower for those who receiving background biologic RA therapy (1.4 mg/dL) compared with those who receiving background non-biologic RA therapy (approximately 2 mg/dL).

3.1.4.5 Results and Conclusions

Four exploratory efficacy measures were assessed prior to dosing on Days 1 (baseline), 85, 169, 253, and 365. These measures were subject global pain assessment, subject global assessment of disease activity, physician global assessment of disease activity (all 3 measured by VAS), and physical function as assessed by the HAQ disability index. Site 151 (Dr. Roth) were excluded from all efficacy analyses due to compliance irregularities.

Table 52. Median Percent Improvement from Baseline in Select ACR Core Components on Day 365 (IM101031)

		All Abatacept N = 948	All Placebo N = 477
Subject Pain Assessment (VAS 100 mm)	n	823	393
	Baseline Median	61.00	64.00
	Post-Baseline Median	29.00	44.00
	Median % Improvement from Baseline	47.50	26.00
	% Improvement from Baseline Percentile (25th,75th)	(12.96, 75.82)	(0.00, 56.36)
Physical Function (HAQ Disability Index)	n	813	381
	Baseline Median	1.50	1.50
	Post-Baseline Median	1.00	1.38
	Median % Improvement from Baseline	28.57	14.29
	% Improvement from Baseline Percentile (25th,75th)	(5.56, 60.00)	(-6.25, 42.86)
Subject Global Assessment (VAS 100 mm)	n	822	393
	Baseline Median	60.00	61.00
	Post-Baseline Median	30.00	42.00
	Median % Improvement from Baseline	47.14	30.38
	% Improvement from Baseline Percentile (25th,75th)	(13.56, 76.06)	(-3.09, 56.90)
Physician Global Assessment (VAS 100 mm)	n	820	390
	Baseline Median	58.00	59.00
	Post-Baseline Median	21.00	31.00
	Median % Improvement from Baseline	63.28	43.02
	% Improvement from Baseline Percentile (25th,75th)	(39.16, 83.33)	(16.00, 67.39)

% Improvement from Baseline = (Baseline - Post-baseline value) / baseline value x 100.

For abatacept-treated subjects, median percent improvements in efficacy measures (subject assessment of pain, subject global assessment, physician global assessment, subject assessment of physical function) at Day 365 were numerically larger than those observed in the placebo-treated group.

Table 53 summarizes the mean percent improvement from baseline with corresponding 95% CIs at Day 365 for each of the four disease outcome variables.

Table 53. Mean Percent Improvement from Baseline in Select ACR Core Components on Day 365 (IM101031)

		All Abatacept N = 948	All Placebo N = 477
Subject Pain Assessment (VAS 100 mm)	n	823	393
	Baseline Mean (SD)	60.96 (20.32)	60.77 (20.70)
	Post-Baseline Mean (SD)	34.58 (24.80)	44.40 (25.01)
	Mean % Improvement from Baseline (SE)	36.88 (2.43)	18.98 (3.10)
	95% CI	(32.12, 41.64)	(12.89, 25.08)
Physical Function (HAQ Disability Index)	n	813	381
	Baseline Mean (SD)	1.51 (0.61)	1.55 (0.61)
	Post-Baseline Mean (SD)	1.04 (0.68)	1.28 (0.71)
	Mean % Improvement from Baseline (SE)	29.41 (1.67)	9.71 (4.84)
	95% CI	(26.14, 32.69)	(0.19, 19.23)
Subject Global Assessment (VAS 100 mm)	n	822	393
	Baseline Mean (SD)	60.18 (19.53)	60.49 (19.87)
	Post-Baseline Mean (SD)	33.03 (23.06)	42.91 (24.20)
	Mean % Improvement from Baseline (SE)	40.68 (1.58)	21.30 (3.11)
	95% CI	(37.58, 43.78)	(15.18, 27.42)
Physician Global Assessment (VAS 100 mm)	n	820	390
	Baseline Mean (SD)	57.46 (17.26)	57.58 (17.35)
	Post-Baseline Mean (SD)	23.95 (18.31)	33.87 (21.01)
	Mean % Improvement from Baseline (SE)	54.96 (1.55)	36.66 (2.51)
	95% CI	(51.92, 58.00)	(31.72, 41.59)

Population: All randomized and treated subjects.

% Improvement from Baseline = (Baseline - Post-baseline value) / baseline value x 100.

Mean percent improvements for each of these variables were numerically larger in the abatacept group than those observed in the placebo-treated group. The 95% CIs were higher in the abatacept group compared with the placebo group without overlap for each of the variables.

Table 54. Mean Change from Baseline for the HAQ Disability Index by Subscale and Visit
HAQ Disability Index (IM101031)

		All Abatacept N = 948	All Placebo N = 477
Day 85	n	911	456
	Baseline Mean (SD)	1.50 (0.63)	1.52 (0.65)
	Post-Baseline Mean (SD)	1.14 (0.68)	1.35 (0.67)
	Mean Change from Baseline (SE)	-0.36 (0.02)	-0.17 (0.02)
	95% CI	(-0.39, -0.32)	(-0.21, -0.12)
Day 169	n	867	417
	Baseline Mean (SD)	1.50 (0.63)	1.51 (0.65)
	Post-Baseline Mean (SD)	1.07 (0.69)	1.29 (0.70)
	Mean Change from Baseline (SE)	-0.42 (0.02)	-0.22 (0.03)
	95% CI	(-0.46, -0.39)	(-0.27, -0.17)
Day 253	n	841	398
	Baseline Mean (SD)	1.49 (0.63)	1.52 (0.65)
	Post-Baseline Mean (SD)	1.03 (0.68)	1.27 (0.70)
	Mean Change from Baseline (SE)	-0.46 (0.02)	-0.24 (0.03)
	95% CI	(-0.50, -0.42)	(-0.30, -0.19)
Day 365	n	824	390
	Baseline Mean (SD)	1.49 (0.63)	1.51 (0.65)
	Post-Baseline Mean (SD)	1.03 (0.69)	1.26 (0.72)
	Mean Change from Baseline (SE)	-0.46 (0.02)	-0.26 (0.03)
	95% CI	(-0.50, -0.42)	(-0.32, -0.20)

Mean changes from baseline were larger in the abatacept group than in the placebo group for the Disability Index of the HAQ and each of its subscales.

3.2 Evaluation of Safety

In the double-blind period of the 5 core RA studies, 1765 subjects were treated with the recommended dose of abatacept (10 mg/kg or a fixed dose that approximated 10 mg/kg), representing 1527.4 person-years of exposure. In the open-label period, all 2285 subjects were exposed to the recommended dose of abatacept, representing 1093.5 person-years of exposure. Safety data from an additional double-blind Phase II study (IM103002) in subjects with RA (N=122) is discussed separately in the RA section and not integrated into the core RA data set for 2 reasons: the treatment period was 3 months and subjects in this study were treated with abatacept or placebo monotherapy. Other supportive safety information includes a) a completed Phase I study in healthy subjects (N=30), b) 4 studies (3 Phase I studies and 1 Phase II study) in subjects with psoriasis (N=177) c) a Phase II study in subjects with multiple sclerosis (MS) (N=128), d) 3 ongoing RA studies (2 Phase I and 1 Phase II study), and 2 ongoing Phase I studies in healthy subjects.

The following is a sponsor's summary of the key safety findings from the core RA studies:

Double-Blind Controlled Study Periods

- The proportion of subjects in the abatacept and placebo groups who died (0.5% vs 0.6%, respectively) or reported serious adverse events (SAEs) (13.6% vs 12.3%) were similar for the overall RA population in the double-blind periods.
- Discontinuations due to SAEs (2.7% vs 1.6%), and discontinuations due to adverse events (AEs) (5.4% vs 3.9%) were slightly higher in the abatacept group compared with the placebo group.

- The most commonly reported AEs in the abatacept and placebo groups were headache (18.2% vs 12.6%), upper respiratory tract infection (12.7% vs 12.0%), nausea (11.5% vs 10.6%), and nasopharyngitis (11.5% vs 9.1%). The majority of these AEs were mild to moderate in severity.
- There were small increases in the overall incidences of infection (53.8% vs 48.3%) and serious infections (3.0% vs 1.9%) in the abatacept group compared with the placebo group, but there was no increase in the proportion of subjects discontinuing due to an infection (1.2% vs 1.0%). There was no increase in the frequency of opportunistic infections and no difference in the medical seriousness or severity of infections observed in both groups.
- Malignant neoplasms were reported in a similar proportion of subjects in the abatacept and placebo groups, (1.2% vs 1.0%). However, given the low event rate, a potential risk for individual tumor types cannot be excluded based on relatively limited sample size and short observation period in clinical trials.
- Safety data from the double-blind periods does not suggest that use of abatacept is associated with an increased risk of developing medically important autoimmune disorders.
- Overall, the peri-infusional tolerability profile of abatacept was excellent, even though the protocols did not require or recommend pretreatment for potential allergic-type reactions.
- Analysis of the safety data from the small subset of subjects receiving abatacept or placebo on a background of biologic RA therapy indicate that the risk of infection is higher in abatacept-treated subjects receiving concomitant biologic RA therapy. Concurrent therapy with abatacept and another biologic RA therapy is not recommended. While transitioning from a biologic RA therapy to abatacept, subjects should be monitored carefully for signs of infection.
- The frequencies and types of AEs and SAEs did not vary substantially between intrinsic and extrinsic factor subgroups, with the exception of SAEs in subjects > 65 years of age (abatacept 24.1% vs placebo 16.2%). While the overall frequency of SAEs was higher, relatively few subjects > 65 years of age reported any individual SAE and there were no medically significant events that would limit the use of abatacept in older patients. Overall, these analyses did not reveal any clinically significant concerns in any intrinsic or extrinsic factor subgroup.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race and Age

Subgroup analyses of key efficacy measures (ACR response, HAQ response and change from baseline erosion score) were performed for: age, gender and race at Day 169. The reviewer used chi-square test to compare between the two groups. If the number is less than 10 then Fisher's exact p values were used. The reviewer used Vander Waerden test (two sided) for the change from baseline erosion score between the two groups. The results were summarized in Tables 55, 56, and 57 for Study IM101102 and in Tables 58 and 59 for Study IM101029.

4.1.1 Study IM101102

Table 55. Subgroup Analysis of ACR 20 Response by baseline demographics (IM101102)

	(1)MTX n=214	(2)Abatacept n=424	p-values
Age			
<65	77/188(41.0)	248/359(69.1)	<0.0001
≥65	8/26 (30.8)	40/65 (61.5)	<0.0001
Gender			
Female	71/176(40.3)	217/331(65.6)	<0.0001
Male	14/38 (36.8)	71/93 (76.3)	<0.0001
Race			
White	71/189(37.6)	249/370(67.3)	<0.0001
Black	2/4 (50)	6/10 (60)	1.000
Asian	4/10 (40)	12/18 (66.7)	0.2425
Other	8/11 (71.7)	21/26 (80.8)	0.6720
Weight (Kg)			
<60	17/57 (29.8)	69/107(64.5)	<0.0001
≥60-≤100	65/145(44.8)	203/288(70.5)	<0.0001
>100	3/12 (25.0)	16/29 (55.2)	0.0977

Table 56. Subgroup Analysis of HAQ Response by baseline demographics (IM101102)

	(1)MTX n=214	(2)Abatacept n=424	p-values
Age			
<65	77/188(40.4)	234/359(65.2)	<0.0001
≥65	8/26 (30.8)	36/65 (55.4)	0.0393
Gender			
Female	71/176(40.3)	210/331(63.4)	<0.0001
Male	13/38 (34.2)	60/93 (64.5)	0.0015
Race			
White	72/189(38.1)	241/370(65.1)	<0.0001
Black	3/4 (75.0)	3/10 (30.0)	0.2448
Asian	4/10 (40.0)	10/18 (55.6)	0.6946
Other	5/11 (45.5)	16/26 (61.5)	0.4753
Weight (Kg)			
<60	21/57 (36.8)	65/107(60.8)	0.0035
≥60-≤100	62/145(42.8)	193/288(67.0)	<0.0001
>100	1/12 (8.3)	12/29 (55.2)	0.0642

Table 57. Subgroup Analysis of the change from baseline of erosion score baseline demographics (IM101102)

	(1)MTX (n=195)			(2)Abatacept (n=391)			p-values
	Mean (SD)	Median	n	Mean (SD)	Median	n	
Age							
<65	1.24 (2.94)	0.27	174	0.71 (1.79)	0.00	330	0.0347
≥65	0.31 (0.82)	0.26	21	0.18 (1.62)	0.00	61	0.4544
Gender							
Female	1.20 (2.92)	0.27	161	0.66 (1.88)	0.00	311	0.0233
Male	0.86 (2.19)	0.27	34	0.51 (1.25)	0.00	80	0.4827
Race							
White	1.15 (2.93)	0.27	170	0.63 (1.84)	0.00	341	0.0210
Black	2.34 (3.05)	1.26	4	0.35 (0.40)	0.26	6	0.0958
Asian	0.54 (0.73)	0.00	10	1.01 (1.52)	0.52	18	0.5134
Other	0.99 (1.93)	0.00	11	0.42 (1.10)	0.00	26	0.5858
Weight (Kg)							
<60	0.77 (1.88)	0.27	51	0.74 (1.96)	0.00	104	0.4852
≥60-≤100	1.39 (3.17)	0.27	132	0.56 (1.65)	0.00	259	0.0041
>100	-0.04 (0.77)	0.00	12	0.90 (2.12)	0.50	28	0.2948

Because the majority of subjects were white females, < 65 years olds, and weighed between 60 and 100kg, efficacy in these groups showed statistically higher responses in the ACR20 and HAQ and less progression of change from baseline erosion score in the abatacept group compared to the placebo group. The efficacy in other subgroups could not be well established due to limited data.

4.1.2 Study IM101029

Table 58. Subgroup Analysis of ACR 20 Response by baseline demographics (IM101029)

	(1)MTX n=133	(2)Adalimumab n=256	p-values
Age			
<65	24/113(21.2)	101/205(49.3)	<0.0001
≥65	2/20 (10.0)	28/51 (54.9)	0.0005
Gender			
Female	23/106(21.7)	96/198(48.5)	<0.0001
Male	3/27 (36.8)	33/58 (56.9)	<0.0001
Race			
White	24/124(19.4)	126/246(51.2)	<0.0001
Black	2/5 (40.0)	3/9 (33.3)	1.000
Asian	0/2	0/0	
Other	0/2	0/1	
Weight (Kg)			
<60	3/25 (12.0)	20/45 (44.4)	0.0074
≥60-≤100	19/89 (21.4)	91/177(51.4)	<0.0001
>100	4/19 (21.1)	18/34 (52.9)	0.0408

Table 59. Subgroup Analysis of HAQ Response by baseline demographics (IM101029)

	(1)MTX n=133	(2)Adalimumab n=256	p-values
Age			
<65	27/113(23.9)	98/205(47.8)	<0.0001
≥65	4/20 (20.0)	23/51 (45.1)	0.0609
Gender			
Female	28/106(26.4)	89/198 (45.0)	0.0016
Male	3/27 (11.1)	32/58 (55.2)	0.0001
Race			
White	28/124(22.6)	119/246(48.4)	<0.0001
Black	0/5	2/9 (22.2)	0.5055
Asian	1/2 (50.0)	0/0	
Other	2/2	0/1	
Weight (Kg)			
<60	5/25 (20.0)	16/45 (35.6)	0.2761
≥60-≤100	21/89 (23.6)	89/177(50.3)	<0.0001
>100	5/19 (26.3)	16/34 (47.1)	0.1581

The subgroup analyses results were similar with the Study IM101102. Because the majority of subjects were white females, < 65 years olds, and weighed between 60 and 100kg, efficacy in these groups showed statistically higher responses in the ACR20 and HAQ. The efficacy in other subgroups could not be well established due to limited data.

4.2 Other Special/Subgroup Populations

Subgroup analyses of key efficacy measures (ACR response; HAQ response and change from baseline erosion score at day 365 (IM101102 only)) were performed for: duration of RA, baseline CRP, swollen joint count, tender joint count, and baseline HAQ disability index for Study IM101102 and Study IM101029. The results are summarized in Tables 60, 61, and 62 for Study IM101102 and in Tables 63 and 64 for Study IM101029.

Appears This Way
On Original

4.2.1 Study IM101102

Table 60. Subgroup Analysis of ACR20 Response by Baseline Disease Activity (IM101102)

	(1)MTX n=214	(2)Adalimumab n=424	p-values
Disease Duration			
1.0-≤2.0	19/41 (46.3)	67/95 (70.5)	0.0073
2.0-≤5.0	20/46 (43.5)	64/91 (70.3)	0.0023
5.0-≤10.0	22/54 (40.7)	72/105(68.6)	0.0007
>10	24/73 (32.9)	85/133(63.9)	<0.0001
Baseline CRP			
≤12	25/66 (37.9)	65/108(60.2)	0.0043
>12-<40	38/99 (38.4)	136/197(69.0)	<0.0001
≥40	22/49 (44.9)	87/119(73.1)	0.0005
Baseline Swollen Joint Count			
≤15	18/54 (33.3)	80/127(63.0)	0.0002
>15-<27	43/103(41.8)	133/191(69.6)	<0.0001
≥ 27	24/57 (42.1)	75/106(70.8)	0.0004
Baseline Tender Joint Count			
≤21	16/51 (31.4)	75/117(64.1)	<0.0001
>21-<41	43/103(41.8)	150/207(72.5)	<0.0001
≥41	26/60 (43.3)	63/100(63.0)	0.0153
Baseline HAQ			
≤1.125	18/54 (33.3)	83/125(66.4)	<0.0001
>1.125-<2.250	39/106(36.8)	118/175(67.4)	<0.0001
≥2.250	28/54 (51.9)	87/124(70.2)	0.0189

Table 61. Subgroup Analysis of HAQ Response by Baseline Disease Activity (IM101102)

	(1)MTX n=214	(2)Adalimumab n=424	p-values
Disease Duration			
1.0-≤2.0	16/41 (39.0)	67/95 (70.5)	0.0005
2.0-≤5.0	22/46 (47.8)	61/91 (67.0)	0.0298
5.0-≤10.0	18/54 (33.3)	65/105(61.9)	0.0006
>10	28/73 (38.4)	77/133(57.9)	0.0073
Baseline CRP			
≤12	22/66 (33.3)	64/108(59.3)	0.0009
>12-<40	39/99 (39.4)	123/197(62.4)	0.0002
≥40	23/49 (46.9)	83/119(69.8)	0.0054
Baseline Swollen Joint Count			
≤15	20/54 (37.0)	67/127(52.8)	0.0528
>15-<27	42/103(40.8)	129/191(67.5)	<0.0001
≥ 27	22/57 (38.6)	74/106(69.8)	0.0001
Baseline Tender Joint Count			
≤21	19/51 (37.3)	66/117(56.4)	0.0224
>21-<41	39/103(37.9)	134/207(64.7)	<0.0001
≥41	26/60 (43.3)	70/100(70.0)	0.0009
Baseline HAQ			
≤1.125	16/55 (29.1)	61/128(66.4)	0.0197
>1.125-<2.250	41/106(38.7)	116/173(67.1)	<0.0001
≥2.250	27/53 (50.9)	93/123(75.6)	0.0013

Table 62. Subgroup Analysis of Change from Baseline Erosion Score by Baseline Disease Activity (IM101102)

	(1)MTX (n=195)			(2)Abatacept (n=391)			p-values
	Mean (SD)	Median	n	Mean (SD)	Median	n	
Disease Duration							
1.0-≤2.0	2.12 (4.83)	0.27	35	0.59 (1.68)	0.00	90	0.0247
2.0-≤5.0	1.34 (2.94)	0.51	42	1.03 (2.54)	0.28	83	0.4427
5.0-≤10.0	0.76 (1.62)	0.27	51	0.54 (1.28)	0.00	99	0.2420
>10	0.79 (1.73)	0.00	67	0.44 (1.49)	0.00	118	0.5619
Baseline CRP							
≤12	0.67 (1.33)	0.27	59	0.32 (1.24)	0.00	100	0.0291
>12-<40	1.15 (2.58)	0.29	92	0.60 (1.80)	0.00	179	0.0222
≥40	1.73 (4.30)	0.00	44	0.95 (2.05)	0.51	112	0.8676
Baseline Swollen Joint Count							
≤15	0.63 (2.31)	0.00	49	0.35 (1.16)	0.00	117	0.4000
>15-<27	1.12 (2.80)	0.27	93	0.70 (1.55)	0.00	175	0.1755
≥27	1.63 (3.18)	0.52	53	0.83 (2.54)	0.00	99	0.1111
Baseline Tender Joint Count							
≤21	0.98 (3.64)	0.26	47	0.33 (0.91)	0.00	132	0.2821
>21-<41	0.99 (2.39)	0.00	90	0.74 (1.65)	0.00	194	0.6169
≥41	1.49 (2.65)	0.53	58	0.72 (2.58)	0.00	91	0.0149
Baseline HAQ							
≤1.125	1.02 (2.86)	0.51	51	0.38 (1.27)	0.00	113	0.0138
>1.125-<2.250	1.16 (2.94)	0.00	95	0.57 (1.41)	0.00	161	0.1522
≥2.250	1.22 (2.51)	0.27	49	0.95 (2.46)	0.26	117	0.6109

The subgroup analyses of ACR20 at 6 months and HAQ responses at 12 months showed that the abatacept group had statistically higher responses compared to the placebo group across all baseline disease activity subgroups except for HAQ responses at 12 months only when the baseline swollen joint count was less than 15 (p=0.0528).

For the efficacy endpoint of the change from baseline erosion score at 12 months, when disease duration was less than 2 years, baseline CRP was less than 40, baseline tender joint count was greater than 41, and baseline HAQ was less than 1.125, the change from baseline erosion score showed statistically less progression in the abatacept group compared to the placebo group. However, the change from baseline erosion score was numerically less progression in the abatacept group compared to the placebo group.

Appears This Way
On Original

4.2.2 Study IM101029

Table 63. Subgroup Analysis of ACR20 Response by Baseline Disease Activity (IM101029)

	(1)MTX n=133	(2)Adalimumab n=256	p-values
Disease Duration			
1.0-≤2.0	3/16 (18.8)	11/31 (35.5)	0.3211
2.0-≤5.0	5/26 (19.2)	18/31 (58.1)	0.0029
5.0-≤10.0	5/25 (20.0)	34/58 (58.6)	0.0012
>10	13/66 (19.7)	66/136(48.5)	<0.0001
Baseline CRP			
≤12	10/27 (37.0)	21/34 (61.8)	0.0550
>12-<40	13/58 (22.4)	54/118(45.8)	0.0027
≥40	3/48 (6.3)	54/104(51.9)	<0.0001
Baseline Swollen Joint Count			
≤15	6/38 (15.8)	36/78 (46.2)	0.0014
>15-<27	13/60 (21.7)	54/114(47.4)	0.0009
≥ 27	7/35 (20.0)	39/69 (60.9)	0.0001
Baseline Tender Joint Count			
≤21	16/51 (31.4)	75/117(64.1)	<0.0001
>21-<41	11/64 (17.2)	59/123(48.0)	<0.0001
≥41	8/38 (21.1)	37/62 (59.7)	0.0002
Baseline HAQ			
≤1.125	6/26 (23.1)	19/47 (40.4)	0.1347
>1.125-<2.250	10/55 (18.2)	66/116(56.9)	<0.0001
≥2.250	10/52 (19.2)	44/93 (47.3)	0.0011

Table 64. Subgroup Analysis of HAQ Response by Baseline Disease Activity (IM101029)

	(1)MTX n=133	(2)Adalimumab n=256	p-values
Disease Duration			
1.0-≤2.0	3/16 (18.8)	10/31 (32.3)	0.3266
2.0-≤5.0	9/26 (34.6)	15/31 (48.4)	0.2942
5.0-≤10.0	3/25 (12.0)	37/58 (63.8)	<0.0001
>10	16/66 (24.2)	59/136(43.4)	0.0083
Baseline CRP			
≤12	5/27 (18.5)	15/34 (44.1)	0.0540
>12-<40	18/58 (31.0)	56/118(47.5)	0.0380
≥40	8/48 (16.7)	50/104(48.1)	0.0003
Baseline Swollen Joint Count			
≤15	8/38 (21.1)	35/78 (44.9)	0.0144
>15-<27	13/60 (21.7)	53/114(46.5)	0.0013
≥ 27	33/64 (51.6)	10/35 (28.6)	0.0345
Baseline Tender Joint Count			
≤21	5/31 (16.1)	29/71 (40.9)	0.0215
>21-<41	17/64 (26.6)	56/123(45.5)	0.0117
≥41	9/38 (23.7)	36/62 (58.1)	0.0009
Baseline HAQ			
≤1.125	6/28 (21.4)	18/52 (34.6)	0.3074
>1.125-<2.50	8/55 (14.6)	58/114(50.9)	<0.0001
≥2.250	17/50 (34.0)	45/90 (50.0)	0.0678

Table 60 shows that for disease duration less than 2 years, baseline CRP level less than 12, and baseline HAQ less than 1.125, the ACR 20 responses at 6 month were numerically higher in the abatacept group compared to that of the placebo group, but there were no statistically significant differences between the two groups.

Table 61 shows that for disease duration less than 5 years, baseline CRP level less than 12, and baseline HAQ less than 1.125 and greater than 2.250, the HAQ responses at 6 month were numerically higher in the abatacept group compared to that of the placebo group, but there were no statistically significant differences between the two groups.

According to subgroup analyses results, the abatacept group showed more effectiveness in the primary efficacy variables, ACR20 responses at 6 months, HAQ responses at 6 months or 12 months, and the change from baseline erosion score at 12 months as compared to that of the placebo group.

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

The higher percentage of dropouts from the placebo groups than the treated groups were observed mainly due to lack of efficacy across all trials. The sponsor's sensitivity analyses including the worst case sensitivity analyses were robust. Reviewer's results for supportive analyses of GEE methods for primary endpoints ACR20 and HAQ responses using repeated measures for Study IM101102 and Study IM101029 were similar.

For the change from baseline in joint erosion score at 12 months which is the one of co-primary objectives in Study IM101102, all the subjects who had a baseline and a day 365 values were used instead of ITT population for the primary analyses. This reviewer performed sensitivity analyses imputing radiographic scores at day 365 and baseline scores to use ITT population. Although the sensitivity analyses of reviewer's radiographic measures were not robust, the results showed that there was numerically less change from baseline of erosion score as well as joint space narrowing score in the abatacept group as compared to the placebo group.

5.2 Conclusions and Recommendations

Based on the results of the statistical analyses of the efficacy data of three principal studies (IM101100, IM101102, and IM101029) and one safety study (IM101031), the primary endpoint of improvement in signs and symptoms measured by the proportion of ACR20 responses at 6 months were statistically significantly higher in the use of abatacept, given intravenously at a fixed dose approximating 10 mg/kg, compared to the placebo groups. The co-primary endpoint of improving physical function measured by HAQ responses at 6 months or 12 months showed significantly higher in the abatacept groups compared to the placebo groups. The co-primary endpoint of inhibiting the progression of structural damage measured by the change from baseline erosion score at 12 months showed also statistically less progression in the abatacept group compared to the placebo group.

The sponsor's secondary efficacy analyses including a major clinical response and sensitivity analyses were statistically significantly higher in the use of abatacept groups compared to the placebo groups.

The reviewer's supportive analyses for ACR20 and HAQ responses using GEE approaches including all visit measures showed consistent results.

Overall, the use of abatacept, given intravenously at a fixed dose approximating 10 mg/kg, shows in reducing the signs and symptoms of RA, inducing a major clinical response, inhibiting the progression of structural damage, and improving physical function in adults with moderately to severely active RA, who have had an inadequate response to one or more disease-modifying antirheumatic drugs (DMARDs), including TNF-blocking agents. Abatacept may be used in combination with other nonbiologic DMARDs.

SIGNATURES/DISTRIBUTION LIST PAGE

Primary Statistical Reviewer: Kyung Yul Lee, Ph.D.

Kyung Yul Lee 9/12/05

Date:

Concurring Reviewer(s):

Boguang Zhen 9/12/05

Statistical Team Leader: Boguang Zhen, Ph.D.

Biometrics Division Director: Aloka Chakravarty, Ph.D.

Aloka Chakravarty 9/12/05

cc:

HFD-109/Mr. Eric Laughner,
HFD-108/Dr. Keith Hull
HFD-711/Dr. Boguang Zhen
HFD-711 /Dr. Aloka Chakravarty
HFD-700/Dr. Chuck Anello
HFD-711/Chron
HFD-99/DCC

c:\NDA\statreview.doc