



Galapagos

---

## STATISTICAL ANALYSIS PLAN

---

**Project Number:** GLPG0634

**Study Number:** GLPG0634-CL-203 (DARWIN1)

**Study Title** Randomized, double-blind, placebo-controlled, multicenter, phase IIb dose finding study of GLPG0634 administered for 24 weeks in combination with methotrexate to subjects with moderately to severely active rheumatoid arthritis who have an inadequate response to methotrexate alone.

**Development Phase:** IIb

**Status:** Final

**Version:** 5.00

**Date:** 18-Jun-2015

**EudraCT:** 2012-003635-31

**IND:** 155,510

**Sponsor:** Galapagos NV  
Industriepark Mechelen Noord  
Generaal De Wittelaan L11 A3  
2800 Mechelen, Belgium

**Biostatistician:** [REDACTED] IS

Tel: [REDACTED]

**Clinical Study Lead:** [REDACTED]

Tel: [REDACTED]  
GSM: [REDACTED]

---

### CONFIDENTIALITY STATEMENT

The information contained in this document is privileged and confidential. It is the property of Galapagos NV and may not be used, disclosed, reproduced or otherwise disseminated within your organization or communicated to any third parties without the express written authorization of Galapagos NV.

## TABLE OF CONTENTS

<b>List of Tables</b> .....	<b>7</b>
<b>List of Figures</b> .....	<b>17</b>
<b>List of Listings</b> .....	<b>18</b>
<b>1. Study Design</b> .....	<b>20</b>
1.1. Study Objectives .....	20
1.1.1. Primary Objective .....	20
1.1.2. Secondary Objectives .....	20
1.2. Study Design .....	20
1.3. Protocol and Protocol Amendments .....	21
1.4. Flowchart .....	22
<b>2. Analysis Populations and treatment groups</b> .....	<b>24</b>
2.1. Analysis chapters .....	24
2.1.1. Efficacy data .....	24
2.1.2. Safety data .....	26
2.2. Further Analysis populations .....	28
2.2.1. All Screened Subjects Population .....	28
2.2.2. All Randomized Population .....	28
2.3. Handling of unblinded subjects .....	28
2.3.1. Safety population .....	28
2.3.2. Intent to Treat Population .....	28
2.3.3. Per Protocol Population .....	28
2.4. Treatment Groups .....	28
2.4.1. Randomized versus Actual Treatment .....	28
2.4.2. Totals across Groups .....	29
<b>3. Analysis Periods and Analysis Time Points</b> .....	<b>29</b>
3.1. Relative Number of Days .....	29
3.2. Analysis Periods for Non-visit Data .....	29
3.3. Algorithm of Allocating Visits to Time Windows .....	30
3.4. Selection of Visits .....	31
<b>4. Handling of Data</b> .....	<b>31</b>
4.1. Calculation of Descriptive Statistics .....	31
4.2. Calculation of Percentages .....	32
4.3. Handling of Values below (or above) a Threshold .....	32
4.3.1. Safety Data .....	32
4.3.2. PK Data .....	32
4.3.3. PD Data .....	32
4.4. Handling Missing Data .....	32
4.4.1. Handling of Missing Date (Time) or Partially Known Date (Time) .....	32
4.4.2. Handling of missing values in scales or summary parameters .....	32
4.4.3. Handling of missing values and early dropouts in the efficacy analyses during the first study period (up to Week 12) .....	32
4.4.4. Handling of missing values and early dropouts in the efficacy analyses during the study period as a whole (full 24 weeks) .....	33
4.4.5. Handling of missing values and early dropouts in the efficacy analyses during the second study period (after Week 12) .....	33

4.4.6.	Handling of missing values and early dropouts in the safety analysis .....	34
4.5.	Handling treatment switches at Week 12 in the efficacy and safety analysis .....	34
4.5.1.	In the efficacy analysis of the full study as a whole (full 24 weeks) .....	34
4.5.2.	In the safety analysis of the full study as a whole (full 24 weeks).....	34
4.5.3.	In the efficacy and safety analysis of the second study period .....	34
4.6.	Handling of Outliers.....	35
4.7.	Handling of seconds in datetime fields .....	35
4.8.	Stratification factors.....	35
<b>5.</b>	<b>Software and Procedures .....</b>	<b>35</b>
5.1.	Software.....	35
5.2.	Procedures .....	35
5.3.	Formats.....	36
<b>6.</b>	<b>Statistical Methods.....</b>	<b>36</b>
6.1.	Planned Analyses, Protocol Amendments Included .....	36
6.1.1.	Demographic and other baseline characteristics .....	36
6.1.2.	Efficacy Variables .....	37
6.1.3.	Safety Variables .....	41
6.1.4.	Pharmacokinetic Analyses.....	41
6.1.5.	Pharmacodynamic Analyses .....	41
6.1.6.	Pharmacokinetic, Efficacy, Safety and Pharmacodynamic Correlations ....	41
6.2.	Interim Analysis.....	41
6.3.	Changes to the Planned Analyses, Not Covered by Protocol Amendments .....	43
6.3.1.	Changes before Interim Database Lock .....	43
6.3.2.	Changes after Interim Database Lock but before Final Database Lock.....	43
6.3.3.	Changes after Final Database Lock.....	44
<b>7.</b>	<b>Definitions of General Tables, Listings and Figures.....</b>	<b>44</b>
7.1.	Subject Disposition .....	44
7.2.	Protocol Deviations and Eligibility.....	47
7.3.	Subjects Excluded from Analysis .....	47
7.4.	Demographics and Baseline Disease Characteristics.....	48
7.4.1.	Parameters.....	48
7.4.2.	Analysis .....	48
7.5.	Baseline disease characteristics .....	50
7.5.1.	Parameters.....	50
7.5.2.	Analysis .....	51
7.6.	Medical History and Concurrent Diseases .....	53
7.7.	Prior and Concomitant Therapies .....	53
7.7.1.	Classification of Therapies .....	53
7.7.2.	Coding of Therapies .....	55
7.7.3.	Analysis .....	55
7.8.	Exposure to Study Medication and Compliance .....	56
7.8.1.	Derivations for placebo/GLPG0634.....	56
7.8.2.	Tables up to Week 12.....	56
7.8.3.	Tables for the whole 24 weeks period.....	57
7.8.4.	Listings .....	57
7.9.	Exposure to methotrexate and Compliance.....	58
7.9.1.	Derivations for methotrexate .....	58
7.9.2.	Tables up to Week 12 .....	58

7.9.3.	Tables for the whole 24 weeks period	59
7.9.4.	Listings	59
<b>8.</b>	<b>Definitions of Efficacy Tables, Listings and Figures</b>	<b>59</b>
8.1.	ACR20, ACR50, ACR70 and ACR-N	60
8.1.1.	Parameters	60
8.1.2.	Tables over the whole 24 weeks period	61
8.1.3.	Figures over the whole 24 weeks period	70
8.1.4.	Tables after Week 12	71
8.1.5.	Figures after Week 12	72
8.1.6.	Listings	73
8.2.	EULAR remission	73
8.2.1.	Parameters	73
8.2.2.	Tables over the whole 24 weeks	73
8.2.3.	Figures over the whole 24 weeks period	73
8.2.4.	Tables after Week 12	74
8.2.5.	Figures after Week 12	74
8.2.6.	Listings	74
8.3.	SDAI and CDAI scores	74
8.3.1.	Parameters	74
8.3.2.	Tables over the whole 24 weeks period	75
8.3.3.	Figures over the whole 24 weeks period	78
8.3.4.	Tables after Week 12	78
8.3.5.	Figures after Week 12	78
8.3.6.	Listings	79
8.4.	DAS28(CRP)	79
8.4.1.	Parameters	79
8.4.2.	Tables over the whole 24 weeks period	80
8.4.3.	Figures over the whole 24 weeks period	84
8.4.4.	Tables after Week 12	84
8.4.5.	Figures after Week 12	84
8.4.6.	Listings	84
8.5.	ACR component: Tender and swollen joint counts	85
8.5.1.	Parameters	85
8.5.2.	Tables over the whole 24 weeks period	85
8.5.3.	Figures over the whole 24 weeks period	86
8.5.4.	Tables after Week 12	86
8.5.5.	Figures after Week 12	86
8.5.6.	Listings	86
8.6.	ACR component: VAS scales	87
8.6.1.	Parameters	87
8.6.2.	Tables over the whole 24 weeks period	87
8.6.3.	Figures over the whole 24 weeks study period	88
8.6.4.	Tables after Week 12	88
8.6.5.	Figures after Week 12	88
8.6.6.	Listings	88
8.7.	ACR component: HAQ-DI	89
8.7.1.	Parameters	89
8.7.2.	Tables over the whole 24 weeks period	89
8.7.3.	Figures over the whole 24 weeks period	90

8.7.4.	Tables after Week 12.....	91
8.7.5.	Figures after Week 12 .....	91
8.7.6.	Listings .....	91
8.8.	ACR component: Serum CRP .....	91
8.8.1.	Parameters.....	91
8.8.2.	Tables over the whole 24 weeks period .....	92
8.8.3.	Figures over the whole 24 weeks period .....	93
8.8.4.	Tables after Week 12.....	93
8.8.5.	Figures after Week 12 .....	93
8.8.6.	Listings .....	93
8.9.	Quality of life: FACIT fatigue scale .....	94
8.9.1.	Parameters.....	94
8.9.2.	Tables over the whole 24 weeks period .....	94
8.9.3.	Figures over the whole 24 weeks period .....	95
8.9.4.	Tables after Week 12.....	95
8.9.5.	Figures after Week 12 .....	95
8.9.6.	Listings .....	95
8.10.	Quality of life: SF-36 .....	96
8.10.1.	Parameters.....	96
8.10.2.	Tables over the whole 24 weeks period .....	101
8.10.3.	Figures over the whole 24 weeks period .....	102
8.10.4.	Tables after Week 12.....	102
8.10.5.	Figures after Week 12 .....	102
8.10.6.	Listings .....	102
<b>9.</b>	<b>Definitions of Pharmacokinetics Tables, Listings and Figures.....</b>	<b>103</b>
<b>10.</b>	<b>Definitions of Pharmacodynamics Tables, Listings and Figures.....</b>	<b>103</b>
10.1.	Parameters .....	103
10.2.	Tables .....	103
10.3.	Figures.....	104
10.4.	Listings .....	105
<b>11.</b>	<b>Definitions of PK/PD Tables, Listings and Figures .....</b>	<b>105</b>
<b>12.</b>	<b>Definitions of safety Tables, Listings and Figures .....</b>	<b>105</b>
12.1.	Adverse Events .....	105
12.1.1.	Treatment-Emergent Principle.....	105
12.1.2.	Treatment Relatedness.....	106
12.1.3.	Worst-Case Principle.....	106
12.1.4.	Adverse Event Onset Day and Duration.....	106
12.1.5.	Calculation of Percentages .....	107
12.1.6.	Tables up to Week 12.....	107
12.1.7.	Tables over the whole 24 weeks period .....	110
12.1.8.	Listings .....	111
12.2.	Laboratory Safety.....	113
12.2.1.	Laboratory Units .....	113
12.2.2.	Derived lab tests.....	113
12.2.3.	Number of Significant Digits.....	114
12.2.4.	Baseline and Change from Baseline.....	114
12.2.5.	Scoring of Laboratory Values .....	114
12.2.6.	Handling of Urinalysis Results .....	116

12.2.7. Last on-treatment result .....	116
12.2.8. Worst-Case Abnormality .....	117
12.2.9. Tables up to Week 12 .....	118
12.2.10. Figures up to Week 12 .....	121
12.2.11. Tables over the whole 24 weeks period .....	122
12.2.12. Figures over the whole 24 weeks period .....	123
12.2.13. Listings .....	123
12.3. ECG .....	124
12.3.1. Available Data .....	124
12.3.2. Calculated Parameters .....	124
12.3.3. Baseline and Change from Baseline .....	124
12.3.4. Visits for analysis .....	124
12.3.5. Normal Ranges .....	125
12.3.6. Analysis tables at Week 12 .....	125
12.3.7. Analysis tables for the whole 24 weeks period .....	126
12.3.8. Figures .....	127
12.3.9. Listings .....	127
12.4. Vital Signs .....	127
12.4.1. Available Data .....	127
12.4.2. Baseline and Change from Baseline .....	127
12.4.3. Normal Ranges .....	127
12.4.4. Last on-treatment result .....	128
12.4.5. Worst-Case Abnormality .....	128
12.4.6. Tables up to Week 12 .....	128
12.4.7. Tables over the whole 24 weeks period .....	129
12.4.8. Figures .....	130
12.4.9. Listings .....	130
12.5. Physical Examinations .....	130
<b>13. References .....</b>	<b>130</b>

## LIST OF TABLES

Table 14.1.1.1:	Subject disposition: Tabulation by region, country and investigator.....	44
Table 14.1.1.2:	Subject disposition: Analysis populations .....	45
Table 14.1.1.3:	Subject disposition: No-treatment subjects: Tabulation of the reasons .....	45
Table 14.1.1.4:	Subject disposition up to Week 12: Tabulation of the number of subject at each time interval.....	45
Table 14.1.1.5:	Subject disposition up to study end: Tabulation of the number of subject at each time interval.....	45
Table 14.1.1.6:	Subject disposition: First and last date in the study .....	46
Table 14.1.1.7:	Subject disposition up to Week 12: Tabulation of the trial termination reasons.....	46
Table 14.1.1.8:	Subject disposition up to study end: Tabulation of the trial termination reasons.....	46
Table 14.1.2.1:	Major protocol deviations.....	47
Table 14.1.3.1:	Demographic data .....	48
Table 14.1.3.2:	Demographic data for ACR20 responders versus non-responders at Week 12.....	49
Table 14.1.3.3:	Baseline laboratory data: Tabulation of abnormalities.....	49
Table 14.1.3.4:	Baseline ECG data: Tabulation of abnormalities.....	49
Table 14.1.3.5:	Baseline vital signs data: Tabulation of abnormalities.....	49
Table 14.1.3.6:	Baseline physical examination: Tabulation .....	49
Table 14.1.3.7:	Screening and baseline disease characteristics .....	51
Table 14.1.3.8:	Screening and baseline disease characteristics by region.....	51
Table 14.1.3.9:	Screening and baseline disease characteristics for ACR20 responders versus non-responders at Week 12 .....	52
Table 14.1.3.10:	Affected joints: Tabulation .....	52
Table 14.1.3.11:	Prior RA medication.....	52
Table 14.1.3.12:	Prior RA medication for ACR20 responders versus non-responders at Week 12 .....	52
Table 14.1.3.13:	Prior methotrexate: Descriptive statistics .....	52
Table 14.1.3.14:	Prior steroid dose at the start of the study: Descriptive statistics.....	52
Table 14.1.3.15:	Medical history: Tabulation.....	53
Table 14.1.3.16:	Concurrent diseases: Tabulation .....	53
Table 14.1.3.17:	Prior therapies: Frequency table.....	55
Table 14.1.3.18:	Concomitant therapies up to Week 12: Frequency table .....	55
Table 14.1.3.19:	Concomitant therapies over the whole study period: Frequency table .....	55
Table 14.1.3.20:	Post-treatment therapies: Frequency table .....	55
Table 14.1.4.1:	Use of study medication up to Week 12: Descriptive statistics.....	56
Table 14.1.4.2:	Use of study medication over the whole study period: Descriptive statistics.....	57
Table 14.1.4.3:	Changes in methotrexate up to Week 12: tabulation.....	58

Table 14.1.4.4:	Use of methotrexate up to Week 12: Descriptive statistics of the average weekly dose.....	58
Table 14.1.4.5:	Changes in methotrexate over the whole study period: tabulation .....	59
Table 14.1.4.6:	Use of methotrexate over the whole study period: Descriptive statistics of the average weekly dose.....	59
Table 14.2.1.1.1:	ACR20: Summary with overall p-value and pairwise comparisons versus placebo .....	61
Table 14.2.1.1.2:	ACR20: Exploratory pairwise comparisons among the dose groups within each regimen .....	62
Table 14.2.1.1.3:	ACR20: Exploratory pairwise comparisons between regimens within each daily dose level.....	62
Table 14.2.1.1.4:	ACR50: Summary with overall p-value and pairwise comparisons versus placebo .....	63
Table 14.2.1.1.5:	ACR50: Exploratory pairwise comparisons among the dose groups within each regimen .....	63
Table 14.2.1.1.6:	ACR50: Exploratory pairwise comparisons between regimens within each daily dose level.....	64
Table 14.2.1.1.7:	ACR70: Summary with overall p-value and pairwise comparisons versus placebo .....	64
Table 14.2.1.1.8:	ACR70: Exploratory pairwise comparisons among the dose groups within each regimen .....	64
Table 14.2.1.1.9:	ACR70: Exploratory pairwise comparisons between regimens within each daily dose level.....	64
Table 14.2.1.1.10:	ACR20 at Week 12: Tabulation versus baseline characteristics.....	65
Table 14.2.1.1.11:	ACR20 at Week 12: Exploratory logistic regression model containing baseline characteristics .....	66
Table 14.2.1.1.12:	ACR-N: Summary with overall p-value and pairwise comparisons versus placebo .....	67
Table 14.2.1.1.13:	ACR-N: Exploratory pairwise comparisons among the dose groups within each regimen .....	67
Table 14.2.1.1.14:	ACR-N: Exploratory pairwise comparisons between regimens within each daily dose level.....	68
Table 14.2.1.1.15:	Time to ACR20 response: Summary with overall p-value and pairwise comparisons versus placebo.....	68
Table 14.2.1.1.16:	Time to ACR20 response: Exploratory pairwise comparisons among the dose groups within each regimen.....	69
Table 14.2.1.1.17:	Time to ACR20 response: Exploratory pairwise comparisons between regimens within each daily dose level.....	69
Table 14.2.1.1.18:	Time to ACR50 response: Summary with overall p-value and pairwise comparisons versus placebo.....	69
Table 14.2.1.1.19:	Time to ACR50 response: Exploratory pairwise comparisons among the dose groups within each regimen.....	70



Table 14.2.1.1.20: Time to ACR50 response: Exploratory pairwise comparisons between regimens within each daily dose level .....	70
Table 14.2.1.1.21: Time to ACR70 response: Summary with overall p-value and pairwise comparisons versus placebo .....	70
Table 14.2.1.1.22: Time to ACR70 response: Exploratory pairwise comparisons among the dose groups within each regimen .....	70
Table 14.2.1.1.23: Time to ACR70 response: Exploratory pairwise comparisons between regimens within each daily dose level .....	70
Table 14.2.1.1.24: ACR20 after Week 12: Frequency table .....	71
Table 14.2.1.1.25: ACR50 after Week 12: Frequency table .....	72
Table 14.2.1.1.26: ACR70 after Week 12: Frequency table .....	72
Table 14.2.1.1.27: ACR-N after Week 12: Descriptive statistics .....	72
Table 14.2.1.2.1: EULAR remission: Summary with overall p-value and pairwise comparisons versus placebo .....	73
Table 14.2.1.2.2: EULAR remission: Exploratory pairwise comparisons among the dose groups within each regimen .....	73
Table 14.2.1.2.3: EULAR remission: Exploratory pairwise comparisons between regimens within each daily dose level .....	73
Table 14.2.1.2.4: EULAR remission after Week 12: Frequency table .....	74
Table 14.2.1.3.1: SDAI and CDAI: Descriptive statistics of the actual values .....	75
Table 14.2.1.3.2: SDAI and CDAI: Descriptive statistics of the change from baseline .....	75
Table 14.2.1.3.3: SDAI and CDAI: Descriptive statistics of the percent change from baseline .....	75
Table 14.2.1.3.4: SDAI and CDAI: Summary with overall p-value and pairwise comparisons versus placebo .....	76
Table 14.2.1.3.5: SDAI and CDAI: Exploratory pairwise comparisons among the dose groups within each regimen .....	76
Table 14.2.1.3.6: SDAI and CDAI: Exploratory pairwise comparisons between regimens within each daily dose level .....	77
Table 14.2.1.3.7: SDAI and CDAI remission: Summary with overall p-value and pairwise comparisons versus placebo .....	77
Table 14.2.1.3.8: SDAI and CDAI remission: Exploratory pairwise comparisons among the dose groups within each regimen .....	78
Table 14.2.1.3.9: SDAI and CDAI remission: Exploratory pairwise comparisons between regimens within each daily dose level .....	78
Table 14.2.1.3.10: SDAI and CDAI after Week 12: Descriptive statistics of the actual values .....	78
Table 14.2.1.3.11: SDAI and CDAI after Week 12: Descriptive statistics of the change from baseline .....	78
Table 14.2.1.3.12: SDAI and CDAI after Week 12: Descriptive statistics of the percent change from baseline .....	78
Table 14.2.1.4.1: DAS28(CRP): Descriptive statistics of the actual values .....	80
Table 14.2.1.4.2: DAS28(CRP): Descriptive statistics of the change from baseline .....	80

Table 14.2.1.4.3: DAS28(CRP): Descriptive statistics of the percent change from baseline .....	80
Table 14.2.1.4.4: DAS28(CRP): Summary with overall p-value and pairwise comparisons versus placebo.....	80
Table 14.2.1.4.5: DAS28(CRP): Exploratory pairwise comparisons among the dose groups within each regimen.....	80
Table 14.2.1.4.6: DAS28(CRP): Exploratory pairwise comparisons between regimens within each daily dose level.....	81
Table 14.2.1.4.7: DAS28(CRP) change from baseline at Week 12: Tabulation versus baseline characteristics .....	81
Table 14.2.1.4.8: DAS28(CRP) change from baseline at Week 12: Exploratory ANCOVA model containing baseline characteristics.....	82
Table 14.2.1.4.9: DAS28(CRP) remission: Summary with overall p-value and pairwise comparisons versus placebo.....	83
Table 14.2.1.4.10: DAS28(CRP) remission: Exploratory pairwise comparisons among the dose groups within each regimen .....	83
Table 14.2.1.4.11: DAS28(CRP) remission: Exploratory pairwise comparisons between regimens within each daily dose level.....	83
Table 14.2.1.4.12: DAS28(CRP) remission and low disease activity: Summary with overall p-value and pairwise comparisons versus placebo .....	83
Table 14.2.1.4.13: DAS28(CRP) remission and low disease activity: Exploratory pairwise comparisons among the dose groups within each regimen .....	83
Table 14.2.1.4.14: DAS28(CRP) remission and low disease activity: Exploratory pairwise comparisons between regimens within each daily dose level .....	83
Table 14.2.1.4.15: DAS28(CRP) EULAR response: Frequency table.....	83
Table 14.2.1.4.16: DAS28(CRP) EULAR response: Exploratory pairwise comparisons among the dose groups within each regimen.....	84
Table 14.2.1.4.17: DAS28(CRP) EULAR response: Exploratory pairwise comparisons between regimens within each daily dose level.....	84
Table 14.2.1.4.18: DAS28(CRP) after Week 12: Descriptive statistics of the actual values .....	84
Table 14.2.1.4.19: DAS28(CRP) after Week 12: Descriptive statistics of the change from baseline.....	84
Table 14.2.1.4.20: DAS28(CRP) after Week 12: Descriptive statistics of the percent change from baseline.....	84
Table 14.2.1.4.21: DAS28(CRP) EULAR response after Week 12: Frequency table.....	84
Table 14.2.1.5.1: Joint counts: Descriptive statistics of the actual values.....	85
Table 14.2.1.5.2: Joint counts: Descriptive statistics of the change from baseline.....	85
Table 14.2.1.5.3: Joint counts: Descriptive statistics of the percent change from baseline .....	85
Table 14.2.1.5.4: Joint counts: Summary with overall p-value and pairwise comparisons versus placebo.....	86

Table 14.2.1.5.5: Joint counts: Exploratory pairwise comparisons among the dose groups within each regimen .....	86
Table 14.2.1.5.6: Joint counts: Exploratory pairwise comparisons between regimens within each daily dose level .....	86
Table 14.2.1.5.7: Joint counts after Week 12: Descriptive statistics of the actual values .....	86
Table 14.2.1.5.8: Joint counts after Week 12: Descriptive statistics of the change from baseline .....	86
Table 14.2.1.5.9: Joint counts after Week 12: Descriptive statistics of the percent change from baseline .....	86
Table 14.2.1.6.1: VAS scores: Descriptive statistics of the actual values .....	87
Table 14.2.1.6.2: VAS scores: Descriptive statistics of the change from baseline .....	87
Table 14.2.1.6.3: VAS scores: Descriptive statistics of the percent change from baseline .....	87
Table 14.2.1.6.4: VAS scores: Summary with overall p-value and pairwise comparisons versus placebo .....	87
Table 14.2.1.6.5: VAS scores: Exploratory pairwise comparisons among the dose groups within each regimen .....	88
Table 14.2.1.6.6: VAS scores: Exploratory pairwise comparisons between regimens within each daily dose level .....	88
Table 14.2.1.6.7: VAS scores after Week 12: Descriptive statistics of the actual values .....	88
Table 14.2.1.6.8: VAS scores after Week 12: Descriptive statistics of the change from baseline .....	88
Table 14.2.1.6.9: VAS scores after Week 12: Descriptive statistics of the percent change from baseline .....	88
Table 14.2.1.7.1: HAQ-DI total and domain scores: Descriptive statistics of the actual values .....	89
Table 14.2.1.7.2: HAQ-DI total and domain scores: Descriptive statistics of the change from baseline .....	90
Table 14.2.1.7.3: HAQ-DI total and domain scores: Descriptive statistics of the percent change from baseline .....	90
Table 14.2.1.7.4: HAQ-DI total and domain scores: Summary with overall p-value and pairwise comparisons versus placebo .....	90
Table 14.2.1.7.5: HAQ-DI total and domain scores: Exploratory pairwise comparisons among the dose groups within each regimen .....	90
Table 14.2.1.7.6: HAQ-DI total and domain scores: Exploratory pairwise comparisons between regimens within each daily dose level .....	90
Table 14.2.1.7.7: HAQ-DI total and domain scores after Week 12: Descriptive statistics of the actual values .....	91
Table 14.2.1.7.8: HAQ-DI total and domain scores after Week 12: Descriptive statistics of the change from baseline .....	91

Table 14.2.1.7.9: HAQ-DI total and domain scores after Week 12: Descriptive statistics of the percent change from baseline.....	91
Table 14.2.1.8.1: Serum CRP (mg/L): Descriptive statistics of the actual values.....	92
Table 14.2.1.8.2: Serum CRP (mg/L): Descriptive statistics of the change from baseline.....	92
Table 14.2.1.8.3: Serum CRP (mg/L): Descriptive statistics of the percent change from baseline.....	92
Table 14.2.1.8.4: Serum CRP (mg/L): Shift table versus baseline.....	92
Table 14.2.1.8.5: Serum CRP (mg/L): Treatment-emergent normalizations per time point.....	92
Table 14.2.1.8.6: Serum CRP (mg/L): Summary with overall p-value and pairwise comparisons versus placebo.....	92
Table 14.2.1.8.7: Serum CRP (mg/L): Exploratory pairwise comparisons among the dose groups within each regimen.....	92
Table 14.2.1.8.8: Serum CRP (mg/L): Exploratory pairwise comparisons between regimens within each daily dose level.....	93
Table 14.2.1.8.9: Serum CRP (mg/L) after Week 12: Descriptive statistics of the actual values.....	93
Table 14.2.1.8.10: Serum CRP (mg/L) after Week 12: Descriptive statistics of the change from baseline.....	93
Table 14.2.1.8.11: Serum CRP (mg/L) after Week 12: Descriptive statistics of the percent change from baseline.....	93
Table 14.2.1.8.12: Serum CRP (mg/L) after Week 12: Shift table versus baseline.....	93
Table 14.2.1.8.13: Serum CRP (mg/L) after Week 12: Treatment-emergent normalizations per time point.....	93
Table 14.2.1.9.1: FACIT fatigue score: Descriptive statistics of the actual values.....	94
Table 14.2.1.9.2: FACIT fatigue score: Descriptive statistics of the change from baseline.....	94
Table 14.2.1.9.3: FACIT fatigue score: Descriptive statistics of the percent change from baseline.....	95
Table 14.2.1.9.4: FACIT fatigue score: Summary with overall p-value and pairwise comparisons versus placebo.....	95
Table 14.2.1.9.5: FACIT fatigue score: Exploratory pairwise comparisons among the dose groups within each regimen.....	95
Table 14.2.1.9.6: FACIT fatigue score: Exploratory pairwise comparisons between regimens within each daily dose level.....	95
Table 14.2.1.9.7: FACIT fatigue score after Week 12: Descriptive statistics of the actual values.....	95
Table 14.2.1.9.8: FACIT fatigue score after Week 12: Descriptive statistics of the change from baseline.....	95
Table 14.2.1.9.9: FACIT fatigue score after Week 12: Descriptive statistics of the percent change from baseline.....	95
Table 14.2.1.10.1: SF-36: Descriptive statistics of the actual values.....	101

Table 14.2.1.10.2: SF-36: Descriptive statistics of the change from baseline.....	101
Table 14.2.1.10.3: SF-36: Descriptive statistics of the percent change from baseline .....	101
Table 14.2.1.10.4: SF-36: Summary with overall p-value and pairwise comparisons versus placebo .....	101
Table 14.2.1.10.5: SF-36: Exploratory pairwise comparisons among the dose groups within each regimen .....	102
Table 14.2.1.10.6: SF-36: Exploratory pairwise comparisons between regimens within each daily dose level.....	102
Table 14.2.1.10.7: SF-36 after Week 12: Descriptive statistics of the actual values.....	102
Table 14.2.1.10.8: SF-36 after Week 12: Descriptive statistics of the change from baseline .....	102
Table 14.2.1.10.9: SF-36 after Week 12: Descriptive statistics of the percent change from baseline.....	102
Table 14.2.3.1: Pharmacodynamics: YKL-40: Descriptive statistics.....	103
Table 14.2.3.2: Pharmacodynamics: YKL-40: Summary with overall p-value and pairwise comparisons versus placebo.....	103
Table 14.2.3.3: Pharmacodynamics: YKL-40: Exploratory pairwise comparisons among the dose groups within each regimen.....	103
Table 14.2.3.4: Pharmacodynamics: YKL-40: Exploratory pairwise comparisons between regimens within each daily dose level.....	104
Table 14.2.3.5: Pharmacodynamics: YKL-40: Descriptive statistics for Week 12 ACR20 responders versus non-responders.....	104
Table 14.2.3.6: Pharmacodynamics: YKL-40: Descriptive statistics for Week 12 ACR50 responders versus non-responders.....	104
Table 14.2.3.7: Pharmacodynamics: YKL-40: Descriptive statistics for Week 12 ACR70 responders versus non-responders.....	104
Table 14.2.3.8: Pharmacodynamics: YKL-40: Correlation with ACR20/50/70.....	104
Table 14.2.3.9: Pharmacodynamics: YKL-40: Correlation with continuous efficacy parameters.....	104
Table 14.3.1.1: Treatment-emergent adverse events up to Week 12: Summary table .....	107
Table 14.3.1.2: Treatment-emergent adverse events up to Week 12: Summary table by region .....	107
Table 14.3.1.3: Treatment-emergent adverse events up to Week 12: Tabulation of all adverse events .....	107
Table 14.3.1.4: Treatment-emergent adverse events up to Week 12: Tabulation of all adverse events by region.....	108
Table 14.3.1.5: Treatment-emergent adverse events up to Week 12: Tabulation of all adverse events by gender .....	108
Table 14.3.1.6: Treatment-emergent adverse events up to Week 12: Tabulation of all adverse events by age category .....	108
Table 14.3.1.7: Treatment-emergent adverse events up to Week 12: Tabulation per intensity.....	108

Table 14.3.1.8:	Treatment-emergent adverse events up to Week 12: Tabulation of all treatment-related events.....	108
Table 14.3.1.9:	Treatment-emergent adverse events up to Week 12: Tabulation of all treatment-related events by gender.....	108
Table 14.3.1.10:	Treatment-emergent adverse events up to Week 12: Tabulation of all treatment-related events by age category.....	108
Table 14.3.1.11:	Treatment-emergent adverse events up to Week 12: Tabulation of the intensity of treatment-related events.....	109
Table 14.3.1.12:	Treatment-emergent adverse events up to Week 12: Tabulation of the drug-relatedness of serious events.....	109
Table 14.3.1.13:	Treatment-emergent adverse events up to Week 12: Tabulation of the drug-relatedness of events leading to discontinuation.....	109
Table 14.3.1.14:	Treatment-emergent adverse events up to Week 12: Tabulation of the drug-relatedness of events leading to a temporary stop.....	109
Table 14.3.1.15:	Treatment-emergent adverse events up to Week 12: Tabulation of the drug-relatedness of individual stopping rules .....	109
Table 14.3.1.16:	Treatment-emergent adverse events over the whole study: Summary table .....	110
Table 14.3.1.17:	Treatment-emergent adverse events over the whole study: Summary table by region.....	110
Table 14.3.1.18:	Treatment-emergent adverse events over the whole study: Tabulation of all adverse events.....	110
Table 14.3.1.19:	Treatment-emergent adverse events over the whole study: Tabulation of all adverse events by region .....	110
Table 14.3.1.20:	Treatment-emergent adverse events over the whole study: Tabulation of all adverse events by gender .....	110
Table 14.3.1.21:	Treatment-emergent adverse events over the whole study: Tabulation of all adverse events by age category .....	111
Table 14.3.1.22:	Treatment-emergent adverse events over the whole study: Tabulation per intensity.....	111
Table 14.3.1.23:	Treatment-emergent adverse events over the whole study: Tabulation of all treatment-related events .....	111
Table 14.3.1.24:	Treatment-emergent adverse events over the whole study: Tabulation of all treatment-related events by gender.....	111
Table 14.3.1.25:	Treatment-emergent adverse events over the whole study: Tabulation of all treatment-related events by age category.....	111
Table 14.3.1.26:	Treatment-emergent adverse events over the whole study: Tabulation of the intensity of treatment-related events.....	111
Table 14.3.1.27:	Treatment-emergent adverse events over the whole study: Tabulation of the drug-relatedness of serious events .....	111
Table 14.3.1.28:	Treatment-emergent adverse events over the whole study: Tabulation of the drug-relatedness of events leading to discontinuation .....	111

Table 14.3.1.29:	Treatment-emergent adverse events over the whole study: Tabulation of the drug-relatedness of events leading to a temporary stop .....	111
Table 14.3.1.30:	Treatment-emergent adverse events over the whole study: Tabulation of the drug-relatedness of individual stopping rules.....	111
Table 14.3.2.1:	Laboratory data up to Week 12: Descriptive statistics of the actual values per time point .....	118
Table 14.3.2.2:	Laboratory data up to Week 12: Descriptive statistics of the changes from baseline per time point .....	118
Table 14.3.2.3:	Laboratory data up to Week 12: Descriptive statistics of the percent changes from baseline per time point.....	118
Table 14.3.2.4:	Laboratory data up to Week 12: Shift table of the lab abnormalities per time point .....	119
Table 14.3.2.5:	Laboratory data up to Week 12: Treatment-emergent lab abnormalities per time point .....	119
Table 14.3.2.6:	Laboratory data up to Week 12: Frequency table of the CTCAE toxicity grades per time point .....	119
Table 14.3.2.7:	Laboratory data up to Week 12: Shift table of the CTCAE toxicity grades per time point .....	119
Table 14.3.2.8:	Laboratory data up to Week 12: Treatment-emergent CTCAE toxicity grades per time point .....	119
Table 14.3.2.9:	Laboratory data up to Week 12: Frequency table of the ATP III classes per time point.....	120
Table 14.3.2.10:	Laboratory data up to Week 12: Shift table of the ATP III classes per time point .....	120
Table 14.3.2.11:	Laboratory data up to Week 12: Treatment-emergent ATP III classes per time point.....	120
Table 14.3.2.12:	Laboratory data up to Week 12: Frequency table of percent changes from baseline per time point in inhibin B and testosterone .....	120
Table 14.3.2.13:	Laboratory data up to Week 12: Shift table of the categorical lab data per time point.....	121
Table 14.3.2.14:	Laboratory data over the whole study: Descriptive statistics of the actual values per time point .....	122
Table 14.3.2.15:	Laboratory data over the whole study: Descriptive statistics of the changes from baseline per time point.....	122
Table 14.3.2.16:	Laboratory data over the whole study: Descriptive statistics of the percent changes from baseline per time point.....	122
Table 14.3.2.17:	Laboratory data over the whole study: Shift table of the lab abnormalities per time point .....	122
Table 14.3.2.18:	Laboratory data over the whole study: Treatment-emergent lab abnormalities per time point .....	122
Table 14.3.2.19:	Laboratory data over the whole study: Frequency table of the CTCAE toxicity grades per time point.....	122

Table 14.3.2.20:	Laboratory data over the whole study: Shift table of the CTCAE toxicity grades per time point .....	122
Table 14.3.2.21:	Laboratory data over the whole study: Treatment-emergent CTCAE toxicity grades per time point .....	122
Table 14.3.2.22:	Laboratory data over the whole study: Frequency table of the ATP III classes per time point.....	122
Table 14.3.2.23:	Laboratory data over the whole study: Shift table of the ATP III classes per time point.....	122
Table 14.3.2.24:	Laboratory data over the whole study: Treatment-emergent ATP III classes per time point.....	122
Table 14.3.2.25:	Laboratory data over the whole study: Frequency table of percent changes from baseline per time point in inhibin B and testosterone.....	122
Table 14.3.2.26:	Laboratory data over the whole study: Shift table of the categorical lab data per time point .....	122
Table 14.3.3.1:	ECG up to Week 12: Descriptive statistics of the actual values per time point .....	125
Table 14.3.3.2:	ECG at Week 12: Descriptive statistics of the changes from baseline ....	125
Table 14.3.3.3:	ECG at Week 12: Shift table of the QT abnormalities .....	125
Table 14.3.3.4:	ECG at Week 12: Treatment-emergent QT abnormalities .....	126
Table 14.3.3.5:	ECG at Week 12: Treatment-emergent abnormal change in QT .....	126
Table 14.3.3.6:	ECG up to Week 12: Frequency table of the eCRF ECG interpretation per time point.....	126
Table 14.3.3.7:	ECG over the whole study: Descriptive statistics of the actual values per time point .....	126
Table 14.3.3.8:	ECG over the whole study: Descriptive statistics of the changes from baseline.....	126
Table 14.3.3.9:	ECG over the whole study: Shift table of the QT abnormalities .....	126
Table 14.3.3.10:	ECG over the whole study: Treatment-emergent QT abnormalities.....	126
Table 14.3.3.11:	ECG over the whole study: Treatment-emergent abnormal change in QT.....	126
Table 14.3.3.12:	ECG over the whole study: Frequency table of the eCRF ECG interpretation per time point.....	126
Table 14.3.4.1:	Vital signs up to Week 12: Descriptive statistics of the actual values .....	129
Table 14.3.4.2:	Vital signs up to Week 12: Descriptive statistics of the changes from baseline.....	129
Table 14.3.4.3:	Vital signs up to Week 12: Summary table of means .....	129
Table 14.3.4.4:	Vital signs up to Week 12: Shift table of the abnormalities per time point.....	129
Table 14.3.4.5:	Vital signs up to Week 12: Treatment-emergent abnormalities per time point .....	129
Table 14.3.4.6:	Vital signs over the whole study: Descriptive statistics of the actual values .....	130



Table 14.3.4.7:	Vital signs over the whole study: Descriptive statistics of the changes from baseline .....	130
Table 14.3.4.8:	Vital signs over the whole study: Summary table of means.....	130
Table 14.3.4.9:	Vital signs over the whole study: Shift table of the abnormalities per time point .....	130
Table 14.3.4.10:	Vital signs over the whole study: Treatment-emergent abnormalities per time point .....	130

## LIST OF FIGURES

Figure 14.2.1.1.1:	ACR20/50/70: Plot per week .....	70
Figure 14.2.1.1.2:	ACR20: Plot over time .....	70
Figure 14.2.1.1.3:	ACR50: Plot over time .....	71
Figure 14.2.1.1.4:	ACR70: Plot over time .....	71
Figure 14.2.1.1.5:	ACR-N: Plot over time .....	71
Figure 14.2.1.1.6:	Time to ACR20 response: Kaplan-Meier plot .....	71
Figure 14.2.1.1.7:	Time to ACR50 response: Kaplan-Meier plot .....	71
Figure 14.2.1.1.8:	Time to ACR70 response: Kaplan-Meier plot .....	71
Figure 14.2.1.2.1:	EULAR remission: Plot per week .....	73
Figure 14.2.1.2.2:	EULAR remission: Plot over time .....	73
Figure 14.2.1.3.1:	SDAI and CDAI: Plot of the actual values over time .....	78
Figure 14.2.1.3.2:	SDAI and CDAI: Plot of the changes from baseline over time .....	78
Figure 14.2.1.3.3:	SDAI and CDAI: Plot of the percent changes from baseline over time .....	78
Figure 14.2.1.4.1:	DAS28(CRP): Plot of the actual values over time .....	84
Figure 14.2.1.4.2:	DAS28(CRP): Plot of the changes from baseline over time .....	84
Figure 14.2.1.4.3:	DAS28(CRP): Plot of the percent changes from baseline over time .....	84
Figure 14.2.1.5.1:	Joint counts: Plot of the actual values over time .....	86
Figure 14.2.1.5.2:	Joint counts: Plot of the changes from baseline over time .....	86
Figure 14.2.1.5.3:	Joint counts: Plot of the percent changes from baseline over time .....	86
Figure 14.2.1.6.1:	VAS scores: Plot of the actual values over time .....	88
Figure 14.2.1.6.2:	VAS scores: Plot of the changes from baseline over time .....	88
Figure 14.2.1.6.3:	VAS scores: Plot of the percent changes from baseline over time .....	88
Figure 14.2.1.7.1:	HAQ-DI total and domain scores: Plot of the actual values over time .....	90
Figure 14.2.1.7.2:	HAQ-DI total and domain scores: Plot of the changes from baseline over time .....	90
Figure 14.2.1.7.3:	HAQ-DI total and domain scores: Plot of the percent changes from baseline over time .....	90
Figure 14.2.1.8.1:	Serum CRP (mg/L): Plot of the actual values over time .....	93
Figure 14.2.1.8.2:	Serum CRP (mg/L): Plot of the changes from baseline over time .....	93

Figure 14.2.1.8.3: Serum CRP (mg/L): Plot of the percent changes from baseline over time .....	93
Figure 14.2.1.9.1: FACIT fatigue score: Plot of the actual values over time .....	95
Figure 14.2.1.9.2: FACIT fatigue score: Plot of the changes from baseline over time .....	95
Figure 14.2.1.9.3: FACIT fatigue score: Plot of the percent changes from baseline over time .....	95
Figure 14.2.1.10.1: SF-36: Plot of the actual values over time .....	102
Figure 14.2.1.10.2: SF-36: Plot of the changes from baseline over time .....	102
Figure 14.2.1.10.3: SF-36: Plot of the percent changes from baseline over time .....	102
Figure 14.2.3.1: Pharmacodynamics: YKL-40: Mean plot over time .....	104
Figure 14.2.3.2: Pharmacodynamics: YKL-40: Correlation plots versus efficacy .....	104
Figure 14.3.2.1: Laboratory data up to Week 12: Mean plots of the actual values over time .....	121
Figure 14.3.2.2: Laboratory data up to Week 12: Mean plots of the changes from baseline over time .....	121
Figure 14.3.2.3: Laboratory data up to Week 12: Mean plots of the percent changes from baseline over time .....	121
Figure 14.3.2.4: Laboratory data over the whole study: Subject profile plots over time – subjects with a treatment-emergent abnormality .....	123
Figure 14.3.2.5: Laboratory data over the whole study: Mean plots of the actual values over time .....	123
Figure 14.3.2.6: Laboratory data over the whole study: Mean plots of the changes from baseline over time .....	123
Figure 14.3.2.7: Laboratory data over the whole study: Mean plots of the percent changes from baseline over time .....	123

## LIST OF LISTINGS

Listing 16.2.1.1: Subject disposition: Randomization .....	44
Listing 16.2.1.2: Subject disposition: Region, country and site identification .....	44
Listing 16.2.1.3: Subject disposition: Number of days in study .....	45
Listing 16.2.1.4: Subject disposition: Trial termination .....	46
Listing 16.2.1.5: Subject disposition: Study analysis periods .....	46
Listing 16.2.2.1: Protocol deviations .....	47
Listing 16.2.2.2: Eligibility criteria: Violations .....	47
Listing 16.2.2.3: Final eligibility statements .....	47
Listing 16.2.3.1: Subjects excluded from the safety analysis .....	47
Listing 16.2.3.2: Subjects excluded from the efficacy ITT analysis .....	47
Listing 16.2.3.3: Subjects excluded from the efficacy per-protocol analysis .....	48
Listing 16.2.4.1: Demographic data .....	49
Listing 16.2.4.2: Prior RA medication .....	53
Listing 16.2.4.3: Medical history .....	53

Listing 16.2.4.4:	Concomitant diseases .....	53
Listing 16.2.4.5:	Prior and concomitant therapies .....	55
Listing 16.2.5.1:	Exposure to study medication: eCRF data.....	57
Listing 16.2.5.2:	Exposure to study medication: Derived data .....	57
Listing 16.2.5.3:	Dispensed kit numbers.....	57
Listing 16.2.5.4:	Exposure to methotrexate: eCRF data.....	59
Listing 16.2.5.5:	Exposure to methotrexate: Derived data .....	59
Listing 16.2.6.1:	ACR20, ACR50, ACR70 and ACR-N .....	73
Listing 16.2.6.2:	EULAR remission .....	74
Listing 16.2.6.3:	SDAI and CDAI .....	79
Listing 16.2.6.4:	DAS28(CRP): Original data .....	84
Listing 16.2.6.5.1:	Joint counts: Identification of unassessable joints .....	86
Listing 16.2.6.5.2:	Joint counts: Derived data.....	87
Listing 16.2.6.6:	VAS scores.....	88
Listing 16.2.6.7:	HAQ-DI: Derived summary scores.....	91
Listing 16.2.6.8:	Serum CRP (mg/L).....	93
Listing 16.2.6.9:	Total FACIT fatigue score.....	95
Listing 16.2.6.10:	SF-36: Derived summary scores.....	102
Listing 16.2.8:	Pharmacodynamics: YKL-40.....	105
Listing 16.2.10.1:	Treatment-emergent adverse events: Summary listing .....	111
Listing 16.2.10.2:	Pre-treatment adverse events: Summary listing .....	112
Listing 16.2.10.3:	Adverse events: Full listing of the serious adverse events .....	112
Listing 16.2.10.4:	Treatment-emergent adverse events: Full listing of the events leading to discontinuation.....	112
Listing 16.2.10.5:	Treatment-emergent adverse events: Full listing of the events leading to a temporary stop.....	112
Listing 16.2.10.6:	Adverse events: Coding information.....	113
Listing 16.2.10.7:	Adverse events: Listing of subjects for whom a narrative is required .....	113
Listing 16.2.11.1:	Laboratory data: Full listing.....	123
Listing 16.2.11.2:	Laboratory data: Abnormalities listing.....	123
Listing 16.2.11.3:	Laboratory data: Clinically significant abnormalities .....	123
Listing 16.2.11.4:	Laboratory data: Urinalysis results .....	123
Listing 16.2.11.5:	Laboratory data: Positive pregnancy tests .....	124
Listing 16.2.12.1:	ECG: Full listing .....	127
Listing 16.2.12.2:	ECG: Abnormalities listing .....	127
Listing 16.2.13.1:	Vital signs: Full listing .....	130
Listing 16.2.13.2:	Vital signs: Abnormalities listing.....	130
Listing 16.2.14:	Physical examinations: Abnormal findings .....	130

## 1. STUDY DESIGN

### 1.1. STUDY OBJECTIVES

#### 1.1.1. Primary Objective

The primary objective of the study is to evaluate the efficacy in terms of the percentage of subjects achieving an ACR20 response, of different doses and dose regimens of GLPG0634 compared to placebo at Week 12.

#### 1.1.2. Secondary Objectives

The secondary objectives of the study are:

- To evaluate the efficacy in terms of the percentage of subjects achieving an ACR20, ACR50, ACR70, ACR-N, DAS28(CRP), EULAR response and ACR/EULAR remission, CDAI, and SDAI with different doses and dose regimens of GLPG0634 compared to placebo at every visit.
- To evaluate the safety and tolerability of different doses and dose regimens of GLPG0634 in comparison with placebo.
- To characterize the population PK and PD of GLPG0634 and its metabolite (G254445) in subjects with rheumatoid arthritis and investigate the relationship between exposure and efficacy/safety/PD.
- To evaluate the effects of different doses and dose regimens of GLPG0634 administration on subjects' disability, fatigue, and quality of life.

## 1.2. STUDY DESIGN

This is a multicenter, phase IIb, double-blind, placebo-controlled add on study in subjects with moderately to severely active RA who have an inadequate response to MTX. A total of 595 subjects will be randomized to one of 6 dose regimens of GLPG0634 or to placebo, on top of their stable dose of MTX. Treatment duration will be 24 weeks. At Week 12, subjects on placebo who have not achieved a 20% improvement in swollen joint count (SJC66) and tender joint count (TJC68) will be re-randomized (automatically via interactive voice/web response [IXRS]) to treatment to receive GLPG0634 100 mg q.d. or 50 mg b.i.d. doses in a blinded fashion, subjects on 50 mg q.d. who have not achieved a 20% improvement in SJC66 and TJC68 will be assigned to 100 mg q.d. and subjects on 25 mg b.i.d. who have not achieved a 20% improvement in SJC66 and TJC68 will be assigned to 50 mg b.i.d. All will continue the study until Week 24. Subjects in the other treatment groups will maintain their randomized treatment until Week 24.

Upon qualification for the study, subjects will be randomized to placebo or one of 6 dosing regimens in a 1:1:1:1:1:1 ratio (N=85 for each treatment group; total N=595), stratified by geographic region and previous use of a biological DMARD during a single clinical study setting.

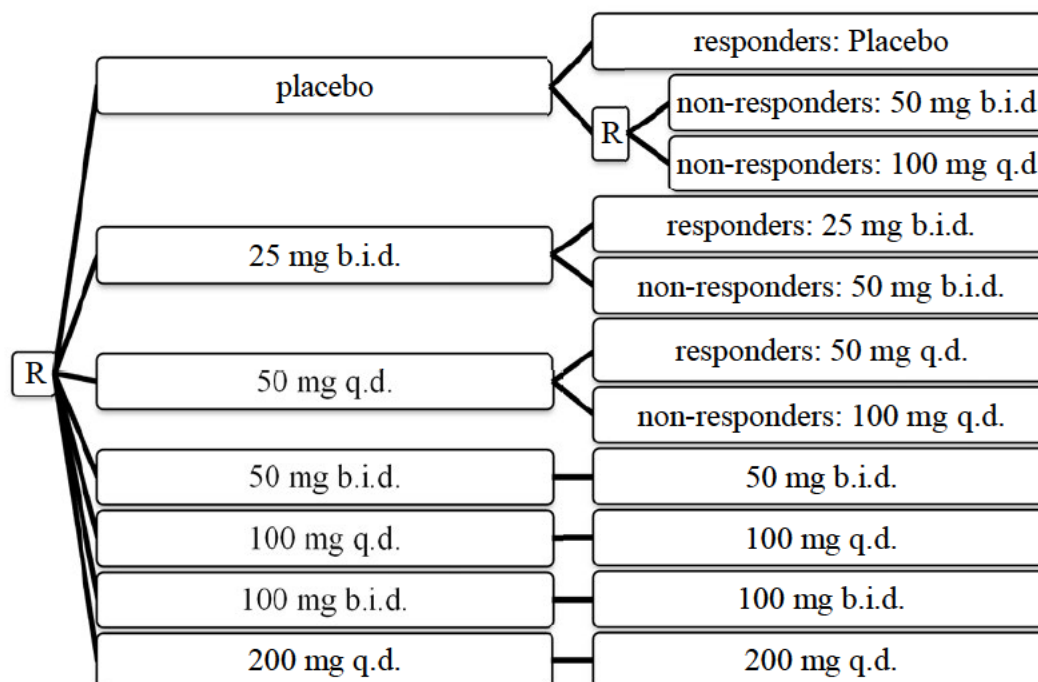
Subjects participating in the study will be requested to attend a maximum of 11 visits throughout the study: Screening visit (up to 28 days before Baseline visit), Baseline visit, Week 1 visit, Week 2 visit, Week 4 visit, Week 8 visit, Week 12 visit, Week 16 visit, Week 20 visit, Week 24 visit (and potentially a follow-up visit). At completion of the 24-week

treatment period, subjects will be offered the option to enter a Long Term Follow-up study (GLPG0634-CL-205). For the subjects not entering the GLPG0634-CL-205, an additional final Follow-up visit will be scheduled 7 to 10 days after end of study treatment.

Consequently, each subject will remain in the study for a maximum of 29 weeks (from Screening visit to Follow-up visit).

Data from the Long Term Follow-up study 205 will be presented in a separate analysis, and is not part of this SAP.

#### GLPG0634-CL-203 (DARWIN 1): add-on treatment to methotrexate:



#### Notes:

- R = (re)randomization.
- Responder: at least 20% drop in TJC68 and SJC66 versus baseline.

### 1.3. PROTOCOL AND PROTOCOL AMENDMENTS

Protocol Versions	Date (ddMMMyyyy)
Final 1	14MAR2013
Amendment 1	11APR2013
Amendment 2	17APR2013
Amendment 3	02AUG2013
Amendment 4	20NOV2013
Amendment 5	20MAY2014

This SAP was based on the latest version of the protocol (Amendment 5, 20MAY2014).

### 1.4. FLOWCHART

Event	Screening	Baseline <sup>1</sup>	Treatment period <sup>2</sup>									FU visit <sup>3</sup>
			D8/ W1	D15/ W2	D29/ W4	D57/ W8	D85/ W12	D113/ W16	D141/ W20	D169/ W24	EDV <sup>4</sup>	
Informed consent	Written informed consent will be obtained before any study-related procedures and/or assessments are performed.											
Inclusion/ exclusion criteria	X	X <sup>5</sup>										
Demographic data and baseline disease characteristics	X											
Medical history/ concomitant illnesses	X											
Physical examination	X <sup>6</sup>	X					X			X	X	X
Serology <sup>7</sup>	X											
TB test and chest X-ray <sup>8</sup>	X											
Pregnancy test <sup>9</sup>	X	X	X	X	X	X	X	X	X	X	X	X
12-Lead ECG	X						X			X	X	X
Vital signs <sup>10</sup>	X	X	X	X	X	X	X	X	X	X	X	X
Clinical laboratory tests <sup>11</sup>	X	X	X	X	X	X	X	X	X	X	X	X
PK blood samples <sup>12</sup>					X		X			X	X	
PD blood samples <sup>13</sup>		X	X		X		X			X	X	

<sup>1</sup> Subjects will begin to take their treatment on the next morning (Day 1) following randomization.

<sup>2</sup> During the treatment period, a visit time window of ± 2 calendar days is allowed.

<sup>3</sup> Follow-up visit will be performed only for subjects discontinuing prematurely from the study and for subjects not entering the Long Term Follow-up study GLPG0634-CL-205.

<sup>4</sup> Early discontinuation visit.

<sup>5</sup> Eligibility criteria check based on the laboratory results from the Screening visit.

<sup>6</sup> At Screening, includes height and weight.

<sup>7</sup> Includes Hepatitis B and C, HIV 1 and 2, RF, and anti-CCP.

<sup>8</sup> X-ray should be performed if results from an X-ray performed in the previous 3 months are not available at the site.

<sup>9</sup> For female subjects only. To be performed on serum at Screening and Week 24/EDV and on urine for other visits.

<sup>10</sup> Vital signs are defined as blood pressure (systolic and diastolic), HR, and oral temperature.

<sup>11</sup> Refer to protocol Section 9.7.1.4.2..

<sup>12</sup> PK samples: 1 blood sample (2 mL) for analysis of GLPG0634 and its main metabolite (G254445) in plasma.

<sup>13</sup> PD samples: 4 blood samples (2x4 mL for serum preparation and 2x2.5 mL in PAXgene tubes) will be collected for analysis of serum (analytes and miRNA profiling) and whole blood (gene expression in circulating leukocytes, mRNA and miRNA profiling).

Event	Screening	Baseline <sup>14</sup>	Treatment period <sup>15</sup>									FU visit <sup>16</sup>
			D8/ W1	D15/ W2	D29/ W4	D57/ W8	D85/ W12	D113/ W16	D141/ W20	D169/ W24	EDV <sup>17</sup>	
Randomization		X <sup>18</sup>					X <sup>19</sup>					
IXRS call	X	X	X	X	X	X	X	X	X	X	X	
Dispense study medication		X	X	X	X	X	X	X	X			
Subject diary card dispensation <sup>20</sup>		X	X	X	X	X	X	X	X			
Subject diary card collection			X	X	X	X	X	X	X	X	X	
Drug accountability check			X	X	X	X	X	X	X	X	X	
Study medication dosing												
Serum CRP	X	X	X	X	X	X	X	X	X	X	X	
SJC66	X	X	X	X	X	X	X	X	X	X	X	
TJC68	X	X	X	X	X	X	X	X	X	X	X	
Physician's Global Assessment		X	X	X	X	X	X	X	X	X	X	
Patient's Global Assessment	X	X	X	X	X	X	X	X	X	X	X	
HAQ-DI		X	X	X	X	X	X	X	X	X	X	
FACIT fatigue scale		X			X		X			X	X	
SF-36 questionnaire		X			X		X			X	X	
AE assessment												
Concomitant medications												

<sup>14</sup> Subjects will begin to take their treatment on the next morning (Day 1) following randomization.

<sup>15</sup> During the treatment period, a visit time window of ± 2 calendar days is allowed.

<sup>16</sup> Follow-up visit will be performed only for subjects discontinuing prematurely from the study and for subjects not entering the Long Term Follow-up study GLPG0634-CL-205.

<sup>17</sup> Early discontinuation visit.

<sup>18</sup> Upon qualification for the study.

<sup>19</sup> At Week 12, subjects on placebo who have not achieved a 20% improvement in SJC66 and TJC68 will be re randomized (automatically via IXRS) to treatment to receive GLPG0634 100 mg q.d. or 50 mg b.i.d. doses in a blinded fashion, subjects on 50 mg q.d. who have not achieved a 20% improvement in SJC66 and TJC68 will be assigned to 100 mg q.d. and subjects on 25 mg b.i.d. who have not achieved a 20% improvement in SJC66 and TJC68 will be assigned to 50 mg b.i.d. All will continue in the study until Week 24. Subjects in the other groups will maintain their randomized treatment until Week 24.

<sup>20</sup> Subject diary card will be dispensed to subjects on D -1 and at every following visit; subjects should be instructed to bring their diary card along with them to all visits.

## 2. ANALYSIS POPULATIONS AND TREATMENT GROUPS

### 2.1. ANALYSIS CHAPTERS

Each part of the analysis will consist of 2 or 3 main chapters:

1. Chapter 1: Period 1 of the study (up to and including Week 12): safety
2. Chapter 2: The study as a whole (periods 1 and 2 together, spanning the entire 24 weeks): safety and efficacy
3. Chapter 3: Period 2 of the study (from Week 12 until the end of the study): only for efficacy and adverse events.

Each analysis chapter will use its own set of “treatment groups” and imputation methods. Later in this document, when all parameters are analyzed separately, will be mentioned which is the primary.

#### 2.1.1. Efficacy data

Study period	Treatment groups	Analysis populations	Imputation method
“Chapter 2”: Periods 1+2 together (full 24 weeks period)	<p><u>The 7 original groups:</u></p> <ul style="list-style-type: none"> <li>– Placebo</li> <li>– GLPG0634 q.d. dosing groups:                             <ul style="list-style-type: none"> <li>• 50 mg q.d.</li> <li>• 100 mg q.d.</li> <li>• 200 mg q.d.</li> </ul> </li> <li>– GLPG0634 b.i.d. dosing groups:                             <ul style="list-style-type: none"> <li>• 25 mg b.i.d.</li> <li>• 50 mg b.i.d.</li> <li>• 100 mg b.i.d.</li> </ul> </li> </ul> <p>For the subjects in the placebo and low dose groups who switched to 100 mg/day during the second study period: they will be handled as if they discontinued at Week 12, and will be imputed from Week 12 onwards in the second study period.                      (Details: see further in the SAP)</p>	<p><u>ITT</u>: all subjects randomized, exposed at least once, and with post-baseline data for at least one efficacy parameter.</p> <p><u>Per-protocol</u>: subpopulation of ITT, excluding all critical (=major) protocol deviations. This is decided during the blind review meeting before final database lock.</p>	<p>Observed case, non-responder imputation (NRI), last observation carried forward (LOCF): depending on the parameter; is specified later when all parameters are discussed.</p>



Study period	Treatment groups	Analysis populations	Imputation method
<p>“Chapter 3”: Period 2, as of Week 12</p>	<p><u>The Week 12 responders, remaining on the same treatment during the second study period: 7 treatment arms:</u></p> <ul style="list-style-type: none"> <li>– Continued placebo</li> <li>– Continued GLPG0634 50 mg q.d.</li> <li>– Continued GLPG0634 100 mg q.d.</li> <li>– Continued GLPG0634 200 mg q.d.</li> <li>– Continued GLPG0634 25 mg b.i.d.</li> <li>– Continued GLPG0634 50 mg b.i.d.</li> <li>– Continued GLPG0634 100 mg b.i.d.</li> </ul> <p>This analysis will show whether Week 12 responders remain responders when treatment is continued unchanged.</p> <p><u>The Week 12 non-responders, either “switching” to 100 mg/day or continuing the same mid/high dose during the second study period:</u></p> <ul style="list-style-type: none"> <li>– Placebo non-responders switched to GLPG0634 100 mg q.d.</li> <li>– GLPG0634 50 mg q.d. non-responders switched to GLPG0634 100 mg q.d.</li> <li>– Continued GLPG0634 100 mg q.d.</li> <li>– Continued GLPG0634 200 mg q.d.</li> <li>– Placebo non-responders switched to GLPG0634 50 mg b.i.d.</li> <li>– GLPG0634 25 mg b.i.d. non-responders switched to GLPG0634 50 mg b.i.d.</li> <li>– Continued GLPG0634 50 mg b.i.d.</li> <li>– Continued GLPG0634 100 mg b.i.d.</li> </ul> <p>This analysis will show whether Week 12 non-responders start to respond when either treated for a longer period with the same mid/high dose, or have a dose increase to 100 mg/day.</p>	<p><u>ITT</u>: all subjects randomized, exposed at least once, and with post-Week 12 data for at least one efficacy parameter. Subjects who discontinued prior the Week 12 visit will be excluded from this analysis.</p>	<p>Observed case, non-responder imputation (NRI), last observation carried forward (LOCF): depending on the parameter; is specified later when all parameters are discussed.</p>

## 2.1.2. Safety data

<b>Study period</b>	<b>Treatment groups</b>	<b>Analysis populations</b>	<b>Imputation method</b>
“Chapter 1”: Period 1, up to Week 12	<u>The 7 original groups:</u> <ul style="list-style-type: none"><li>– Placebo</li><li>– GLPG0634 q.d. dosing groups:<ul style="list-style-type: none"><li>• 50 mg q.d.</li><li>• 100 mg q.d.</li><li>• 200 mg q.d.</li></ul></li><li>– GLPG0634 b.i.d. dosing groups:<ul style="list-style-type: none"><li>• 25 mg b.i.d.</li><li>• 50 mg b.i.d.</li><li>• 100 mg b.i.d.</li></ul></li></ul>	<u>Safety:</u> all subjects randomized and exposed at least once.	Observed-case only (no imputation)

Study period	Treatment groups	Analysis populations	Imputation method
<p>“Chapter 2”: Periods 1+2 together (full 24 weeks period)</p>	<p>The 11 groups of the second study period, taking into account treatment switches:</p> <ul style="list-style-type: none"> <li>– Subjects starting on placebo                             <ul style="list-style-type: none"> <li>• Continued placebo</li> <li>• Switched to 100 mg q.d.</li> <li>• Switched to 50 mg b.i.d.</li> </ul> </li> <li>– GLPG0634 continued dosing in responders                             <ul style="list-style-type: none"> <li>• 50 mg q.d.</li> <li>• 25 mg b.i.d.</li> </ul> </li> <li>– GLPG0634 continued dosing in non-responders                             <ul style="list-style-type: none"> <li>• 50 mg q.d. increased to 100 mg q.d.</li> <li>• 25 mg b.i.d. increased to 50 mg b.i.d.</li> </ul> </li> <li>– GLPG0634 continued q.d. dosing groups                             <ul style="list-style-type: none"> <li>• 100 mg q.d.</li> <li>• 50 mg b.i.d.</li> </ul> </li> <li>– GLPG0634 continued b.i.d. dosing groups                             <ul style="list-style-type: none"> <li>• 200 mg q.d.</li> <li>• 100 mg b.i.d.</li> </ul> </li> </ul> <p>For the subjects in the placebo and low dose groups who discontinued prior to Week 12: they will be included in the “continued placebo”, “continued 25 mg b.i.d.” and “continued 50 q.d.” groups.</p> <p>Note that AEs will be presented overall, but also split up in the two study periods.</p>	<p><u>Safety</u>: all subjects randomized and exposed at least once during the study.</p>	<p>Observed-case only (no imputation)</p>

## **2.2. FURTHER ANALYSIS POPULATIONS**

### **2.2.1. All Screened Subjects Population**

All subjects who signed an informed consent form.

### **2.2.2. All Randomized Population**

All subjects randomized into the study.

## **2.3. HANDLING OF UNBLINDED SUBJECTS**

### **2.3.1. Safety population**

In case a particular subject would be unblinded (e.g., in case of an SAE), all safety data of this subject recorded in the database will still be included in the safety analysis.

### **2.3.2. Intent to Treat Population**

In case a particular subject would be unblinded (e.g., in case of an SAE), all efficacy data of this subject up to the point of unblinding will still be included in the ITT efficacy analysis. Recorded efficacy data from the moment of unblinding onwards will be disqualified from the analysis. Such data will still be part of the listings, but will be labeled as use=no. Missing data after the moment of unblinding will be imputed using the same rules as used for other subjects.

### **2.3.3. Per Protocol Population**

In case a particular subject would be unblinded (e.g., in case of an SAE), all efficacy data of this subject up to the point of unblinding will still qualify for the per-protocol efficacy analysis selection according to the above rule of excluding critical protocol deviations. Recorded efficacy data from the moment of unblinding onwards will be disqualified from the analysis. Such data will still be part of the listings, but will be labeled as use=no. Missing data after the moment of unblinding will be imputed using the same rules as used for other subjects.

## **2.4. TREATMENT GROUPS**

### **2.4.1. Randomized versus Actual Treatment**

For efficacy parameters, the treatment group as assigned by the randomization will be used in the analysis (i.e., as-randomized analysis, considered closest to a worst-case intention-to-treat principle).

For safety parameters, the treatment that was actually used by the subject will be applied in the analysis (i.e., as-treated analysis, considered the fairest worst-case method to allocate potential safety problems to the treatment that was actually used). Such an as-treated analysis will only be done in case more than 20% of the actually used study drug was different than planned and is in a higher dose than planned. If the subject uses less than 20% of incorrect kits or the incorrect kit is a lower dose, then an as-randomized analysis will be applied. The

“20%” mentioned here is over the related study period: period 1 only, or full periods 1+2 combined, or period 2 only (for AE), depending on the analysis chapter.

During the second part of the study: if the subject passed the re-randomization visit (Week 12) but it is unknown whether the subject actually used any study drug during the second part of the study, then it is assumed that the subject used at least one dose of study drug during the second part. If the subject passed the re-randomization visit and it is known that the subject did not take any study drug during the second part of the study, then the re-randomization will be ignored and the subject will only have a study part 1 in the analysis.

Differences between as-treated and as-randomized will be flagged in the listing on subject allocation.

## **2.4.2. Totals across Groups**

A total across all groups will be presented for the General (subject disposition, demographics, etc) part of the analysis. For the Safety parts of the analysis, a total across all GLPG0634 groups will be presented for data up to and including Week 12, and for data after Week 12 or data across the entire 24 weeks of the study a total across all groups will be presented. No totals will be presented for PD or Efficacy. Totals will only be shown on tables, but not on listings or figures.

# **3. ANALYSIS PERIODS AND ANALYSIS TIME POINTS**

## **3.1. RELATIVE NUMBER OF DAYS**

The relative day (DY) is calculated as follows:

= Visit date – reference date + 1 day, when the visit date is on or after the reference date

= Visit date – reference date, when the visit date is before the reference date.

The reference date in the study is the first study medication intake date, which by definition has DY=1. There is no DY=0.

## **3.2. ANALYSIS PERIODS FOR NON-VISIT DATA**

These analysis periods are to be used for allocation of adverse events and concomitant medications into periods.

Analysis period	Period start date	Period end date
Screening	Date of signing the initial ICF.	First study drug administration date – 1 day.
First 12 weeks	First study drug administration date.	Date of the W12 visit.
Second 12 weeks	Date of the W12 visit + 1 day.	Max ( trial termination date ; date of last dose + 30 days ).
Full 24 weeks	First study drug administration date.	Max ( trial termination date ; date of last dose + 30 days ).

The trial termination date is the date of last contact with the subject, as recorded in the eCRF.

Note that the last analysis period in case of early termination will always be ended by: *Maximum ( trial termination date ; date of last dose + 30 days)*. In case the subject rolls over to DARWIN3, the trial termination date will be used (no 30 days will be added since that would be during DARWIN3).

### 3.3. ALGORITHM OF ALLOCATING VISITS TO TIME WINDOWS

All visits (including early discontinuation visits and unscheduled visits but excluding visits without data) will be placed into time windows according to their relative day (DY) in the study, according to the following allocation table:

Time point label	Target day	Interval lower bound	Interval upper bound
Screening (3)	NA	NA	NA
Baseline (1)	NA	NA	NA
Week 1	8	2	11
Week 2	15	12	21
Week 4	29	22	42
Week 8	57	43	70
Week 12	85	71	98
Week 16	113	99	126
Week 20	141	127	154
Week 24	169	155	182
>Week 24 (2)	NA	183	+∞
Follow-up (3)	7-10 days after last visit	NA	NA

(NA = not applicable).

**Notes:**

- (1) The actual baseline reference value will be determined per parameter as the last available pre-dosing data point, so might differ from this “baseline” visit interval. There will be only one baseline time point in this study: the last nonmissing value prior to first dosing of study drug.
- (2) Visits falling really late in the study (relative day > 182; excluding the follow-up visit) will be allocated to a “>Week 24” interval, and will not be shown in tables nor figures, but will only be listed.
- (3) For the screening and follow-up visits, no time window is defined. The visit as recorded in the eCRF will be used. Follow-up visits of subjects who discontinue the study prior to the Week 12 re-randomization will be shown separately as well in the tables that present the first part of the study only. For tables that present the study as a whole, all follow-up visits will be shown.

Tables, figures and listings will present the time points, not the visits.

For ECGs, a different method will be used. See section 12.3 for the details.

### **3.4. SELECTION OF VISITS**

It is possible that more than one visit gets allocated into the same time window. In that case, only one visit will be selected for Tables and figures. The nonselected visit(s) will only be listed and flagged as use=no.

The visit with a relative day (DY) closest to the target day will be selected. If there are multiple visits at the same distance of the scheduled visit day (meaning: equal  $ABS(DY - \text{target day})$ ), then the one latest in time is selected.

In case of an additional screening result, the last pre-baseline measurement is selected as screening value for analysis. This value (not the original screening value) was also used in the clinical center to include the subject in the study.

In case more than one parameter is measured per time point (e.g. for lab), the selection is performed per parameter and per time point, not per “sample” and per time point. Missing values are removed before the selection is made.

A missing baseline value will be imputed with the last nonmissing value of a preceding visit. This can be a screening visit or an unscheduled visit prior to the first dose of study medication; whichever comes latest. If there is no nonmissing result at such screening or unscheduled visits, then the missing baseline value remains missing.

In case the screening visit is used to impute the baseline, the visit will be duplicated in the derived dataset. The original screening visit as well as the imputed baseline will be presented in the tables. In the listings, only original screening data will be shown: the original screening visit will be shown and flagged as “reference” visit. The imputed baseline will not be shown.

In case an unscheduled visit is used to impute the baseline, the visit will be renamed to “baseline” in the derived dataset. The imputed baseline will be presented in the tables. In the listings, the original unscheduled data will be shown and flagged as “reference” visit.

## **4. HANDLING OF DATA**

### **4.1. CALCULATION OF DESCRIPTIVE STATISTICS**

For continuous parameters, descriptive statistics will be presented when  $N \geq 2$ . When  $N=1$ , the observation will not be shown in the table but only in the listing.

Descriptive statistics will include at least the following:

- the number of nonmissing data points (N)
- the arithmetic mean
- the standard error (SE)
- the median, minimum and maximum
- 95% confidence interval of the mean (only when requested).

## **4.2. CALCULATION OF PERCENTAGES**

Missing values will not be included in the denominator count when computing percentages.

Imputed missing values are no longer considered to be missing values.

## **4.3. HANDLING OF VALUES BELOW (OR ABOVE) A THRESHOLD**

### **4.3.1. Safety Data**

Values below (above) the detection limit will be imputed by the value of the detection limit itself. Listings will always present the original value.

Example: if the database contains values like “<0.04”, then for the descriptive statistics the value of the detection limit (0.04) shall be used. A value like “>1000” will be imputed by “1000”.

### **4.3.2. PK Data**

Values below the detection limit will be imputed by a zero.

### **4.3.3. PD Data**

Values below the detection limit will be imputed by half of the value of the detection limit itself. Listings will always present the original value.

## **4.4. HANDLING MISSING DATA**

### **4.4.1. Handling of Missing Date (Time) or Partially Known Date (Time)**

No imputation will be done of missing date (time) fields, nor of the missing parts of partially known date (time) fields. When needed, a worst-case selection will be made. Such worst-case selection rules are elaborated in the specific sections (e.g., AE).

### **4.4.2. Handling of missing values in scales or summary parameters**

Missing items in scales or summary parameters will be imputed to allow for the derivation of the total scale/parameter score. The appropriate imputation rules will be described in the sections where the total scores are defined.

### **4.4.3. Handling of missing values and early dropouts in the efficacy analyses during the first study period (up to Week 12)**

This is for the “efficacy chapter 2” analysis.



For binary data (e.g., ACR20), two different imputations will be done for missing visits as well as early dropouts: non-responder imputation (NRI) and LOCF, up to and including Week 12. The NRI imputation will be primary, and LOCF will be secondary.

For other data (ordinal or continuous), missing visits as well as early dropouts will be imputed using the last observation carried forward (LOCF) algorithm, up to and including Week 12.

Definitions of the imputation methods:

- LOCF (all types of parameters, from binary to continuous): to impute a missing value, the last preceding nonmissing value will be used (which could even be a baseline value).
- NRI (binary response parameters only): to impute a missing response, the subject will be assumed to be a non-responder.

For all parameters, an observed-case analysis will also be performed (i.e., without any imputation). The observed-case analysis will serve as a sensitivity analysis (to ascertain the impact of early dropouts on the results) and will basically be limited to:

- ACR20/50/70/N: descriptive at W12+W24 and inferential at W12
- EULAR remission: descriptive at W12+W24
- SDAI/CDAI: descriptive at BSL+W12+W24
- DAS28: descriptive at BSL+W12+W24 and inferential at W12
- All ACR components (VAS, joint counts, CRP, HAQ-DI): descriptive at BSL+W12+W24
- QoL (FACIT,SF-36): descriptive at BSL+W12+W24.

#### **4.4.4. Handling of missing values and early dropouts in the efficacy analyses during the study period as a whole (full 24 weeks)**

This is for the “efficacy chapter 2” analysis.

For binary data (e.g., ACR20), two different imputations will be done for missing visits as well as early dropouts: non-responder imputation (NRI) and LOCF, up to and including Week 24. The NRI imputation will be primary, and LOCF will be secondary.

For other data (ordinal or continuous), missing visits as well as early dropouts will be imputed using the last observation carried forward (LOCF) algorithm, up to and including Week 24.

For all parameters, an observed-case analysis will also be performed (i.e., without any imputation). This observed-case will not be shown for subjects who switched treatment. See previous section for the parts where the observed-case analysis will be presented.

#### **4.4.5. Handling of missing values and early dropouts in the efficacy analyses during the second study period (after Week 12)**

This is for the “efficacy chapter 3” analysis.

Subjects in this analysis must have data after the Week 12 potential switch in dosing. Subjects who discontinued the study during the first period of the study will not be included in this analysis and will therefore also not be imputed.

For binary data (e.g., ACR20), two different imputations will be done for missing visits as well as early dropouts during period 2: non-responder imputation (NRI) and LOCF, up to and including Week 24. The NRI imputation will be primary, and LOCF will be secondary.

For other data (ordinal or continuous), missing visits as well as early dropouts during period 2 will be imputed using the last observation carried forward (LOCF) algorithm, up to and including Week 24.

Note that no Period 1 efficacy data will be carried forward into this chapter 3 analysis.

For all parameters, an observed-case analysis will also be performed (i.e., without any imputation).

#### **4.4.6. Handling of missing values and early dropouts in the safety analysis**

No imputation is done of missing values or missing visits. The analysis will be observed-case only.

### **4.5. HANDLING TREATMENT SWITCHES AT WEEK 12 IN THE EFFICACY AND SAFETY ANALYSIS**

This section only applies to the Week 12 non-responders in the placebo and low dose groups, where there is a dose increase for the remaining 12 weeks until the end of the study, and also to the Week 12 responders in the placebo and low dose groups, who remain on the same treatment for the remaining 12 weeks until the end of the study.

#### **4.5.1. In the efficacy analysis of the full study as a whole (full 24 weeks)**

The efficacy analysis will apply the original 7 treatment groups (original randomization at baseline). Subjects who switch treatment at Week 12 will be regarded as having discontinued treatment at Week 12: their data after Week 12 will not be used in this analysis. Their efficacy data will be handled for the remainder of the study as described above in section 4.4.4. The placebo, 50 mg q.d. and 25 mg b.i.d. groups will thus contain both data under this treatment (continued treatment groups), as well as Week 12 imputed information from non-responders who switch to a higher dose of GLPG0634. It is recognized that these 3 groups (placebo, 50 mg q.d. and 25 mg b.i.d.) will only be an approximation, so the formal comparisons between all groups after Week 12 should be interpreted with care. LOCF, NRI and observed-case will be produced.

#### **4.5.2. In the safety analysis of the full study as a whole (full 24 weeks)**

The treatment groups that will be presented in the tables will distinguish treatment switches at Week 12 as separate arms: the period 1 & 2 treatment sequence will be used.

#### **4.5.3. In the efficacy and safety analysis of the second study period**

The treatment arms will be presented descriptively only, as separate columns.

In the safety analysis, this only applies to the adverse events where the treatment-emergent AEs that emerge during the second part of the study will be presented separately. The lab data, ECG and vital signs are already presented in chapter 2 of the analysis.

#### **4.6. HANDLING OF OUTLIERS**

All measured values will be included in the analyses.

#### **4.7. HANDLING OF SECONDS IN DATETIME FIELDS**

If a datetime field contains seconds, these will be cut off (i.e., rounded down to the minute) prior to data analysis. The analysis will only use the date and time up to the minute.

#### **4.8. STRATIFICATION FACTORS**

In case there is a discrepancy in randomization stratification stratum between IWRS and the actual eCRF data, then:

1. The actual eCRF stratum will be used in analyses and analysis models, and
2. Any such discrepancy will be flagged in the listing on randomizations (listing 16.2.1.1).

### **5. SOFTWARE AND PROCEDURES**

#### **5.1. SOFTWARE**

SAS version 9.3 (or higher) will be used for programming.

#### **5.2. PROCEDURES**

Analyses will comply with ICH regulations, in particular: (ICH-E3) , (ICH-E6) and (ICH-E9).

The following Galapagos SOPs will be followed:

- SOP-CLI-001: Developing Clinical Study Documents (version 1.0)
- SOP-CLI-003: Managing Data Processing Activities (version 1.0)
- SOP-CLI-004: Managing DSMB Activities (version 1.0)

The following ████ STAT SOPs will be followed:

<b>Location</b>	<b>Document Type</b>	<b>Document Number</b>	<b>Effective Date</b>	<b>Document Title</b>
Global	SOP	03.002.01	03-Feb-2014	Providing Statistical Input to CSRs and Clinical Summaries for Submission Dossiers
Global	SOP	03.003.00	06-Feb-2012	Generation of define.xml
Global	SOP	03.004.01	03-Feb-2014	Statistical Input to Clinical Trial Protocols
Global	SOP	03.005.00	29-Mar-2012	Planning and Conducting Interim Analyses
Global	SOP	03.007.01	01-Oct-2013	Development of a Statistical Analysis Plan
Global	SOP	03.009.01	01-Oct-2013	Quality Control of SAS® Programming
Global	SOP	03.010.01	01-Oct-2013	Development of SAS® Programs
Global	SOP	03.011.00	30-Apr-2012	Randomization and Unblinding
Global	SOP	03.012.02	02-Dec-2013	Biostatistics and Statistical Programming Directory Structure Creation and Setup
Global	SOP	03.013.01	01-Oct-2013	Development and Production of Analysis Datasets
Global	SOP	03.014.01	29-Apr-2014	Biostatistics and Statistical Programming Functional Lead Responsibilities
Global	SOP	03.015.01	29-Apr-2014	Performing Blind Review of Data

### **5.3. FORMATS**

Derived data will be following CDISC ADaM 2.1 (ADaMIG 1.0) format.

Tables, listings and figures will follow the Mock TLFs, as provided in a separate document.

## **6. STATISTICAL METHODS**

### **6.1. PLANNED ANALYSES, PROTOCOL AMENDMENTS INCLUDED**

#### **6.1.1. Demographic and other baseline characteristics**

Subject disposition including the reasons for study termination will be tabulated.

Demographic characteristics will be listed. Demographic characteristics will be summarized using appropriate descriptive statistics, as applicable.

## 6.1.2. Efficacy Variables

### 6.1.2.1. Definition of primary efficacy endpoint: ACR20 at Week 12

The primary endpoint is the percentage of subjects achieving an ACR20 response at Week 12. Other time points will be regarded as secondary endpoints. The non-responder imputation (NRI) of the ACR20 will be the primary parameter, with observed-case and last observation carried forward (LOCF) imputation as secondary confirmatory analyses.

The ACR response is a measurement of improvement in multiple disease assessment criteria. The ACR20 response is defined as:

≥20% improvement from baseline in SJC66 (66 joints) and TJC68 (68 joints)

AND

≥20% improvement from baseline in at least 3 of the following 5 assessments

- Pain (VAS) in cm (from HAQ-DI).
- Patient's Global Assessment of Disease Activity (VAS) in cm.
- Physician's Global Assessment of Disease Activity (VAS) in cm.
- Patient's Assessment of Physical Function as measured by HAQ-DI.
- CRP in mg/L.

### 6.1.2.2. Definition of secondary efficacy endpoints

#### ACR50 and ACR70

ACR50 and ACR70 are similarly defined as ACR20, except the improvement threshold from baseline is 50% and 70%, respectively.

#### ACR-N

The ACR-N is the smallest percentage improvement in swollen and tender joints and the median of the remaining 5 core parameters, and is expected to be more sensitive to change than the ACR20, ACR50 or ACR70. It is a number varying between 0 and 100.

ACR-N = MIN [ %improvement in TJC68, %improvement in SJC66, MED %improvement in (patient's global assessment, physician's global assessment, patient's assessment of pain, HAQ-DI, CRP) ], cut off at 0 in case it would be a negative number.

#### ACR/EULAR Remission

A subject's disease activity status can be defined as being in remission if:

when scores on the TJC28, SJC28, CRP (actual value in mg/dL) and Patient Global Assessment of Disease Activity (cm) are all ≤1.

Simplified Disease Activity Index (SDAI)

The SDAI is the numerical sum of 5 outcome parameters: TJC28, SJC28, Patient Global Assessment of Disease Activity (in cm), Physician's Global Assessment of Disease Activity (in cm), and CRP (mg/dL).

$$\text{SDAI} = \text{TJC28} + \text{SJC28} + \text{Patient's Global Assessment of Disease Activity (VAS in cm)} + \text{Physician's Global Assessment of Disease Activity (VAS in cm)} + \text{CRP (mg/dL)}$$

The SDAI can be categorized:

- High disease activity:  $\text{SDAI} > 26$
- Moderate disease activity:  $]11, 26]$
- Low disease activity:  $]3.3, 11]$
- Remission:  $\leq 3.3$

Clinical Disease Activity Index (CDAI)

The CDAI is the SDAI modified to exclude CRP.

$$\text{CDAI} = \text{TJC28} + \text{SJC28} + \text{Patient's Global Assessment of Disease Activity (VAS in cm)} + \text{Physician's Global Assessment of Disease Activity (VAS in cm)}$$

The CDAI can also be categorized:

- High disease activity:  $> 22$
- Moderate disease activity:  $]10, 22]$
- Mild disease activity:  $]2.8, 10]$
- Remission:  $\leq 2.8$

Disease Activity Score 28 Joints Based on CRP (DAS28[CRP])

The DAS28(CRP) is a statistically derived index combining tender joints (28 joints), swollen joints (28 joints), CRP and Patient's Global Assessment of Disease Activity (general health, GH).

DAS28(CRP) is defined as follows:

$$\text{DAS28(CRP)} = 0.56 \times \text{SQRT}(\text{TJC28}) + 0.28 \times \text{SQRT}(\text{SJC28}) + 0.36 \times \text{Ln}(\text{CRP}+1) + 0.014 \times \text{GH} + 0.96,$$

Where:

- TJC28 is 28 joint count for tenderness
- SJC28 is 28 joint count for swelling.
- $\text{Ln}(\text{CRP}+1)$  is the natural logarithm of (CRP value [mg/L] + 1)
- SQRT is square root
- GH is the Patient's Global Assessment of Disease Activity on a 100 mm VAS

Categorization of the DAS28(CRP) scores:

- High disease activity:  $> 5.1$
- Moderate disease activity:  $] 3.2, 5.1 ]$
- Low disease activity:  $[ 2.6, 3.2 ]$
- Remission:  $< 2.6$

### EULAR Response

A second categorization of the DAS28(CRP) will be done according to the following table:

Actual DAS28(CRP)	Improvement in DAS28(CRP) from Baseline		
	$> 1.2$	$]0.6, 1.2]$	$\leq 0.6$
$\leq 3.2$	Good	Moderate	None
$]3.2, 5.1]$	Moderate	Moderate	None
$> 5.1$	Moderate	None	None

### Quality of Life: FACIT Fatigue Scale and SF-36

Quality of life will be assessed using the FACIT fatigue scale and the SF-36 questionnaire. The appropriate (sub)totals will be derived according to the scale's scoring algorithm.

#### **6.1.2.3. Methods of analysis**

Efficacy data (ACR20, ACR50, ACR70, ACR-N, DAS28(CRP), EULAR response, and ACR/EULAR remission, and components of the ACR, CDAI and SDAI at each post dosing visit) will be analyzed descriptively.

The analysis will present the results in 2 chapters:

1. Chapter 2: Analysis of the full 24 weeks study period.
2. Chapter 3: Analysis of the second study period, from Week 12 onwards.

All chapters will use the original baseline visit as reference.

Handling of discontinued subjects during the 24 weeks treatment period:

- Subjects who discontinue within the first week so that ACR20/50/70 responses cannot be determined will be classified as non-responders.
- For all discontinued subjects, visits after their discontinuation will be imputed for a duration depending on the analysis chapter:
  - Chapter 2: subjects discontinuing during the first study period will have their remaining visits imputed up to and including Week 12.
  - Chapter 2: subjects discontinuing the study will be imputed up to and including Week 24, irrespective of whether they discontinued during the first or second study period.
  - Chapter 3: subjects discontinuing during the first study period will not be imputed during the second study period. Subjects discontinuing during the second study period will be imputed up to and including Week 24.

The method of imputation will depend on the parameter: NRI and LOCF for binary data; LOCF for ordinal or continuous data. Observed-case will always be presented as well.

Non-responding subjects from the placebo and lowest dose groups who are re-randomized at Week 12 will be handled as follows, again depending on the analysis chapter:

- Chapter 2: their original treatment group of the first study period will be used. The Week 12 value will be carried forward to all further time points (and between-group comparisons) after Week 12.
- Chapter 3: their actual assessments on Weeks 16-20-24 will be analyzed descriptively using the new treatment groups.
- See also: overview table, section 2.1.

Between-group comparisons will be done for each dose group versus the placebo group. Hommel's closed-testing correction procedure (as implemented in SAS PROC MULTTEST) will be applied to adjust for multiplicity.

This study is not powered for any formal comparison among the dose groups, nor for comparisons between q.d. and b.i.d. regimens. However, differences between q.d. and b.i.d. regimens will be calculated and presented with a 95% confidence interval. Differences between q.d. and b.i.d. regimens will also be tested exploratively via a model containing the total daily dose (50/100/200 mg) as well as the regimen (q.d./b.i.d.). This model will not use the placebo data. Differences among the q.d. groups and among the b.i.d. groups will be described using the same model. No adjustment for multiplicity will be done for these exploratory differences.

Analysis methods per post-baseline time point:

- Binary parameters (such as ACR20): logistic regression model with factors treatment, region and previous use of biologics.
- Continuous parameters: both the changes from baseline and the percent changes from baseline will be described. Formal between-group comparisons will be performed on the changes from baseline only, using an analysis of (co)variance (AN[C]OVA) model with factors treatment, baseline value, region and previous use of biologics. Baseline will be the last nonmissing value prior to first dosing in the study.
- Time to (first) response (ACR20/50/70) will be analyzed using Kaplan-Meier survival techniques, and groups will be compared against placebo using a Cox proportional hazards regression model with factors treatment, region and previous use of biologics.

Quality of Life data (FACIT fatigue scale and SF-36 questionnaire) will be analyzed descriptively. Exploratory between-group comparisons will also be done at each post-baseline time point using AN(C)OVA models on the changes from baseline with factors treatment, baseline value, region and previous use of biologics.

A descriptive sensitivity analysis by baseline characteristic (such as country, geographic region, previous use of biological DMARDs etc) will be performed for the ACR20 response and the change in DAS28(CRP). Additionally an exploratory regression model will be fitted using a stepwise selection algorithm on the ACR20 response and the change in DAS28(CRP) at Week 12, using treatment and baseline characteristics as explanatory factors to investigate which baseline factors have an influence on the response.

Graphical presentations are added to facilitate the overall interpretation of the study results.



### 6.1.3. Safety Variables

Clinical safety will be evaluated by assessing treatment-emergent AEs, physical examinations, laboratory assessments, ECG, and vital signs results in a descriptive manner. Original results, changes and percent changes from baseline (or from screening, depending on the parameter) will be summarized for the laboratory data, vital signs, and ECG values. Values will be categorized as low/normal/high according to normal ranges, and shift tables versus baseline will be created to determine treatment-emergent abnormalities.

The analysis will present the results in 2 chapters:

1. Chapter 1: Analysis of the first study period, up to and including the Week 12 visit.
2. Chapter 2: Analysis of the full 24 weeks study period.

Non-responding subjects from the placebo and lowest dose groups who are re-randomized at Week 12: the actual treatments will be presented. (See also: overview table, section 2.1.)

### 6.1.4. Pharmacokinetic Analyses

This will be elaborated in a separate SAP by [REDACTED].

### 6.1.5. Pharmacodynamic Analyses

Serum levels of YKL-40 will be described as actual values, changes from baseline and percent changes from baseline. Between-group comparisons will be performed using the same ANCOVA models as used for the efficacy data. To further explore the data, scatterplots of the percent change from baseline in YKL-40 values versus efficacy parameters will be created.

Further analyses may be done at a later time, but these will not be part of the CSR.

### 6.1.6. Pharmacokinetic, Efficacy, Safety and Pharmacodynamic Correlations

This will be elaborated in a separate SAP by [REDACTED].

## 6.2. INTERIM ANALYSIS

Note that a DSMB is installed to monitor unblinded safety data during the course of the study. This is described in a separate DSMB Charter and DSMB SAP.

An unblinded interim analysis will be performed when all subjects have reached Week 12 (or have discontinued before Week 12). Only data from this study period will be included in the interim analysis.

The analysis will not present everything, but only a selection of parameters:

- General: region and county distribution, trial termination reasons, demographics and baseline disease characteristics.
- Efficacy: ACR20/50/70, ACR-N, SDAI and CDAI, DAS28(CRP), the ACR components (joint counts, CRP, HAQ-DI and the 3 VAS scores).
- Safety: adverse events, laboratory safety and vital signs.

The exact tables and figures to be produced for this interim analysis are identified in the next sections where all the TLFs are defined. The interim analysis is intended to support preliminary dose selection for the GLPG0634 phase III program.

Because this efficacy analysis consists of parts of the full Week 12 primary analysis, there is no inflation of the type-I error rate, and therefore no statistical correction is applied.

To preserve the blind for Team members who have access to raw patient data, the following approach will be used:

The statistical analysis of the unblinded data will be handled by a statistician at [REDACTED] who is independent from the regular study team. The regular study team (including [REDACTED] lead statistician and corresponding [REDACTED] biometrics team) will remain blinded.

This interim analysis output will comprise summary tables and plots, but no listings. These summary results will be viewed by the sponsor, but no individual subject data (such as listings or datasets) will be exchanged.

To prevent accidental unblinding of specific subjects, the following rules will also be applied:

- For descriptive statistics: min-max lines will be suppressed.
- For frequency tables containing small numbers of subjects in some categories (e.g., trial termination reasons, demographics/baseline characteristics, AEs & SAEs, and lab & vital signs abnormalities), the following rules will be programmed in the tables:
  - If there are events in multiple groups, including the placebo group: the full details will be shown in the table, per treatment group. This does not cause any unblinding of specific subjects.
  - If there are events in only one of the groups (either a GLPG0634 dose group or the placebo group): this would unblind these subjects. To prevent this, the individual groups will be left blank in the frequency table, and only the “overall total” group will be populated.
  - If there are events in multiple groups, but none in the placebo group: this would partially unblind these subjects, indicating that all of them received GLPG0634. To prevent this, the individual groups will be left blank in the frequency table, and only the “overall total” group will be populated.

The analysis TLFs will be checked for potential unblinding details by the [REDACTED] independent statistician prior to delivery to the sponsor.

## **6.3. CHANGES TO THE PLANNED ANALYSES, NOT COVERED BY PROTOCOL AMENDMENTS**

### **6.3.1. Changes before Interim Database Lock**

No changes to the protocol-specified analyses were made. The SAP provides further elaborations, clarifications and specifications on the protocol-specified analyses.

Version changes after 1.0:

Change from 1.0 to 3.0 SAP and Mocks due to technical issues in Galapagos' electronic document management system (Plaza) without any change to document contents.

Change from 3.0 to 4.0 SAP and Mocks: corrections of the actual contents of the interim analysis (inconsistencies within the SAP and with DARWIN2). Also solving inconsistencies between SAP and Mocks, and between DARWIN1 and DARWIN2, and providing more clarification on the expected observed-case analysis.

### **6.3.2. Changes after Interim Database Lock but before Final Database Lock**

Changes from 4.0 to 5.0 SAP:

Corrections:

- The analysis periods were updated from the interim SAP: the actual Week 12 visit date day is now included in the first 12 weeks of the study period. In the interim SAP it was mistakenly excluded from the first 12 weeks of the study.
- Randomization stratification factors: it is now more clearly specified that in case of discrepancies between IWRS and eCRF, the real underlying eCRF stratum will be used in the analyses and analysis models. This was already so in the interim analysis, but possibly not clear enough in this SAP.
- For safety data, the overall total was replaced by a total across the GLPG0634 groups only. This will allow easier comparison of the overall GLPG0634 safety versus the placebo group.
- Baseline HAQ-DI, Baseline SDAI and Baseline CDAI will be added in the table on baseline disease characteristics.
- VAS plots on percent changes from baseline: from the interim analysis it seemed that the mean values were not so much influenced by the few outliers as initially anticipated. Therefore medians are no longer needed: plots will be mean plots just like all other efficacy parameters.
- Scatterplots between PK plasma levels and safety (lab safety and ECG QTcF) were removed. Because PK plasma sampling can happen at any time relative to the last drug intake ('sparse sampling'), the result is not always representing the subject's exposure. Making a correlation between the PK finding and safety would then be misleading.

Additions:

- Pharmacodynamic analyses were added and specified for the final analysis.
- DAS28(CRP) remission and low disease activity, SDAI remission, CDAI remission: between-group inferential statistics were added, similar to the ACR20 analysis, to allow a more detailed interpretation of the results.

### 6.3.3. Changes after Final Database Lock

None.

## 7. DEFINITIONS OF GENERAL TABLES, LISTINGS AND FIGURES

Mock tables, listings and figures are provided in a separate document.

All analyses will be run on the Safety population, unless specified otherwise.

### 7.1. SUBJECT DISPOSITION

#### Listing 16.2.1.1: Subject disposition: Randomization

Listing of subject numbers, randomization and re-randomization groups, population flags, any information on code breaking, stratification factors (region, previous use of biologics), country and site.

All discrepancies (as-randomized versus as-treated, or discrepancies in stratification factors) will be flagged.

Regions used in the central randomization in the DARWIN program are:

Region	Countries belonging to the region
1: Latin America	Argentina, Chile, Colombia, Guatemala, Mexico
2: Central and Eastern Europe, EU	Bulgaria, Hungary, Latvia, Poland
3: Central and Eastern Europe, non-EU	Russia, Ukraine, Moldova
4: West and Asia Pac	Australia, Austria, Belgium, Czech Republic, France, Germany, Israel, New Zealand, Spain, USA

Population: all randomized population.

#### Listing 16.2.1.2: Subject disposition: Region, country and site identification

Listing of region, country and site numeric codes used for the numbering of the subjects, and the meaning of these codes.

Population: all screened population.

#### Table 14.1.1.1: Subject disposition: Tabulation by region, country and investigator

Tabulation per initial treatment group (and overall) of the number of subjects in each of the regions, countries and investigator sites. For the interim analysis, only regions and countries will be presented.

Population: all randomized.

Table to be created for the Week 12 interim analysis as well.

**Table 14.1.1.2: Subject disposition: Analysis populations**

Tabulation per initial treatment group (and overall) of the number of subjects in each of the analysis populations defined in section 3.1, separated per study period.

Population: all screened population.

**Table 14.1.1.3: Subject disposition: No-treatment subjects: Tabulation of the reasons**

Tabulation of discontinuations and the reasons for discontinuation.

Population: all screened population, minus the safety population (so selecting only subjects who were never exposed to the study medication: subjects either never randomized or randomized but not exposed).

**Table 14.1.1.4: Subject disposition up to Week 12: Tabulation of the number of subject at each time interval**

Tabulation per treatment group (and overall) and per time point (up to and including Week 12).

Population: safety.

**Table 14.1.1.5: Subject disposition up to study end: Tabulation of the number of subject at each time interval**

Repeat table after Week 12 and follow-up.

Population: safety subjects who entered the second part of the study.

**Listing 16.2.1.3: Subject disposition: Number of days in study**

Listing per initial treatment group, per subject and per time point the number of days in study at the time of the visit (derivation of these “days”: see section 3.1), with population flags.

Population: safety.

**Table 14.1.1.6: Subject disposition: First and last date in the study**

List the following (overall, not per subject):

- Date of the first signature on study ICF
- Date of first and last screening visit
- Screening period duration: last screening date – first screening date + 1 day
- Date of first study drug administration (periods 1 and 2)
- Last visit date (all visits; including unscheduled visits)
- Last date of contact in the study with any subject.

Population: all screened population.

**Table 14.1.1.7: Subject disposition up to Week 12: Tabulation of the trial termination reasons**

Tabulation per treatment group (and overall) of completion/discontinuations during the first period of the study, with the reasons for discontinuation. Subjects who moved on past the Week 12 visit will be shown as “completed”.

Population: safety.

Table to be created for the Week 12 interim analysis as well.

**Table 14.1.1.8: Subject disposition up to study end: Tabulation of the trial termination reasons**

Repeat table for the second period of the study.

Population: safety subjects who entered the second part of the study.

**Listing 16.2.1.4: Subject disposition: Trial termination**

Listing per treatment group and per subject of the reason for completion/discontinuation and the number of days since first study treatment administration at trial termination. In case the discontinuation was due to AE, the AE preferred term will be presented in this listing. If there is a (verbatim) explanation on the discontinuation reason, this will also be presented in this listing.

Population: safety.

**Listing 16.2.1.5: Subject disposition: Study analysis periods**

Listing per treatment group and per subject of the analysis periods in the study, together with the start and end dates of each analysis period. With the date of first and last use of study medication in each study period, population flags, the ICF signature date and trial termination date. (Analysis periods: see section 3.2)

Population: safety.

## 7.2. PROTOCOL DEVIATIONS AND ELIGIBILITY

Protocol deviations are determined and recorded whilst the study is ongoing, and the list is finalized prior to final database lock (and unblinding). All deviations are classified as either “critical” or “non-critical”. Only the critical (=major) protocol deviations will be presented in the tables. For more details, please refer to the Protocol Deviation Plan.

### Table 14.1.2.1: Major protocol deviations

Tabulation per treatment group (and overall) of the number (and percentage) of subjects having a major/minor protocol deviation, together with the detailed protocol deviation categories of the major deviations only.

Population: safety.

### Listing 16.2.2.1: Protocol deviations

Listing per treatment group and per subject of all protocol deviations, with indication critical/non-critical.

Population: safety.

### Listing 16.2.2.2: Eligibility criteria: Violations

Only violated in- and exclusion criteria will be listed per treatment group and per subject. These are criteria as ticked by the investigator in the eCRF.

Population: safety.

### Listing 16.2.2.3: Final eligibility statements

Listing per treatment group and per subject of the final eligibility statements in the eCRF at screening and at baseline.

Population: safety.

## 7.3. SUBJECTS EXCLUDED FROM ANALYSIS

### Listing 16.2.3.1: Subjects excluded from the safety analysis

Listing of all subjects that were randomized and/or not treated: the trial termination reason and/or the reason for being a no-treatment subject will be listed, whichever is available.

Population: all screened population, minus the Safety population.

### Listing 16.2.3.2: Subjects excluded from the efficacy ITT analysis

Listing of all subjects that were exposed but who were not included in the ITT population: the trial termination reason and the reason for being excluded will be listed.

Population: Safety, minus the ITT population.

**Listing 16.2.3.3: Subjects excluded from the efficacy per-protocol analysis**

Listing of all subjects that were in the ITT population but not in the per-protocol population: the trial termination reason and the reason for being excluded will be listed.

Population: ITT population, minus the Per-protocol population.

**7.4. DEMOGRAPHICS AND BASELINE DISEASE CHARACTERISTICS****7.4.1. Parameters**

- sex
- age (years) as recorded in the eCRF
- age categories:
  - < 45 ;
  - [45,65[ ;
  - [65,75[ ;
  - ≥ 75 years
- year of birth: only listed
- race
- ethnicity
- height (cm)
- weight (kg)
- body mass index BMI = (weight in kg) / (height in m)<sup>2</sup> (kg/m<sup>2</sup>): the BMI will be recalculated and rounded to the nearest integer, even when already available in the database. The original BMI will not be used in that case.
- BMI categories:
  - ≤18.5;
  - ]18.5,25.0] ;
  - ]25.0,30.0[ ;
  - ≥30.0 kg/m<sup>2</sup>

**7.4.2. Analysis****Table 14.1.3.1: Demographic data**

Continuous parameters: descriptive statistics per treatment group (and overall). Overall p-value to compare all groups: 1-way ANOVA model with factor treatment.

Categorical parameters: frequency tabulation per treatment group (and overall). Overall p-value to compare all groups: Cochran-Mantel-Haenszel test for general association.

Population: safety, ITT, per-protocol.

Table to be created for the Week 12 interim analysis as well, but only for Safety and ITT populations.



**Table 14.1.3.2: Demographic data for ACR20 responders versus non-responders at Week 12**

Same descriptive table as above, but split up for ACR20 responders and non-responders at Week 12. See further on for the definition of ACR20.

Population: ITT.

Table to be created for the Week 12 interim analysis as well.

**Listing 16.2.4.1: Demographic data**

Listing per treatment group and per subject of all demographic parameters.

Population: safety.

**Table 14.1.3.3: Baseline laboratory data: Tabulation of abnormalities**

Frequency tabulation per lab test and per treatment group (and overall) of the baseline result, categorized as low/normal/high according to the normal ranges (see section 12.2.5). Exclude the CRP.

Population: safety.

**Table 14.1.3.4: Baseline ECG data: Tabulation of abnormalities**

Frequency tabulation per parameter (QT and QTcF) per treatment group (and overall) of the baseline result, categorized as (same as in section 12.3.5):

≤450 ms ;

]450,480] ms ;

]480,500] ms ;

>500 ms.

Population: safety.

**Table 14.1.3.5: Baseline vital signs data: Tabulation of abnormalities**

Frequency tabulation per parameter and per treatment group (and overall) of the baseline result, categorized as low/normal/high according to the normal ranges (see section 12.4.3).

Population: safety.

**Table 14.1.3.6: Baseline physical examination: Tabulation**

Frequency tabulation per CRF body system and per treatment group (and overall) of the baseline result, categorized as normal/abnormal.

Population: safety.

## 7.5. BASELINE DISEASE CHARACTERISTICS

### 7.5.1. Parameters

- Duration of RA (years) =  $\frac{(\text{date of initial diagnosis}) - (\text{screening date}) + 1}{365.25}$ , rounded to the nearest integer. If the date of initial diagnosis is incomplete, then the following rules will be applied: Missing day: use the first of the month. Missing month: use January.
- Duration of RA, categorized:
  - <0.5;
  - [0.5,2[ ;
  - [2,5[ ;
  - [5,10[ ;
  - [10,20[ ;
  - ≥20 years
- Screening and baseline CRP
- Screening and baseline CRP, categorized:
  - <0.7 x ULN;
  - [0.7,1.0] x ULN;
  - ]1.0,1.2[ x ULN;
  - ≥ 1.2 x ULN
- Change in CRP (baseline – screening)
- Screening and baseline TJC68 and SJC66
- Change in TJC68 and SJC66 = baseline – screening
- Screening and baseline DAS28(CRP) – see further sections for the definition of the DAS28(CRP)
- Change in DAS28(CRP) = baseline – screening
- Screening and baseline patient's global evaluation of disease activity (0-100 mm VAS)
- Change in patient's global evaluation of disease activity = baseline – screening
- Baseline investigator's global evaluation of disease activity (0-100 mm VAS)
- Baseline patient's pain score (0-100 mm VAS)
- Baseline HAQ-DI
- Baseline SDAI
- Baseline CDAI
- RF status, classified as positive/negative based on the upper limit of the normal range (ULN):
  - o Positive: value > ULN
  - o Negative: < value ≤ ULN
  - o Missing: value is missing for any reason.
- Anti-CCP status, classified as positive/negative based on the upper limit of the normal range (ULN):
  - o Positive: value > ULN
  - o Negative: < value ≤ ULN
  - o Missing: value is missing for any reason.
- Affected joints identification
- Prior RA medication:

- Generic names, also classified as: biological (anti-TNF or other biological DMARD), anti-TNF, other biological DMARD, conventional DMARD, NSAID, steroid, analgesic, other RA medication.
- MTX duration (study screening date – date of start of MTX + 1 day), expressed in years. If there were multiple dosing periods of MTX, then the “duration” will add up all such episodes.
- MTX duration, categorized as:
  - <0.5;
  - [0.5,2[ ;
  - [2,5[ ;
  - [5,10[ ;
  - [10,20[ ;
  - ≥20 years
- MTX dose at the start of the study, expressed in mg/week
- MTX dose at the start of the study, categorized as:
  - <10 ;
  - [10,15[ ;
  - ]15,20[ ;
  - >20 mg/week
- Steroids (as prednisolone-equivalent): total daily dose at the start of the study
- Steroids (as prednisolone-equivalent): total daily dose at the start of the study, categorized as:
  - <5 ;
  - [5,10[ ;
  - [10,15[ ;
  - [15,20[ ;
  - ≥20 mg/day.

## 7.5.2. Analysis

### Table 14.1.3.7: Screening and baseline disease characteristics

Continuous parameters: descriptive statistics per treatment group (and overall). Overall p-value to compare all groups: ANOVA model with factor treatment.

Categorical parameters: frequency tabulation per treatment group (and overall). Overall p-value to compare all groups: Cochran-Mantel-Haenszel test for general association.

Population: safety, ITT, per-protocol.

Table to be created for the Week 12 interim analysis as well, but only for Safety and ITT populations.

### Table 14.1.3.8: Screening and baseline disease characteristics by region

Same descriptive table as above, but split up for the 4 geographical regions.

Population: ITT.

Table to be created for the Week 12 interim analysis as well.

**Table 14.1.3.9: Screening and baseline disease characteristics for ACR20 responders versus non-responders at Week 12**

Same descriptive table as above, but split up for ACR20 responders and non-responders at Week 12.

Population: ITT.

Table to be created for the Week 12 interim analysis as well.

**Table 14.1.3.10: Affected joints: Tabulation**

Frequency table of the detailed affected joints, separately for: tender / swollen. No formal comparisons of treatment groups.

Population: Safety.

**Table 14.1.3.11: Prior RA medication**

Tabulation of the number of prior RA medications (split up for anti-TNF biological DMARD, other biological DMARD, conventional DMARD, NSAID, steroid, analgesic, other RA medication) and the generic terms per treatment group and overall. No formal comparisons of treatment groups.

Population: safety.

**Table 14.1.3.12: Prior RA medication for ACR20 responders versus non-responders at Week 12**

Same descriptive table as above, but split up for ACR20 responders and non-responders at Week 12. No formal comparisons of treatment groups.

Population: ITT.

**Table 14.1.3.13: Prior methotrexate: Descriptive statistics**

Descriptive statistics per treatment group (and overall) of the total prior MTX duration (in years) and the MTX dose at the start of the study (mg/week).

Frequency per treatment group (and overall) of the total prior MTX duration (in years).

Frequency per treatment group (and overall) of the MTX dose at the start of the study (mg/week).

Population: Safety.

**Table 14.1.3.14: Prior steroid dose at the start of the study: Descriptive statistics**

Descriptive statistics per treatment group (and overall) of the prior steroid dose at the start of the study (mg/day prednisolone-equivalent).

Frequency per treatment group (and overall) of the prior steroid dose at the start of the study.

Population: Safety.

#### **Listing 16.2.4.2: Prior RA medication**

Listing per treatment group and per subject of the prior RA medications.

Population: safety.

## **7.6. MEDICAL HISTORY AND CONCURRENT DISEASES**

#### **Table 14.1.3.15: Medical history: Tabulation**

Frequency tabulation per treatment group (and overall) of the system organ classes and preferred terms, selecting only the medical history findings (i.e., condition no longer present at the start of the study).

Population: safety.

#### **Table 14.1.3.16: Concurrent diseases: Tabulation**

Frequency tabulation per treatment group (and overall) of the system organ classes and preferred terms.

Population: safety.

#### **Listing 16.2.4.3: Medical history**

Listing per treatment group and per subject of the medical history data findings (i.e., condition no longer present at the start of the study): original terms as well as coded terms.

Population: safety.

#### **Listing 16.2.4.4: Concomitant diseases**

Listing per treatment group and per subject of the concomitant diseases data findings (i.e., condition still present or unknown): original terms as well as coded terms.

Population: safety.

## **7.7. PRIOR AND CONCOMITANT THERAPIES**

Note that prior RA medications are presented in the section on baseline disease characteristics. They are not repeated here.

### **7.7.1. Classification of Therapies**

All prior and concomitant therapies will be allocated into exactly one of the following categories (using the analysis periods as defined in section 3.2):

- Prior only: the therapy ended before the first study treatment administration.
- Concomitant only: the therapy started on or after the first study drug administration.
- Post-treatment: the therapy started after the last study drug administration.
- Prior and concomitant: the therapy started before the first study drug administration, and ended on or after the first study drug administration.

When the start date is completely missing, the therapy is assumed to have started on the same day as the first administration of study medication.

When the end date is missing, the therapy is assumed to be still ongoing after the end of the study.

When the start date is only partially known, the missing parts will be imputed by the first day of the month/year, or the day of the first medication intake in case the nonmissing part matches.

When the end date is only partially known, the missing parts will be imputed by the last day of the month/year, or the day of the last medication intake in case the nonmissing part matches.

Imputation of start and end dates is only done for the classification into prior / concomitant / prior + concomitant. The listing will present the original start and end dates, also when incomplete. Fields like start and stop days will not be derived from imputed incomplete dates.

The relative study day of therapy start and stop will be derived as follows (only listed):

### **Start day**

= (Therapy start date) – (date of first study medication administration) + 1, when the start date of therapy is known and complete, and when the therapy start date is on or after the date of first study medication;

= (Therapy start date) – (date of first study medication administration), when the start date of therapy is known and complete, and when the therapy start date is before the date of first study medication;

= missing when the start date of therapy is unknown, missing, or incomplete.

### **Stop day**

= (Therapy stop date) – (date of first study medication administration) + 1, when the end date of therapy is known and complete, and the end date of therapy is on or after the date of first study medication;

= (Therapy stop date) – (date of first study medication administration), when the end date of therapy is known and complete, and the end date of therapy is before the date of first study medication;

= (Trial termination date) – (date of first study medication administration) + 1, when the therapy is still ongoing when the subject leaves the study; in such cases the stop day will be presented as “>XX days” in the listing;

= missing when the end date of therapy is unknown, missing or incomplete and the therapy isn't ongoing after the subject left the study.

### 7.7.2. Coding of Therapies

All therapies are coded using WHO-DRUG. In the table(s), the generic term will be used. Multiple records of the same generic term for the same subject with the same categorization will be counted only once. The table will therefore present subjects, not occurrences.

### 7.7.3. Analysis

#### **Table 14.1.3.17: Prior therapies: Frequency table**

Frequency tabulation per treatment group (and overall) of the ATC classes and generic terms, of "prior" and "prior+concomitant" therapies only.

These are all the therapies that were used before the first intake of study medication.

Population: safety.

#### **Table 14.1.3.18: Concomitant therapies up to Week 12: Frequency table**

Frequency tabulation per treatment group (and overall) of the ATC classes and generic terms, selecting both "concomitant" (with a start date during the first study period, or a missing start date) and "prior+concomitant" (all of them) therapies.

These are all the therapies that were used during the first 12 weeks of the study, concomitantly to the study medication.

Population: safety.

#### **Table 14.1.3.19: Concomitant therapies over the whole study period: Frequency table**

Frequency tabulation per treatment group (and overall) of the ATC classes and generic terms, selecting "concomitant" and "prior+concomitant" therapies only.

These are all the therapies that were used during the study, concomitantly to the study medication.

Population: safety.

#### **Table 14.1.3.20: Post-treatment therapies: Frequency table**

Frequency tabulation per treatment group (and overall) of the ATC classes and generic terms, selecting "post-treatment" therapies only.

These are all the therapies that were used during the study, but started after the last dose of study medication.

Population: safety.

#### **Listing 16.2.4.5: Prior and concomitant therapies**

Listing per treatment group and per subject of all data on prior and concomitant therapies. Flags for prior/concomitant/post-treatment will be added. In case the therapy was due to an AE, the AE preferred term will be mentioned in this listing.

Population: safety.

## 7.8. EXPOSURE TO STUDY MEDICATION AND COMPLIANCE

### 7.8.1. Derivations for placebo/GLPG0634

For the first period of the study (baseline – Week 12):

- Total treatment duration (days) = (last dose administration date – first dose administration date) + 1 day, where the last dose administration date is cut off on the day prior to the Week 12 visit. (“Prior to the Week 12 visit”, because at the Week 12 visit the re-randomization takes place and it is not known when exactly on that day the new study drug is being taken.)
- Total treatment duration, excluding days off drug (days): sum of all durations (last – first +1) in the drug log pages where there is a morning and/or evening dose >0 capsules, and where the last duration is cut off on the day prior to the Week 12 visit.
- Percentage days with an intake =  $100\% \times \left( \frac{\text{total treatment duration, excluding days off drug}}{\text{total duration}} \right)$
- Total compliant treatment duration (days): sum of all durations (last – first +1) in the drug log pages where there is a morning and/or evening dose with exactly 2 capsules, and where the last duration is cut off on the day prior to the Week 12 visit.
- Percentage compliance =  $100\% \times \left( \frac{\text{total compliant treatment duration}}{\text{total duration}} \right)$

The “total duration” used as denominator in the above formulae equals: (date of Week 12 visit) – (date of first drug intake). This total duration is not including the Week 12 visit itself, just like the above derivations, because of the treatment switch happening at Week 12 making it uncertain which drug intakes actually happened at the day of the visit itself.

For the total 24 weeks study period:

- Total treatment duration (days) = (last dose administration date – first dose administration date) + 1 day.
- Total treatment duration, excluding days off drug (days): sum of all durations (last – first +1) in the drug log pages where there is a morning and/or evening dose >0 capsules.
- Percentage days with an intake =  $100\% \times \left( \frac{\text{total treatment duration, excluding days off drug}}{\text{total duration}} \right)$
- Total compliant treatment duration (days): sum of all durations (last – first +1) in the drug log pages where there is a morning and/or evening dose with exactly 2 capsules.
- Percentage compliance =  $100\% \times \left( \frac{\text{total compliant treatment duration}}{\text{total duration}} \right)$

### 7.8.2. Tables up to Week 12

#### Table 14.1.4.1: Use of study medication up to Week 12: Descriptive statistics



Descriptive statistics per treatment group (and overall) of the total treatment duration (days) and the overall compliance (%).

Frequency tabulation per treatment group (and overall) of the two treatment duration parameters, categorized in weeks.

Frequency tabulation per treatment group (and overall) of the days with intake and compliance, categorized as:

<80% ;

[80%,100%[ ;

100% ;

]100%,120%] ;

>120%.

Population: safety.

### **7.8.3. Tables for the whole 24 weeks period**

#### **Table 14.1.4.2: Use of study medication over the whole study period: Descriptive statistics**

Similar to the preceding table, also using the 11 treatment sequences of Chapter 3.

### **7.8.4. Listings**

#### **Listing 16.2.5.1: Exposure to study medication: eCRF data**

Listing per treatment group and per subject of all data related to the use of study medication as recorded in the eCRF.

Population: safety.

#### **Listing 16.2.5.2: Exposure to study medication: Derived data**

Listing per treatment group and per subject of all derived parameters related to the use of study medication.

Population: safety.

#### **Listing 16.2.5.3: Dispensed kit numbers**

Listing per treatment group and per subject of all dispensed placebo/GLPG0634 kit numbers, flagging any mistakes in dispensing.

Population: safety.

## 7.9. EXPOSURE TO METHOTREXATE AND COMPLIANCE

### 7.9.1. Derivations for methotrexate

The average weekly dose (mg) of methotrexate (MTX) in each of the two study periods is derived from summing up the durations of loglines with a weekly dose >0. The eCRF collects the use of MTX in terms of total weekly doses rather than daily doses.

For the first period of the study (screening – Week 12):

- MTX dose change: decrease / no change / increase: at any time during the entire period, versus the dose used at the start of the study.
- Total dose, excluding days off drug (days): sum of all doses (*last – first + 1*) x (*weekly dose in that period*) in the MTX drug log pages where there is a weekly dose >0 mg, and where the first duration starts with the screening date and the last duration is cut off on the day of the Week 12 visit.

$$\text{Average weekly dose (mg)} = \frac{\text{total dose}}{(\text{date of Week 12} - \text{date of screening} + 1) / 7}$$

For the whole 24 weeks study period:

- MTX dose change: decrease / no change / increase: at any time during the entire period, versus the dose used at the start of the study.
- Total dose, excluding days off drug (days): sum of all doses (*last – first + 1*) x (*weekly dose in that period*) in the MTX drug log pages where there is a weekly dose >0 mg, and where the first duration starts with the screening date and the last duration is cut off on the day of the Week 24/FU/Early discontinuation visit (whichever applies).

$$\text{Average weekly dose (mg)} = \frac{\text{total dose}}{(\text{date of Week 24/FU/Early term visit} - \text{date of Week 12}) / 7}$$

### 7.9.2. Tables up to Week 12

#### Table 14.1.4.3: Changes in methotrexate up to Week 12: tabulation

Tabulation per treatment group (and overall) of the change in MTX dose (decrease/no change/increase).

Population: safety.

#### Table 14.1.4.4: Use of methotrexate up to Week 12: Descriptive statistics of the average weekly dose

Descriptive statistics per treatment group (and overall) and per study period of the average weekly dose (mg/week).

Frequency tabulation per treatment group (and overall) of the average weekly dose, categorized as:

<10 ;  
 [10,15] ;  
 ]15,20] ;  
 >20 mg/week.

Population: safety.

### **7.9.3. Tables for the whole 24 weeks period**

Similar tables, using the 11 treatment arms of Chapter 3:

**Table 14.1.4.5: Changes in methotrexate over the whole study period: tabulation**

**Table 14.1.4.6: Use of methotrexate over the whole study period: Descriptive statistics of the average weekly dose**

### **7.9.4. Listings**

**Listing 16.2.5.4: Exposure to methotrexate: eCRF data**

Listing per treatment group and per subject of all data related to the use of methotrexate, as recorded in the eCRF.

Population: safety.

**Listing 16.2.5.5: Exposure to methotrexate: Derived data**

Listing per treatment group and per subject of all derived parameters related to the use of methotrexate.

Population: safety.

## **8. DEFINITIONS OF EFFICACY TABLES, LISTINGS AND FIGURES**

Mock tables, listings and figures are provided in a separate document.

All analyses will be run on the ITT population, unless specified otherwise.

General comment on the efficacy data listings: listings will present the observed-case time points, with imputed data (for components) clearly flagged. The LOCF- or NRI-imputed time points will not be shown.

General comments on the observed-case analysis:

- Is done for efficacy tables and listings, but not for efficacy plots.
- Is done for Baseline, Week 12 and Week 24 in the descriptive statistics.
- Is done for Week 12 only in the inferential statistics.
- Is done for ITT only; not for per-protocol.

## 8.1. ACR20, ACR50, ACR70 AND ACR-N

### 8.1.1. Parameters

ACR20 response when:

- $\geq 20\%$  improvement from baseline in SJC66, and
- $\geq 20\%$  improvement from baseline in TJC68, and
- $\geq 20\%$  improvement from baseline in at least 3 of the following 5 items:
  - Pain VAS (taken from the HAQ-DI)
  - Patient's Global Assessment of Disease Activity VAS
  - Physician's Global Assessment of Disease Activity VAS
  - Total HAQ-DI score
  - CRP.

Subjects who do not meet this criterion are non-responders.

If any of the 7 components of the ACR20 would be missing, then this missing component will be imputed using the value of the last preceding nonmissing visit. If all components are missing, then the derived ACR20 will remain missing. Imputation of such cases for the tables will be done there using LOCF or NRI like any other missing value.

Subjects who discontinue within the first week so that ACR20/50/70 responses cannot be determined will be classified as non-responders.

The ACR50 and ACR70 responses are derived using the same algorithm, but with 50% and 70% cutoffs respectively.

**Note that the primary endpoint of the study is the ACR20 at Week 12 in the ITT population, using NRI imputation.**

The ACR-N is derived using the following formula:

$$\text{ACR-N} = \text{MIN} [ \% \text{ improvement in TJC68, } \% \text{ improvement in SJC66, MED } \% \text{ improvement in (patient's global assessment, physician's global assessment, patient's assessment of pain, HAQ-DI, CRP) } ]$$

So this is taking the smallest of the percentage increases in SJC66, TJC68, and the third of the remaining 5 parameters. In case the ACR-N is negative (i.e., the subject actually worsens instead of improves), then the ACR-N is replaced by a 0. The ACR-N is therefore a number between 0 and 100.

The ACR-N represents the percentage NN% according to which the subjects would have been an "ACRNN" responder.

(See further sections on how to derive these components.)

#### The time to ACR20 response

= (date of first time ACR20 response is reached) – (date of first medication intake) + 1 day, if the subject achieved ACR20 response during the initial 12 weeks treatment period, or  
= MIN (date of the Week 12 visit , date of trial termination) – (date of first medication intake) + 1 day, in case the subject did not achieve ACR20 response at any time during the initial 12 weeks treatment period.

Next to this, the visit during which the ACR20 response was first recorded will be analyzed as well:

- Week 1
- Week 2
- Week 4
- Week 8
- Week 12
- Never (i.e., subject did not achieve ACR20 response during the initial 12 weeks of treatment).

The time to ACR50 response and the time to ACR70 response are defined in the same way.

Note that the study period 2 cannot be used for the time-to-event derivation, because of the predefined potential treatment shifts at Week 12.

Note that if for a specific visit both the regular CRP as well as the high-sensitive CRP are available, the high-sensitive CRP will be used for the derivation of the percentage change from baseline and also for the ACR20/50/70/N.

Note that the percentage change from baseline remains missing if the baseline value itself equals zero.

### **8.1.2. Tables over the whole 24 weeks period**

Tables will present Weeks 1-2-4-8-12-16-20-24, unless specified otherwise. Tables will present the imputed 7 treatment groups, as mentioned in section 2.1.1.

#### **Table 14.2.1.1.1: ACR20: Summary with overall p-value and pairwise comparisons versus placebo**

Logistic regression model at each time point, modeling the responders. Factors in the model are: treatment, region, and previous use of biologics.

Per time point, the table will present:

- Overall type-III p-value for all explanatory factors in the model.
- Per group: number and percentage of responders, difference in percentage responders versus the placebo group (GLPG0634 minus placebo), together with a 95% confidence interval of this difference (as provided by SAS PROC FREQ).
- Per group: p-value for pairwise comparison versus the placebo group, derived from contrasts within the overall regression model. These p-values will be presented both as uncorrected p-values as well as corrected for multiplicity according to Hommel's closed-testing method (SAS PROC MULTTEST).

The "raw" SAS PROC output must be provided in a statistical appendix listing.

The table will show:

- NRI in ITT and in per-protocol
- LOCF in ITT and in per-protocol
- observed-case in ITT (descriptive at Week 12 + 24; inferential statistics at Week 12 only).

Table to be created for the Week 12 interim analysis as well, cut off at Week 12. ITT only.

**Table 14.2.1.1.2: ACR20: Exploratory pairwise comparisons among the dose groups within each regimen**

Logistic regression model at each time point, modeling the responders. Factors in the model are: treatment, region, and previous use of biologics. Placebo won't be used, so can be omitted from the overall model.

Per time point, the table will present:

- Per group: difference in percentage responders versus the other GLPG0634 groups within the same regimen (q.d. or b.i.d.), together with a 95% confidence interval of this difference (as provided by SAS PROC FREQ).
- Per group: p-value for pairwise comparison versus the other GLPG0634 groups. These p-values will be presented both as uncorrected p-values, and are derived as contrasts within the overall logistic regression model. As this is a purely exploratory analysis, no adjustment for multiplicity will be performed.

The "raw" SAS PROC output must be provided in a statistical appendix listing.

The table will show:

- NRI in ITT and in per-protocol
- LOCF in ITT and in per-protocol
- observed-case in ITT (inferential statistics at Week 12 only).

Table to be created for the Week 12 interim analysis as well, cut off at Week 12. ITT only.

**Table 14.2.1.1.3: ACR20: Exploratory pairwise comparisons between regimens within each daily dose level**

Logistic regression model at each time point, modeling the responders. Factors in the model are: regimen, dose, regimen\*dose interaction, region, and previous use of biologics. Placebo

subjects will not be used in this model (as there is no q.d./b.i.d. regimen available for placebo). The regimen\*dose interaction will be removed from the final model when not significant at level 0.10 (two-sided type-III p-value). Regimen will have 2 categories: q.d. or b.i.d. Dose will have 3 categories, corresponding to the total daily dose: 50 mg/day, 100 mg/day, or 200 mg/day.

Per time point, the table will present:

- Per dose: difference in percentage responders of q.d. versus b.i.d. within the same dose level (50, 100 or 200 mg/day), together with a 95% confidence interval of this difference (as provided by SAS PROC FREQ).
- Per group: p-value for pairwise comparison of q.d. versus b.i.d. These p-values will be presented both as uncorrected p-values, and are derived as contrasts within the overall logistic regression model. As this is a purely exploratory analysis, no adjustment for multiplicity will be performed.

The “raw” SAS PROC output must be provided in a statistical appendix listing.

The table will show:

- NRI in ITT and in per-protocol
- LOCF in ITT and in per-protocol
- observed-case in ITT (inferential statistics at Week 12 only).

Table to be created for the Week 12 interim analysis as well, cut off at Week 12. ITT only.

**Table 14.2.1.1.4: ACR50: Summary with overall p-value and pairwise comparisons versus placebo**

Repeat table of the ACR20 table.

The table will show:

- NRI in ITT
- LOCF in ITT
- observed-case in ITT (descriptive at Week 12 + 24; inferential statistics at Week 12 only).

Table to be created for the Week 12 interim analysis as well, cut off at Week 12.

**Table 14.2.1.1.5: ACR50: Exploratory pairwise comparisons among the dose groups within each regimen**

Repeat table of the ACR20 table.

The table will show:

- NRI in ITT
- LOCF in ITT
- observed-case in ITT (inferential statistics at Week 12 only).

Table to be created for the Week 12 interim analysis as well, cut off at Week 12.

**Table 14.2.1.1.6: ACR50: Exploratory pairwise comparisons between regimens within each daily dose level**

Repeat table of the ACR20 table.

The table will show:

- NRI in ITT
- LOCF in ITT
- observed-case in ITT (inferential statistics at Week 12 only).

Table to be created for the Week 12 interim analysis as well, cut off at Week 12.

**Table 14.2.1.1.7: ACR70: Summary with overall p-value and pairwise comparisons versus placebo**

Repeat table of the ACR20 table.

The table will show:

- NRI in ITT
- LOCF in ITT
- observed-case in ITT (descriptive at Week 12 + 24; inferential statistics at Week 12 only).

Table to be created for the Week 12 interim analysis as well, cut off at Week 12.

**Table 14.2.1.1.8: ACR70: Exploratory pairwise comparisons among the dose groups within each regimen**

Repeat table of the ACR20 table.

The table will show:

- NRI in ITT
- LOCF in ITT
- observed-case in ITT (inferential statistics at Week 12 only).

Table to be created for the Week 12 interim analysis as well, cut off at Week 12.

**Table 14.2.1.1.9: ACR70: Exploratory pairwise comparisons between regimens within each daily dose level**

Repeat table of the ACR20 table.

The table will show:

- NRI in ITT
- LOCF in ITT
- observed-case in ITT (inferential statistics at Week 12 only).

Table to be created for the Week 12 interim analysis as well, cut off at Week 12.



**Table 14.2.1.1.10: ACR20 at Week 12: Tabulation versus baseline characteristics**

The table will show two methods of imputing missing values: LOCF and NRI.

Frequency tabulation of the ACR20 response rate at Week 12 per treatment group, split up according to the following baseline categories:

- Demographics:
  - Region (as used in the central randomization)
  - Country
  - Gender
  - Age categories:
    - < 45 ;
    - [45,65[ ;
    - [65,75[ ;
    - ≥ 75 years
  - BMI categories:
    - ≤ 18.5;
    - ] 18.5 , 25.0 ] ;
    - ] 25 , 30 [,
    - ≥ 30 kg/m<sup>2</sup>
- Baseline disease characteristics:
  - Duration of RA:
    - < 1;
    - [ 1 , 2 [ ;
    - [ 2 , 5 [ ;
    - [ 5 , 10 [ ;
    - [ 10 , 20 [ ;
    - ≥ 20 years
  - Baseline CRP:
    - <0.7 x ULN;
    - [0.7,1.0] x ULN;
    - ]1.0,1.2[ x ULN;
    - ≥ 1.2 x ULN
  - Baseline TJC68: < 20, ≥ 20
  - Baseline SJC66: < 20, ≥ 20
  - Baseline DAS28(CRP): high, moderate, low, remission (see further section on DAS28 for the definition of these categories)
  - RF status
  - Anti-CCP status
  - Previous use of:
    - Biological (anti-TNF or other biological DMARD) yes/no,
    - anti-TNF yes/no,
    - other biological DMARD yes/no,
    - conventional DMARD yes/no,
    - NSAID yes/no,
    - steroid yes/no,

- analgesic yes/no,
- other RA medication yes/no.

Population: ITT.

Table to be created for the Week 12 interim analysis as well.

**Table 14.2.1.1.11: ACR20 at Week 12: Exploratory logistic regression model containing baseline characteristics**

The table will show two methods of imputing missing values: LOCF and NRI.

Logistic regression model of the ACR20 response at Week 12, with the following exploratory factors:

- Treatment group
- Demographics:
  - Region
  - Country
  - Gender
  - Age as continuous variable
  - BMI as continuous variable
- Baseline disease characteristics:
  - Duration of RA as continuous variable
  - Baseline CRP as continuous variable
  - Baseline TJC68 and SJC66 as continuous variables
  - Baseline DAS28(CRP) as continuous variable
  - RF status
  - anti-CCP status
  - Previous use of:
    - Biological (anti-TNF or other biological DMARD) yes/no,
    - anti-TNF yes/no,
    - other biological DMARD yes/no,
    - conventional DMARD yes/no,
    - NSAID yes/no,
    - steroid yes/no,
    - analgesic yes/no,
    - other RA medication yes/no.

All these factors will be tested using a stepwise selection approach. Significance level to enter the model: 0.10; significance level to stay in the model: 0.05. The treatment groups will always be kept in the model, even when not significant at level 0.05. The table will present only the final model with type-III p-value and odds ratio estimates. The “raw” SAS PROC output must be provided in a statistical appendix listing, and will contain all details on the model building steps.

Population: ITT.

**Table 14.2.1.1.12: ACR-N: Summary with overall p-value and pairwise comparisons versus placebo**

P-values are generated from an ANOVA model at each time point. Factors in the model are: treatment, region, and previous use of biologics.

Per time point, the table will present:

- Per group: sample regular statistics.
- Per group: difference in LS-means versus the placebo group (GLPG0634 minus placebo), together with a 95% confidence interval of this difference.
- Overall type-III p-value for “treatment”.
- Per group: p-value for pairwise comparison versus the placebo group, derived from contrasts within the overall model. These p-values will be presented both as uncorrected p-values as well as corrected for multiplicity according to Hommel’s closed-testing method (SAS PROC MULTTEST).

Frequency per time point (weeks 12 and 24 only), and per treatment group. ACR-N will be presented in the following categories:

<10,  
 [10,20[,  
 [20,30[,  
 [30,40[,  
 [40,50[,  
 [50,60[,  
 [60,70[,  
 [70,80[,  
 [80,90[,  
 [90,100].

The “raw” SAS PROC output must be provided in a statistical appendix listing.

The table will show:

- LOCF in ITT
- observed-case in ITT (descriptive at Week 12 + 24; inferential statistics at Week 12 only).

Table to be created for the Week 12 interim analysis as well, cut off at Week 12.

**Table 14.2.1.1.13: ACR-N: Exploratory pairwise comparisons among the dose groups within each regimen**

ANOVA model at each time point. Factors in the model are: treatment, region, and previous use of biologics. Placebo won’t be used, so can be omitted from the overall model.

Per time point, the table will present:

- Per group: difference in LS-means versus the other GLPG0634 groups within the same regimen (q.d. or b.i.d.), together with a 95% confidence interval of this difference.
- Per group: p-value for pairwise comparison versus the other GLPG0634 groups. These p-values will be presented both as uncorrected p-values, and are derived as contrasts within

the overall model. As this is a purely exploratory analysis, no adjustment for multiplicity will be performed.

The “raw” SAS PROC output must be provided in a statistical appendix listing.

The table will show:

- LOCF in ITT
- observed-case in ITT (inferential statistics at Week 12 only).

Table to be created for the Week 12 interim analysis as well, cut off at Week 12.

**Table 14.2.1.1.14: ACR-N: Exploratory pairwise comparisons between regimens within each daily dose level**

ANCOVA model at each time point. Factors in the model are: regimen, dose, regimen\*dose interaction, region, and previous use of biologics. Placebo subjects will not be used in this model (as there is no q.d./b.i.d. regimen available for placebo). The regimen\*dose interaction will be removed from the final model when not significant at level 0.10 (two-sided type-III p-value). Regimen will have 2 categories: q.d. or b.i.d. Dose will have 3 categories, corresponding to the total daily dose: 50 mg/day, 100 mg/day, or 200 mg/day.

Per time point, the table will present:

- Per dose: difference in LS-means of q.d. versus b.i.d. within the same dose level (50, 100 or 200 mg/day), together with a 95% confidence interval of this difference.
- Per group: p-value for pairwise comparison of q.d. versus b.i.d. These p-values will be presented both as uncorrected p-values, and are derived as contrasts within the overall model. As this is a purely exploratory analysis, no adjustment for multiplicity will be performed.

The “raw” SAS PROC output must be provided in a statistical appendix listing.

The table will show:

- LOCF in ITT
- observed-case in ITT (inferential statistics at Week 12 only).

Table to be created for the Week 12 interim analysis as well, cut off at Week 12.

**Table 14.2.1.1.15: Time to ACR20 response: Summary with overall p-value and pairwise comparisons versus placebo**

Descriptive statistics per treatment group.

Between-group comparison (overall and all pairwise comparisons versus placebo)

P-values are generated from a Cox proportional hazards regression model. Factors in the model are: treatment, region, and previous use of biologics.

The table will present:

- Per group: N (number of events, number of censored, number in total), mean (SE), median, quartiles.

- Overall type-III p-values for treatment, region and previous use of biologics.
- Per group: p-value for pairwise comparison versus the placebo group. These p-values will be corrected for multiplicity according to Hommel's closed-testing method as implemented in SAS PROC MULTTEST. The uncorrected p-values will also be presented.

A frequency table of the time-to-event in categories (Week 1, Week 2, Week 4, Week 8, Week 12, never) will also be presented.

The "raw" SAS PROC output must be provided in a statistical appendix listing.

Population: ITT.

**Table 14.2.1.1.16: Time to ACR20 response: Exploratory pairwise comparisons among the dose groups within each regimen**

Cox proportional hazards regression model. Factors in the model are: treatment, region, and previous use of biologics. Placebo won't be used, so can be omitted from the overall model.

The table will present per group: p-value for pairwise comparison versus the other GLPG0634 groups. These p-values will be presented both as uncorrected p-values, and are derived as contrasts within the overall model. As this is a purely exploratory analysis, no adjustment for multiplicity will be performed.

The "raw" SAS PROC output must be provided in a statistical appendix listing.

Population: ITT.

**Table 14.2.1.1.17: Time to ACR20 response: Exploratory pairwise comparisons between regimens within each daily dose level**

Cox proportional hazards regression model. Factors in the model are: regimen, dose, regimen\*dose interaction, region, and previous use of biologics. Placebo subjects will not be used in this model (as there is no q.d./b.i.d. regimen available for placebo). The regimen\*dose interaction will be removed from the final model when not significant at level 0.10 (two-sided type-III p-value). Regimen will have 2 categories: q.d. or b.i.d. Dose will have 3 categories, corresponding to the total daily dose: 50 mg/day, 100 mg/day, or 200 mg/day.

The table will present per group: p-value for pairwise comparison of q.d. versus b.i.d. These p-values will be presented both as uncorrected p-values, and are derived as contrasts within the overall model. As this is a purely exploratory analysis, no adjustment for multiplicity will be performed.

The "raw" SAS PROC output must be provided in a statistical appendix listing.

Population: ITT.

**Table 14.2.1.1.18: Time to ACR50 response: Summary with overall p-value and pairwise comparisons versus placebo**

Repeat table of the ACR20 table.

Population: ITT.

**Table 14.2.1.1.19: Time to ACR50 response: Exploratory pairwise comparisons among the dose groups within each regimen**

Repeat table of the ACR20 table.

Population: ITT.

**Table 14.2.1.1.20: Time to ACR50 response: Exploratory pairwise comparisons between regimens within each daily dose level**

Repeat table of the ACR20 table.

Population: ITT.

**Table 14.2.1.1.21: Time to ACR70 response: Summary with overall p-value and pairwise comparisons versus placebo**

Repeat table of the ACR20 table.

Population: ITT.

**Table 14.2.1.1.22: Time to ACR70 response: Exploratory pairwise comparisons among the dose groups within each regimen**

Repeat table of the ACR20 table.

Population: ITT.

**Table 14.2.1.1.23: Time to ACR70 response: Exploratory pairwise comparisons between regimens within each daily dose level**

Repeat table of the ACR20 table.

Population: ITT.

### **8.1.3. Figures over the whole 24 weeks period**

**Figure 14.2.1.1.1: ACR20/50/70: Plot per week**

Vertical bar chart of the ACR20, ACR50 and ACR70 response rates for the 7 treatment groups. Separate plot per week.

The plot will show two methods of imputing missing values: LOCF and NRI.

Population: ITT.

**Figure 14.2.1.1.2: ACR20: Plot over time**

Line plot of the ACR20 response rates, with a separate line for each of the 7 treatment groups.

The plot will show two methods of imputing missing values: LOCF and NRI.

Population: ITT.

**Figure 14.2.1.1.3: ACR50: Plot over time**

Same plot for the ACR50 response.

The plot will show two methods of imputing missing values: LOCF and NRI.

Population: ITT.

**Figure 14.2.1.1.4: ACR70: Plot over time**

Same plot for the ACR70 response.

The plot will show two methods of imputing missing values: LOCF and NRI.

Population: ITT.

**Figure 14.2.1.1.5: ACR-N: Plot over time**

Mean (without SE) plot of the ACR-N, with a separate line for each of the 7 treatment groups. LOCF only.

Population: ITT.

**Figure 14.2.1.1.6: Time to ACR20 response: Kaplan-Meier plot**

Ascending Kaplan-Meier plot per treatment group.

Population: ITT.

**Figure 14.2.1.1.7: Time to ACR50 response: Kaplan-Meier plot**

Ascending Kaplan-Meier plot per treatment group.

Population: ITT.

**Figure 14.2.1.1.8: Time to ACR70 response: Kaplan-Meier plot**

Ascending Kaplan-Meier plot per treatment group.

Population: ITT.

**8.1.4. Tables after Week 12**

Tables will present only Weeks 12-16-20-24, unless specified otherwise. Tables will show the 15 arms of the “third efficacy analysis chapter”.

**Table 14.2.1.1.24: ACR20 after Week 12: Frequency table**

Frequency table per time point, and per treatment group (15 arms).

The table will show:

- NRI in ITT
- LOCF in ITT
- observed-case in ITT (weeks 12 + 24)

**Table 14.2.1.1.25: ACR50 after Week 12: Frequency table**

Repeat table of the ACR20 table.

The table will show:

- NRI in ITT
- LOCF in ITT
- observed-case in ITT (weeks 12 + 24)

**Table 14.2.1.1.26: ACR70 after Week 12: Frequency table**

Repeat table of the ACR20 table.

The table will show:

- NRI in ITT
- LOCF in ITT
- observed-case in ITT (weeks 12 + 24)

**Table 14.2.1.1.27: ACR-N after Week 12: Descriptive statistics**

Descriptive statistics per time point, and per treatment group (15 arms).

Frequency at Weeks 12 and 24: per treatment group (15 arms). ACR-N will be presented in the following categories:

<10,  
[10,20[,  
[20,30[,  
[30,40[,  
[40,50[,  
[50,60[,  
[60,70[,  
[70,80[,  
[80,90[,  
[90,100].

The table will show:

- LOCF in ITT
- observed-case in ITT (weeks 12 + 24)

**8.1.5. Figures after Week 12**

No figures planned.



## 8.1.6. Listings

### Listing 16.2.6.1: ACR20, ACR50, ACR70 and ACR-N

Listing per treatment group, per subject and per time point of the ACR20/50/70 and ACR-N and their components.

Population: ITT.

## 8.2. EULAR REMISSION

### 8.2.1. Parameters

EULAR remission:

- “YES” when  $TJC28 \leq 1$  and  $SJC28 \leq 1$  and  $CRP \leq 1$  mg/dL and patient’s global assessment of disease activity  $VAS \leq 1$  cm (or 10 mm).
- “NO” when at least 1 of the 4 components has a value  $> 1$ . Note: in case of missing components, it is still possible to qualify as “non remission” if at least 1 of the nonmissing components has a value  $> 1$ .
- “MISSING” in all remaining cases.

If any of the 4 components of the EULAR would be missing, then this missing component will be imputed using the value of the last preceding nonmissing visit.

### 8.2.2. Tables over the whole 24 weeks

Tables will present Baseline-Weeks 1-2-4-8-12-16-20-24. Similar analysis tables as for ACR20 will be made. Analysis population: ITT. NRI, LOCF and observed case (descriptive only at weeks 12 + 24).

**Table 14.2.1.2.1: EULAR remission: Summary with overall p-value and pairwise comparisons versus placebo**

**Table 14.2.1.2.2: EULAR remission: Exploratory pairwise comparisons among the dose groups within each regimen**

**Table 14.2.1.2.3: EULAR remission: Exploratory pairwise comparisons between regimens within each daily dose level**

### 8.2.3. Figures over the whole 24 weeks period

The same bar chart and line plot as for the ACR20 will be made. Population: ITT. NRI and LOCF imputations only.

**Figure 14.2.1.2.1: EULAR remission: Plot per week**

**Figure 14.2.1.2.2: EULAR remission: Plot over time**

## 8.2.4. Tables after Week 12

### Table 14.2.1.2.4: EULAR remission after Week 12: Frequency table

Table will present Weeks 12-16-20-24 for the 15 arms of “efficacy chapter 3”. A similar descriptive analysis as for ACR20 will be performed. Analysis population: ITT. NRI, LOCF and observed case (weeks 12 + 24).

## 8.2.5. Figures after Week 12

No figures planned.

## 8.2.6. Listings

### Listing 16.2.6.2: EULAR remission

Listing per treatment group, per subject and per time point of the EULAR remission category and its components.

Population: ITT.

## 8.3. SDAI AND CDAI SCORES

### 8.3.1. Parameters

$SDAI = TJC28 + SJC28 + \text{patient's disease activity VAS (in cm)} + \text{investigator's disease activity VAS (in cm)} + CRP \text{ (in mg/dL)}$ .

$CDAI = TJC28 + SJC28 + \text{patient's disease activity VAS (in cm)} + \text{investigator's disease activity VAS (in cm)}$ .

(TJC28 and SJC28 are the 28 joint counts for tenderness and swelling. See further sections on how to derive these components.)

CDAI and SDAI scores can only be derived when all components are nonmissing. If any of the components would be missing, then this missing component will be imputed using the value of the last preceding nonmissing visit.

Note that if for a specific visit both the regular CRP as well as the high-sensitive CRP are available, the high-sensitive CRP will be used for the derivation of the SDAI.

Apart from the actual scores, the changes from baseline and percent changes from baseline will be derived.

The SDAI is also categorized:

- High disease activity: SDAI > 26
- Moderate disease activity: SDAI in ] 11 , 26 ]
- Low disease activity: SDAI in ] 3.3 , 11 ]
- Remission: SDAI  $\leq$  3.3

The CDAI is also categorized:

- High disease activity: CDAI > 22
- Moderate disease activity: CDAI in ] 10 , 22 ]
- Mild disease activity: CDAI in ] 2.8 , 10 ]
- Remission: CDAI  $\leq$  2.8

Note that the percentage change from baseline remains missing if the baseline value itself equals zero.

### 8.3.2. Tables over the whole 24 weeks period

Tables will present Baseline-Weeks 1-2-4-8-12-16-20-24, unless specified otherwise.

Analysis population: ITT. LOCF and observed case (descriptive only at Baseline and Weeks 12 + 24).

#### Table 14.2.1.3.1: SDAI and CDAI: Descriptive statistics of the actual values

Descriptive statistics of the actual values per time point, and per treatment group.

Frequency per time point (all time points), and per treatment group. SDAI and CDAI will be presented in the categories defined above (remission, low, moderate, high).

The table will show:

- LOCF in ITT
- observed-case in ITT (Baseline, Week 12 and Week 24 only)

#### Table 14.2.1.3.2: SDAI and CDAI: Descriptive statistics of the change from baseline

Descriptive statistics of the changes from baseline per time point, and per treatment group.

The table will show:

- LOCF in ITT
- observed-case in ITT (weeks 12 + 24 only)

#### Table 14.2.1.3.3: SDAI and CDAI: Descriptive statistics of the percent change from baseline

Descriptive statistics of the percent changes from baseline per time point, and per treatment group.

Frequency table at Weeks 12 and 24 with cumulative percentages of the percent change from baseline, categorized as:

$\geq 0$ ,  
 [ -10 , 0 [,  
 [ -20 , -10 [,  
 [ -30 , -20 [,  
 [ -40 , -30 [,  
 [ -50 , -40 [,  
 [ -60 , -50 [,  
 [ -70 , -60 [,  
 [ -80 , -70 [,  
 [ -90 , -80 [,  
 [ -100 , -90 [.

The table will show:

- LOCF in ITT
- observed-case in ITT (weeks 12 + 24 only)

**Table 14.2.1.3.4: SDAI and CDAI: Summary with overall p-value and pairwise comparisons versus placebo**

ANCOVA model on the change from baseline at each time point. Factors in the model are: treatment, baseline value, region, and previous use of biologics.

Per time point, the table will present:

- Overall type-III p-value for all explanatory factors in the model.
- Per group: difference in LS-means versus the placebo group (GLPG0634 minus placebo), together with a 95% confidence interval of this difference.
- Per group: p-value for pairwise comparison versus the placebo group, derived from contrasts within the overall model. These p-values will be presented both as uncorrected p-values as well as corrected for multiplicity according to Hommel's closed-testing method (SAS PROC MULTTEST).

The "raw" SAS PROC output must be provided in a statistical appendix listing.

The table will show:

- LOCF in ITT

Table to be created for the Week 12 interim analysis as well, cut off at Week 12.

**Table 14.2.1.3.5: SDAI and CDAI: Exploratory pairwise comparisons among the dose groups within each regimen**

ANCOVA model on the changes from baseline at each time point. Factors in the model are: treatment, baseline value, region, and previous use of biologics. Placebo won't be used, so can be omitted from the overall model.

Per time point, the table will present:

- Per group: difference in LS-means versus the other GLPG0634 groups within the same regimen (q.d. or b.i.d.), together with a 95% confidence interval of this difference.
- Per group: p-value for pairwise comparison versus the other GLPG0634 groups. These p-values will be presented both as uncorrected p-values, and are derived as contrasts within the overall model. As this is a purely exploratory analysis, no adjustment for multiplicity will be performed.

The “raw” SAS PROC output must be provided in a statistical appendix listing.

The table will show:

- LOCF in ITT

Table to be created for the Week 12 interim analysis as well, cut off at Week 12.

**Table 14.2.1.3.6: SDAI and CDAI: Exploratory pairwise comparisons between regimens within each daily dose level**

ANCOVA model on the change from baseline at each time point. Factors in the model are: baseline value, regimen, dose, regimen\*dose interaction, region, and previous use of biologics. Placebo subjects will not be used in this model (as there is no q.d./b.i.d. regimen available for placebo). The regimen\*dose interaction will be removed from the final model when not significant at level 0.10 (two-sided type-III p-value). Regimen will have 2 categories: q.d. or b.i.d. Dose will have 3 categories, corresponding to the total daily dose: 50 mg/day, 100 mg/day, or 200 mg/day.

Per time point, the table will present:

- Per dose: difference in LS-means of q.d. versus b.i.d. within the same dose level (50, 100 or 200 mg/day), together with a 95% confidence interval of this difference.
- Per group: p-value for pairwise comparison of q.d. versus b.i.d. These p-values will be presented both as uncorrected p-values, and are derived as contrasts within the overall model. As this is a purely exploratory analysis, no adjustment for multiplicity will be performed.

The “raw” SAS PROC output must be provided in a statistical appendix listing.

The table will show:

- LOCF in ITT

Table to be created for the Week 12 interim analysis as well, cut off at Week 12.

**Table 14.2.1.3.7: SDAI and CDAI remission: Summary with overall p-value and pairwise comparisons versus placebo**

Similar table as for ACR20, presenting the rates of SDAI remission and CDAI remission at Week 12 and Week 24.

The table will show:

- NRI in ITT
- LOCF in ITT

- observed-case in ITT (descriptive at Week 12 + 24; inferential statistics at Week 12 only).

**Table 14.2.1.3.8: SDAI and CDAI remission: Exploratory pairwise comparisons among the dose groups within each regimen**

**Table 14.2.1.3.9: SDAI and CDAI remission: Exploratory pairwise comparisons between regimens within each daily dose level**

### **8.3.3. Figures over the whole 24 weeks period**

LOCF only. Population: ITT.

**Figure 14.2.1.3.1: SDAI and CDAI: Plot of the actual values over time**

Mean (without SE) plot of the actual values over time up to (and including) Week 24, with a separate line for each of the 7 treatment groups.

Population: ITT.

**Figure 14.2.1.3.2: SDAI and CDAI: Plot of the changes from baseline over time**

Mean (without SE) plot of the changes from baseline over time up to (and including) Week 24, with a separate line for each of the 7 treatment groups. With a horizontal reference line at 0, indicating no change.

Population: ITT.

**Figure 14.2.1.3.3: SDAI and CDAI: Plot of the percent changes from baseline over time**

Repeat plot for the percent change.

Population: ITT.

### **8.3.4. Tables after Week 12**

Tables will present Weeks 12-16-20-24, unless specified otherwise, for the 15 arms of “efficacy chapter 3”. A similar descriptive analysis will be performed. Analysis population: ITT. LOCF and observed case (weeks 12 +24).

**Table 14.2.1.3.10: SDAI and CDAI after Week 12: Descriptive statistics of the actual values**

**Table 14.2.1.3.11: SDAI and CDAI after Week 12: Descriptive statistics of the change from baseline**

**Table 14.2.1.3.12: SDAI and CDAI after Week 12: Descriptive statistics of the percent change from baseline**

### **8.3.5. Figures after Week 12**

No figures planned.

### 8.3.6. Listings

#### Listing 16.2.6.3: SDAI and CDAI

Listing per treatment group, per subject and per time point of the SDAI and CDAI scores and their components. SDAI and CDAI: actual values, changes from baseline, percent changes from baseline, and the categorization.

Population: ITT.

## 8.4. DAS28(CRP)

### 8.4.1. Parameters

$$\text{DAS28(CRP)} = 0.56 \times \sqrt{\text{TJC28}} + 0.28 \times \sqrt{\text{SJC28}} + 0.36 \times \text{Ln}[1 + \text{CRP (in mg/L)}] + 0.014 \times \text{patient's disease activity VAS (in mm)} + 0.96$$

(TJC28 and SJC28 are the 28 joint counts for tenderness and swelling. See further sections on how to derive these components. Ln = natural logarithm.)

DAS28(CRP) scores can only be derived when all components are nonmissing. If any of the components of the DAS28(CRP) would be missing, then this missing component will be imputed using the value of the last preceding nonmissing visit.

Note that if for a specific visit both the regular CRP as well as the high-sensitive CRP are available, the high-sensitive CRP will be used for the derivation of the DAS28(CRP).

Note that the percentage change from baseline remains missing if the baseline value itself equals zero.

Next to the actual DAS28(CRP) value, also the changes and percent changes from baseline will be derived.

Categorization of the DAS28(CRP) scores:

- High disease activity: > 5.1
- Moderate disease activity: ] 3.2 , 5.1 ]
- Low disease activity: [ 2.6 , 3.2 ]
- Remission: < 2.6

A second categorization of the DAS28(CRP) into EULAR response categories (none, moderate, good) will be done according to the following:

Actual DAS28(CRP)	Improvement in DAS28(CRP) from baseline:		
	> 1.2	] 0.6, 1.2]	≤ 0.6
≤ 3.2	Good	Moderate	None
]3.2, 5.1]	Moderate	Moderate	None
> 5.1	Moderate	None	None

#### 8.4.2. Tables over the whole 24 weeks period

Tables will present Baseline-Weeks 1-2-4-8-12-16-20-24, unless specified otherwise. A similar analysis as to SDAI and CDAI will be performed. Analysis population: ITT. LOCF and observed case (descriptive at Baseline, Week 12 and Week 24; inferential statistics only at Week 12).

##### Table 14.2.1.4.1: DAS28(CRP): Descriptive statistics of the actual values

(descriptive statistics, including a frequency table at all time points)

Analysis population: ITT and per-protocol.

Table to be created for the Week 12 interim analysis as well, cut off at Week 12. ITT only.

##### Table 14.2.1.4.2: DAS28(CRP): Descriptive statistics of the change from baseline

(descriptive statistics, including a shift table at all time points)

Analysis population: ITT and per-protocol.

Table to be created for the Week 12 interim analysis as well, cut off at Week 12. ITT only.

##### Table 14.2.1.4.3: DAS28(CRP): Descriptive statistics of the percent change from baseline

(descriptive statistics, including a frequency table at Weeks 12 + 24)

Analysis population: ITT and per-protocol.

Table to be created for the Week 12 interim analysis as well, cut off at Week 12. ITT only.

##### Table 14.2.1.4.4: DAS28(CRP): Summary with overall p-value and pairwise comparisons versus placebo

Analysis population: ITT and per-protocol.

Table to be created for the Week 12 interim analysis as well, cut off at Week 12. ITT only.

##### Table 14.2.1.4.5: DAS28(CRP): Exploratory pairwise comparisons among the dose groups within each regimen

Analysis population: ITT and per-protocol.



Table to be created for the Week 12 interim analysis as well, cut off at Week 12. ITT only.

**Table 14.2.1.4.6: DAS28(CRP): Exploratory pairwise comparisons between regimens within each daily dose level**

Analysis population: ITT and per-protocol.

Table to be created for the Week 12 interim analysis as well, cut off at Week 12. ITT only.

**Table 14.2.1.4.7: DAS28(CRP) change from baseline at Week 12: Tabulation versus baseline characteristics**

The table will show LOCF only.

Descriptive statistics of the DAS28(CRP) change from baseline at Week 12 per treatment group, split up according to the following baseline categories:

- Demographics:
  - Region (as used in the central randomization)
  - Country
  - Gender
  - Age categories:
    - < 45 ;
    - [45,65[ ;
    - [65,75[ ;
    - ≥ 75 years
  - BMI categories:
    - ≤ 18.5;
    - ] 18.5 , 25.0 ] ;
    - ] 25 , 30 [,
    - ≥ 30 kg/m<sup>2</sup>
- Baseline disease characteristics:
  - Duration of RA:
    - < 1;
    - [ 1 , 2 [ ;
    - [ 2 , 5 [ ;
    - [ 5 , 10 [ ;
    - [ 10 , 20 [ ;
    - ≥ 20 years
  - Baseline CRP:
    - <0.7 x ULN;
    - [0.7,1.0] x ULN;
    - ]1.0,1.2[ x ULN;
    - ≥ 1.2 x ULN
  - Baseline TJC68: < 20, ≥ 20
  - Baseline SJC66: < 20, ≥ 20
  - Baseline DAS28(CRP): high, moderate, low, remission (see further section on DAS28 for the definition of these categories)
  - RF status

- Anti-CCP status
- Previous use of:
  - Biological (anti-TNF or other biological DMARD) yes/no
  - anti-TNF yes/no,
  - other biological DMARD yes/no,
  - conventional DMARD yes/no,
  - NSAID yes/no,
  - steroid yes/no,
  - analgesic yes/no,
  - other RA medication yes/no.

Population: ITT.

Table to be created for the Week 12 interim analysis as well.

**Table 14.2.1.4.8: DAS28(CRP) change from baseline at Week 12: Exploratory ANCOVA model containing baseline characteristics**

The table will show LOCF only.

ANCOVA model of the DAS28(CRP) change from baseline at Week 12, with the following exploratory factors:

- Treatment group
- Demographics:
  - Region
  - Country
  - Gender
  - Age as continuous variable
  - BMI as continuous variable
- Baseline disease characteristics:
  - Duration of RA as continuous variable
  - Baseline CRP as continuous variable
  - Baseline TJC68 and SJC66 as continuous variables
  - Baseline DAS28(CRP) as continuous variable
  - RF status
  - anti-CCP status
  - Previous use of:
    - Biological (anti-TNF or other biological DMARD) yes/no
    - anti-TNF yes/no,
    - other biological DMARD yes/no,
    - conventional DMARD yes/no,
    - NSAID yes/no,
    - steroid yes/no,
    - analgesic yes/no,
    - other RA medication yes/no.

All these factors will be tested using a stepwise selection approach. Significance level to enter the model: 0.10; significance level to stay in the model: 0.05. The treatment groups will always be kept in the model, even when not significant at level 0.05. The table will present only the final model with type-III p-value and LS-mean estimates. The “raw” SAS PROC output must be provided in a statistical appendix listing, and will contain all details on the model building steps.

Population: ITT.

**Table 14.2.1.4.9: DAS28(CRP) remission: Summary with overall p-value and pairwise comparisons versus placebo**

Similar table as for ACR20, presenting the rates of DAS28(CRP) remission at Week 12 and Week 24.

The table will show:

- NRI in ITT
- LOCF in ITT
- observed-case in ITT (descriptive at Week 12 + 24; inferential statistics at Week 12 only).

**Table 14.2.1.4.10: DAS28(CRP) remission: Exploratory pairwise comparisons among the dose groups within each regimen**

**Table 14.2.1.4.11: DAS28(CRP) remission: Exploratory pairwise comparisons between regimens within each daily dose level**

**Table 14.2.1.4.12: DAS28(CRP) remission and low disease activity: Summary with overall p-value and pairwise comparisons versus placebo**

Similar table as for the remission, but now combining “remission” and “low disease activity” into one percentage.

**Table 14.2.1.4.13: DAS28(CRP) remission and low disease activity: Exploratory pairwise comparisons among the dose groups within each regimen**

**Table 14.2.1.4.14: DAS28(CRP) remission and low disease activity: Exploratory pairwise comparisons between regimens within each daily dose level**

**Table 14.2.1.4.15: DAS28(CRP) EULAR response: Frequency table**

Frequency per time point, and per treatment group. EULAR response will be presented in the categories defined above (none, moderate, good).

Overall p-value and pairwise comparisons versus placebo: Cochran-Mantel-Haenszel test for row mean score differences, controlling for region and prior biologics. These p-values will be corrected for multiplicity according to Hommel’s closed-testing method as implemented in SAS PROC MULTTEST. The uncorrected p-values will also be presented.

Population: ITT and per-protocol.

**Table 14.2.1.4.16: DAS28(CRP) EULAR response: Exploratory pairwise comparisons among the dose groups within each regimen**

Analysis population: ITT and per-protocol.

Comparisons between the doses within each regimen: Cochran-Mantel-Haenszel test for row mean score differences, controlling for region and prior biologics. As this is purely exploratory, no correction for multiplicity will be done.

**Table 14.2.1.4.17: DAS28(CRP) EULAR response: Exploratory pairwise comparisons between regimens within each daily dose level**

Analysis population: ITT and per-protocol.

Comparisons between the regimens within each daily dose level: Cochran-Mantel-Haenszel test for row mean score differences, controlling for region and prior biologics. The placebo subjects will not be used for this comparison. As this is purely exploratory, no correction for multiplicity will be done.

**8.4.3. Figures over the whole 24 weeks period**

Same type of mean line plots as for the SDAI and CDAI. Population: ITT. LOCF only.

**Figure 14.2.1.4.1: DAS28(CRP): Plot of the actual values over time**

**Figure 14.2.1.4.2: DAS28(CRP): Plot of the changes from baseline over time**

**Figure 14.2.1.4.3: DAS28(CRP): Plot of the percent changes from baseline over time**

**8.4.4. Tables after Week 12**

Tables will present Weeks 12-16-20-24, unless specified otherwise, for the 15 arms of “efficacy chapter 3”. A similar descriptive analysis as to SDAI and CDAI will be performed. Analysis population: ITT. LOCF and observed case (weeks 12 + 24).

**Table 14.2.1.4.18: DAS28(CRP) after Week 12: Descriptive statistics of the actual values**

**Table 14.2.1.4.19: DAS28(CRP) after Week 12: Descriptive statistics of the change from baseline**

**Table 14.2.1.4.20: DAS28(CRP) after Week 12: Descriptive statistics of the percent change from baseline**

**Table 14.2.1.4.21: DAS28(CRP) EULAR response after Week 12: Frequency table**

**8.4.5. Figures after Week 12**

No figures planned.

**8.4.6. Listings**

**Listing 16.2.6.4: DAS28(CRP): Original data**

Listing per treatment group, per subject and per time point of the DAS28(CRP) score (actual values, changes from baseline, percent changes from baseline, and the two categorizations), as well as the DAS28 components.

Population: ITT.

## 8.5. ACR COMPONENT: TENDER AND SWOLLEN JOINT COUNTS

### 8.5.1. Parameters

$$\text{TJC68} = (\text{total number of tender joints}) \times \left( \frac{68}{\text{Total number of assessed joints}} \right)$$

$$\text{SJC66} = (\text{total number of swollen joints}) \times \left( \frac{66}{\text{Total number of assessed joints}} \right)$$

TJC28 and SJC28 will be calculated using the same formula but only selecting the 28 relevant<sup>1</sup> joints:

$$\text{TJC28} = (\text{total number of tender joints}) \times \left( \frac{28}{\text{Total number of assessed joints}} \right)$$

$$\text{SJC28} = (\text{total number of swollen joints}) \times \left( \frac{28}{\text{Total number of assessed joints}} \right)$$

For all 4 parameters, the actual values, changes from baseline and percent changes from baseline will be derived and analyzed.

The above formulae already deal with any potential missing values. If more than 50% of the joints is scored temporarily or permanently unassessable or simply missing, then the total joint count will be set to “missing”.

Note that the percentage change from baseline remains missing if the baseline value itself equals zero.

### 8.5.2. Tables over the whole 24 weeks period

Tables will present Screening-Baseline-Weeks 1-2-4-8-12-16-20-24, unless specified otherwise. The same analysis tables as DAS28(CRP) will be made. Analysis population: ITT. LOCF and observed case (descriptive only; baseline and weeks 12 + 24 only).

**Table 14.2.1.5.1: Joint counts: Descriptive statistics of the actual values**

**Table 14.2.1.5.2: Joint counts: Descriptive statistics of the change from baseline**

**Table 14.2.1.5.3: Joint counts: Descriptive statistics of the percent change from baseline**

---

<sup>1</sup> Relevant 28 joints are: shoulder, elbow, wrist, metacarpophalangeal and proximal interphalangeal; knee.

(descriptive statistics, and a frequency table at weeks 12 + 24)

**Table 14.2.1.5.4: Joint counts: Summary with overall p-value and pairwise comparisons versus placebo**

Table to be created for the Week 12 interim analysis as well, cut off at Week 12.

**Table 14.2.1.5.5: Joint counts: Exploratory pairwise comparisons among the dose groups within each regimen**

Table to be created for the Week 12 interim analysis as well, cut off at Week 12.

**Table 14.2.1.5.6: Joint counts: Exploratory pairwise comparisons between regimens within each daily dose level**

Table to be created for the Week 12 interim analysis as well, cut off at Week 12.

### **8.5.3. Figures over the whole 24 weeks period**

Same analysis as for the DAS28(CRP). Population: ITT. LOCF only.

**Figure 14.2.1.5.1: Joint counts: Plot of the actual values over time**

**Figure 14.2.1.5.2: Joint counts: Plot of the changes from baseline over time**

**Figure 14.2.1.5.3: Joint counts: Plot of the percent changes from baseline over time**

### **8.5.4. Tables after Week 12**

Tables will present Weeks 12-16-20-24, unless specified otherwise, for the 15 arms of “efficacy analysis chapter 3”. The same descriptive analysis as DAS28(CRP) will be performed. Analysis population: ITT. LOCF and observed case (weeks 12 + 24).

**Table 14.2.1.5.7: Joint counts after Week 12: Descriptive statistics of the actual values**

**Table 14.2.1.5.8: Joint counts after Week 12: Descriptive statistics of the change from baseline**

**Table 14.2.1.5.9: Joint counts after Week 12: Descriptive statistics of the percent change from baseline**

### **8.5.5. Figures after Week 12**

No figures planned.

### **8.5.6. Listings**

**Listing 16.2.6.5.1: Joint counts: Identification of unassessable joints**

Listing per treatment group, per subject and per time point of the original 68/66 joints, selecting only the temporarily or permanently unassessable joints. There will be no listing of all assessed joints.

Population: ITT.

#### **Listing 16.2.6.5.2: Joint counts: Derived data**

Listing per treatment group, per subject and per time point of the TJC68, SJC66, TJC28 and SJC28: actual values, changes from baseline and percent changes from baseline.

Population: ITT.

## **8.6. ACR COMPONENT: VAS SCALES**

### **8.6.1. Parameters**

Assessments (each in mm, ranging from 0 to 100 with smaller scores indicating a better health condition):

- Patient's pain VAS
- Patient's global assessment of disease activity VAS
- Investigator's global assessment of disease activity VAS

For all 3 parameters, the actual values, changes from baseline and percent changes from baseline will be derived and analyzed.

Note that the percentage change from baseline remains missing if the baseline value itself equals zero.

Blinded data prior to database lock has revealed extreme values in the %changes from baseline for subjects with a very low baseline value (0, 1, 2 ...). No data will be excluded from the analysis; the data will go into the analysis as they are.

### **8.6.2. Tables over the whole 24 weeks period**

Tables will present Baseline-Weeks 1-2-4-8-12-16-20-24, unless specified otherwise. The screening visit will only be listed (is already analyzed in the baseline disease characteristics tables). The same analysis tables as DAS28(CRP) will be made. Analysis population: ITT. LOCF and observed case (descriptive only at Baseline and Weeks 12 + 24).

**Table 14.2.1.6.1: VAS scores: Descriptive statistics of the actual values**

**Table 14.2.1.6.2: VAS scores: Descriptive statistics of the change from baseline**

**Table 14.2.1.6.3: VAS scores: Descriptive statistics of the percent change from baseline**

(descriptive statistics, and a frequency table at weeks 12 + 24)

**Table 14.2.1.6.4: VAS scores: Summary with overall p-value and pairwise comparisons versus placebo**

Table to be created for the Week 12 interim analysis as well, cut off at Week 12.

**Table 14.2.1.6.5: VAS scores: Exploratory pairwise comparisons among the dose groups within each regimen**

Table to be created for the Week 12 interim analysis as well, cut off at Week 12.

**Table 14.2.1.6.6: VAS scores: Exploratory pairwise comparisons between regimens within each daily dose level**

Table to be created for the Week 12 interim analysis as well, cut off at Week 12.

**8.6.3. Figures over the whole 24 weeks study period**

Same analysis as for the DAS28(CRP). Population: ITT. LOCF only.

**Figure 14.2.1.6.1: VAS scores: Plot of the actual values over time****Figure 14.2.1.6.2: VAS scores: Plot of the changes from baseline over time****Figure 14.2.1.6.3: VAS scores: Plot of the percent changes from baseline over time****8.6.4. Tables after Week 12**

Tables will present Weeks 12-16-20-24, unless specified otherwise, for the 15 arms of “efficacy chapter 3”. The same descriptive analysis as DAS28(CRP) will be performed. Analysis population: ITT. LOCF and observed case (weeks 12 + 24).

**Table 14.2.1.6.7: VAS scores after Week 12: Descriptive statistics of the actual values****Table 14.2.1.6.8: VAS scores after Week 12: Descriptive statistics of the change from baseline****Table 14.2.1.6.9: VAS scores after Week 12: Descriptive statistics of the percent change from baseline****8.6.5. Figures after Week 12**

No figures planned.

**8.6.6. Listings****Listing 16.2.6.6: VAS scores**

Listing per treatment group, per subject and per time point of the 3 VAS scores: actual values, changes from baseline and percent changes from baseline.

Population: ITT.



## 8.7. ACR COMPONENT: HAQ-DI

### 8.7.1. Parameters

The HAQ-DI consists of 20 individual items.

The pain VAS of the HAQ-DI is analyzed separately as component of the ACR20/50/70, not in this chapter on HAQ-DI.

The HAQ-DI total scores are:

- 8 domain subscales: dressing, arising, eating, hygiene, walking, reaching, grip, activities
- Overall total score

Derivation of these summary scores:

- There are four response options ranging from “Without any difficulty” to “Unable to do”, scored from 0 to 3.
- Step 1: Derive the 8 category scores by using the highest sub-category score from each domain.
- Step 2: Adjust for use of aids/devices and/or help from another person when indicated:
  - Adjust the score for a domain by increasing a 0 or a 1 to a 2.
  - If a patient's highest score for that sub-category is a 2 it remains a 2, and if a 3, it remains a 3.
- Step 3: Divide the summed domain scores by the number of domains answered to obtain a HAQ-DI score of 0-3 (3=worst functioning). There must be a nonmissing score for at least six of the eight domains. Otherwise a HAQ-DI score cannot be computed.

If any of the 20 items would be missing, then this missing component will be imputed using the value of the last preceding nonmissing visit.

The analysis will present the actual values, changes from baseline and percent changes from baseline.

The analysis will present the subscores only at baseline, Week 12 and Week 24. Not at intermediate time points. (Except in the mean plots, where all time points will be presented.)

Note that the percentage change from baseline remains missing if the baseline value itself equals zero.

### 8.7.2. Tables over the whole 24 weeks period

Tables will present Baseline-Weeks 1-2-4-8-12-16-20-24, unless specified otherwise. The same analysis tables as DAS28(CRP) will be made. Analysis population: ITT. LOCF and observed case (descriptive only at baseline and weeks 12 + 24).

#### **Table 14.2.1.7.1: HAQ-DI total and domain scores: Descriptive statistics of the actual values**

Descriptive statistics at all time points. Frequency table at all time points of values  $\leq$  or  $>$  0.5 in the total HAQ-DI score. Subjects with a score below 0.5 are considered to have normal physical function.

**Table 14.2.1.7.2: HAQ-DI total and domain scores: Descriptive statistics of the change from baseline**

Descriptive statistics at all time points. Frequency table at all time points of the changes in total HAQ-DI score:

$\leq$  -0.8  
]-0.8 , -0.5]  
]-0.5 , -0.3]  
]-0.3 , -0.22]  
 $>$  -0.22

**Table 14.2.1.7.3: HAQ-DI total and domain scores: Descriptive statistics of the percent change from baseline**

Descriptive statistics, and a frequency table (weeks 12 + 24).

**Table 14.2.1.7.4: HAQ-DI total and domain scores: Summary with overall p-value and pairwise comparisons versus placebo**

Table to be created for the Week 12 interim analysis as well, cut off at Week 12.

**Table 14.2.1.7.5: HAQ-DI total and domain scores: Exploratory pairwise comparisons among the dose groups within each regimen**

Table to be created for the Week 12 interim analysis as well, cut off at Week 12.

**Table 14.2.1.7.6: HAQ-DI total and domain scores: Exploratory pairwise comparisons between regimens within each daily dose level**

Table to be created for the Week 12 interim analysis as well, cut off at Week 12.

### **8.7.3. Figures over the whole 24 weeks period**

Same analysis as for the DAS28(CRP). Population: ITT. LOCF only.

**Figure 14.2.1.7.1: HAQ-DI total and domain scores: Plot of the actual values over time**

**Figure 14.2.1.7.2: HAQ-DI total and domain scores: Plot of the changes from baseline over time**

**Figure 14.2.1.7.3: HAQ-DI total and domain scores: Plot of the percent changes from baseline over time**

#### **8.7.4. Tables after Week 12**

Tables will present Weeks 12-16-20-24, unless specified otherwise, for the 15 arms of “efficacy chapter 3”. The same descriptive analysis as DAS28(CRP) will be performed. Analysis population: ITT. LOCF and observed case (weeks 12 + 24).

**Table 14.2.1.7.7: HAQ-DI total and domain scores after Week 12: Descriptive statistics of the actual values**

**Table 14.2.1.7.8: HAQ-DI total and domain scores after Week 12: Descriptive statistics of the change from baseline**

**Table 14.2.1.7.9: HAQ-DI total and domain scores after Week 12: Descriptive statistics of the percent change from baseline**

#### **8.7.5. Figures after Week 12**

No figures planned.

#### **8.7.6. Listings**

**Listing 16.2.6.7: HAQ-DI: Derived summary scores**

Listing per treatment group, per subject and per time point of all derived summary scores (the overall score and the 8 domain scores): actual values, changes from baseline and percent changes from baseline. Total HAQ-DI values of  $\leq 0.5$  as well as changes from baseline in total HAQ-DI of  $\leq -0.22$  will be flagged in the listing. Note that the pain VAS is already listed elsewhere so will be omitted from the HAQ-DI listing.

Population: ITT.

### **8.8. ACR COMPONENT: SERUM CRP**

#### **8.8.1. Parameters**

Serum CRP (mg/L) will be analyzed as raw values, change from baseline and percent change from baseline.

Note that if for a specific visit both the regular CRP as well as the high-sensitive CRP are available, the high-sensitive CRP will be used in the analysis tables and figures. The regular CRP will then only be listed.

CRP and hsCRP values will also be categorized according to the lab’s normal ranges: normal (value  $\leq$  ULN) or high (value  $>$  ULN), where ULN = upper limit of the normal range as provided by the central lab. Note that the same ULN (8.999) will be applied to both CRP and hsCRP.

Note that CRP and hsCRP will be excluded from the lab safety analysis.

Note that the percentage change from baseline remains missing if the baseline value itself equals zero.

## 8.8.2. Tables over the whole 24 weeks period

Tables will present Screening-Baseline-Weeks 1-2-4-8-12-16-20-24, unless specified otherwise. The same analysis tables as DAS28(CRP) will be made. Analysis population: ITT. LOCF and observed case (descriptive only at baseline and weeks 12 + 24).

### **Table 14.2.1.8.1: Serum CRP (mg/L): Descriptive statistics of the actual values**

### **Table 14.2.1.8.2: Serum CRP (mg/L): Descriptive statistics of the change from baseline**

### **Table 14.2.1.8.3: Serum CRP (mg/L): Descriptive statistics of the percent change from baseline**

(descriptive statistics at all time points; frequency table at weeks 12 + 24)

Overall as well as split up for subjects with a baseline CRP > ULN versus  $\leq$  ULN.

### **Table 14.2.1.8.4: Serum CRP (mg/L): Shift table versus baseline**

Shift table per treatment group and time point. The table will present the shift in category (normal/high/missing) at each post-baseline time point versus the baseline Day 1 category (normal/high/missing).

Population: ITT population.

### **Table 14.2.1.8.5: Serum CRP (mg/L): Treatment-emergent normalizations per time point**

Frequency table of the treatment-emergent CRP normalizations (i.e., a value shifting from an abnormally high baseline to a normal value a post-baseline time point) per treatment group and time point. Subjects with a missing or normal baseline CRP will not be used. Time points with a missing post-baseline result will not be used either. No formal between-group comparisons.

Note: normal/abnormal is derived according to the central lab's upper limit of normal.

Population: ITT population.

Table to be created for the Week 12 interim analysis as well, cut off at Week 12.

### **Table 14.2.1.8.6: Serum CRP (mg/L): Summary with overall p-value and pairwise comparisons versus placebo**

Table to be created for the Week 12 interim analysis as well, cut off at Week 12.

### **Table 14.2.1.8.7: Serum CRP (mg/L): Exploratory pairwise comparisons among the dose groups within each regimen**

Table to be created for the Week 12 interim analysis as well, cut off at Week 12.

**Table 14.2.1.8.8: Serum CRP (mg/L): Exploratory pairwise comparisons between regimens within each daily dose level**

Table to be created for the Week 12 interim analysis as well, cut off at Week 12.

**8.8.3. Figures over the whole 24 weeks period**

Same analysis as for the DAS28(CRP). Population: ITT. LOCF only.

**Figure 14.2.1.8.1: Serum CRP (mg/L): Plot of the actual values over time****Figure 14.2.1.8.2: Serum CRP (mg/L): Plot of the changes from baseline over time****Figure 14.2.1.8.3: Serum CRP (mg/L): Plot of the percent changes from baseline over time****8.8.4. Tables after Week 12**

Tables will present Weeks 12-16-20-24, unless specified otherwise, for the 15 arms of “efficacy chapter 3”. The same descriptive analysis as DAS28(CRP) will be performed. Analysis population: ITT. LOCF and observed case (weeks 12 + 24).

**Table 14.2.1.8.9: Serum CRP (mg/L) after Week 12: Descriptive statistics of the actual values****Table 14.2.1.8.10: Serum CRP (mg/L) after Week 12: Descriptive statistics of the change from baseline****Table 14.2.1.8.11: Serum CRP (mg/L) after Week 12: Descriptive statistics of the percent change from baseline****Table 14.2.1.8.12: Serum CRP (mg/L) after Week 12: Shift table versus baseline****Table 14.2.1.8.13: Serum CRP (mg/L) after Week 12: Treatment-emergent normalizations per time point****8.8.5. Figures after Week 12**

No figures planned.

**8.8.6. Listings****Listing 16.2.6.8: Serum CRP (mg/L)**

Listing per treatment group, per subject and per time point of the raw CRP and hsCRP values, changes from baseline and percent changes from baseline. A flag “H” will be added for values above the upper limit of normal.

Population: ITT.

## 8.9. QUALITY OF LIFE: FACIT FATIGUE SCALE

### 8.9.1. Parameters

The scale (version 4) consists of 13 items:

1. I feel fatigued
2. I feel weak all over
3. I feel listless (“washed out”)
4. I feel tired
5. I have trouble starting things because I am tired
6. I have trouble finishing things because I am tired
7. I have energy
8. I am able to do my usual activities
9. I need to sleep during the day
10. I am too tired to eat
11. I need help doing my usual activities
12. I am frustrated by being too tired to do the things I want to do
13. I have to limit my social activity because I am tired

Each item is scored on an ordinal 5-point scale: Not At All; A Little Bit; Somewhat; Quite A Bit; Very Much.

The items must be scored as follows: 4=Not At All; 3=A Little Bit; 2=Somewhat; 1=Quite A Bit; 0=Very Much. Items 7 and 8 which are scored reversely.

The total FACIT fatigue score is then derived, and ranges from 0 to 52 with a higher score indicating a better quality of life:

$$\text{FACIT fatigue score} = (\text{Sum of all the scores}) \times \left( \frac{13}{\text{Number of items actually answered}} \right)$$

The FACIT fatigue score will only be derived when at least 7 of the 13 items are available, and will remain missing when 6 or less of the 13 items are available.

The analysis will use actual values, changes from baseline and percent changes from baseline.

Note that the percentage change from baseline remains missing if the baseline value itself equals zero.

### 8.9.2. Tables over the whole 24 weeks period

Tables will present Weeks 4, 12 and 24, unless specified otherwise. The same analysis as DAS28(CRP) will be performed. Analysis population: ITT. LOCF and observed case (descriptive only at baseline, week 12 and week 24).

**Table 14.2.1.9.1: FACIT fatigue score: Descriptive statistics of the actual values**

**Table 14.2.1.9.2: FACIT fatigue score: Descriptive statistics of the change from baseline**

**Table 14.2.1.9.3: FACIT fatigue score: Descriptive statistics of the percent change from baseline**

(descriptive statistics at all time points; frequency table at weeks 12 + 24)

**Table 14.2.1.9.4: FACIT fatigue score: Summary with overall p-value and pairwise comparisons versus placebo****Table 14.2.1.9.5: FACIT fatigue score: Exploratory pairwise comparisons among the dose groups within each regimen****Table 14.2.1.9.6: FACIT fatigue score: Exploratory pairwise comparisons between regimens within each daily dose level****8.9.3. Figures over the whole 24 weeks period**

Line plots showing baseline, Week 4, 12 and 24. Same analysis as for the DAS28(CRP). Population: ITT. LOCF only.

**Figure 14.2.1.9.1: FACIT fatigue score: Plot of the actual values over time****Figure 14.2.1.9.2: FACIT fatigue score: Plot of the changes from baseline over time****Figure 14.2.1.9.3: FACIT fatigue score: Plot of the percent changes from baseline over time****8.9.4. Tables after Week 12**

Tables will present Weeks 4, 12 and 24, unless specified otherwise, for the 15 arms of “efficacy chapter 3”. The same descriptive analysis as DAS28(CRP) will be performed. Analysis population: ITT. LOCF and observed case (weeks 12 + 24).

**Table 14.2.1.9.7: FACIT fatigue score after Week 12: Descriptive statistics of the actual values****Table 14.2.1.9.8: FACIT fatigue score after Week 12: Descriptive statistics of the change from baseline****Table 14.2.1.9.9: FACIT fatigue score after Week 12: Descriptive statistics of the percent change from baseline****8.9.5. Figures after Week 12**

No figures planned.

**8.9.6. Listings****Listing 16.2.6.9: Total FACIT fatigue score**

Listing per treatment group, per subject and per time point of the total score: actual values, changes from baseline and percent changes from baseline. The original 13 items will not be listed.

Population: ITT.

## 8.10. QUALITY OF LIFE: SF-36

### 8.10.1. Parameters

The SF-36 (version 2) consists of 36 items.

The 36 questions:

- *1: In general, would you say your health is*
- *2: Compared to one year ago, how would you rate your health in general now*
- *3A: Vigorous activities, such as running, lifting heavy objects, participating in strenuous sports*
- *3B: Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf*
- *3C: Lifting or carrying groceries*
- *3D: Climbing several flights of stairs*
- *3E: Climbing one flight of stairs*
- *3F: Bending, kneeling, or stooping*
- *3G: Walking more than a mile*
- *3H: Walking several hundred yards*
- *3I: Walking one hundred yards*
- *3J: Bathing or dressing yourself*
- *4A: As a result of your physical health, cut down on the amount of time you spent on work or other activities*
- *4B: As a result of your physical health, accomplished less than you would like*
- *4C: As a result of your physical health, were limited in the kind of work or other activities*
- *4D: As a result of your physical health, had difficulty performing the work or other activities (for example, it took extra effort)*
- *5A: As a result of any emotional problems, cut down the amount of time you spent on work or other activities*
- *5B: As a result of any emotional problems, accomplished less than you would like*
- *5C: As a result of any emotional problems, did work or other activities less carefully than usual*
- *6: During the past 4 weeks, to what extent has your physical health or emotional problems interfered with your normal social activities with family, friends, neighbors, or groups*
- *7: How much bodily pain have you had during the past 4 weeks*
- *8: During the past 4 weeks how much did pain interfere with your normal work (including both work outside the home and housework)*
- *9A: How much of the time during the past 4 weeks: did you feel full of life*



- 9B:How much of the time during the past 4 weeks: have you been very nervous
- 9C:How much of the time during the past 4 weeks: have you felt so down in the dumps that nothing could cheer you up
- 9D:How much of the time during the past 4 weeks: have you felt calm and peaceful
- 9E:How much of the time during the past 4 weeks: did you have a lot of energy
- 9F:How much of the time during the past 4 weeks: have you felt downhearted and depressed
- 9G:How much of the time during the past 4 weeks: did you feel worn out
- 9H:How much of the time during the past 4 weeks: have you been happy
- 9I:How much of the time during the past 4 weeks: did you feel tired
- 10:During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting your friends, relatives, etc.)
- 11A:I seem to get sick a little easier than other people
- 11B:I am as healthy as anybody I know
- 11C:I expect my health to get worse
- 11D:My health is excellent

**General comments:** The coding and scoring for each component (PCS, MCS) is done according to the tables below.

#### Coding and Scoring for the 8 SF-36 (Version 2) Scales

Scale	Item	Coding (makes use of QSRCCD)	Range of raw score	Normalization*
Physical functioning	Items 3a to 3j	Yes, limited a lot=1 Yes, limited a little=2 No, not limited=3	10-30	(S-10)/20x100
Role-physical	Items 4a to 4d	All of the time = 1 Most of the time = 2 Some of the time = 3 A little of the time = 4 None of the time = 5	4-20	(S-4)/16x100
Bodily pain	Item 7	None=6 Very mild=5.4 Mild=4.2 Moderate=3.1 Severe=2.2 Very severe=1	2-12	(S-2)/10x100
	Item 8 if both items 7 and 8 are answered	Not at all and item7 eq 'None' = 6 Not at all and item7 not equal 'None' = 5		

Scale	Item	Coding (makes use of QSRCCD)	Range of raw score	Normalization*
		A little bit = 4 Moderately = 3 Quite a bit = 2 Extremely = 1		
	<i>Item 8 if item 7 is not answered</i>	Not at all =6 A little bit = 4.75 Moderately =3.5 Quite a bit = 2.25 Extremely = 1		
General health	<i>Item 1</i>	Excellent = 5 Very good = 4.4 Good = 3.4 Fair = 2 Poor = 1	5-25	(S-5)/20x100
	<i>Items 11a and 11c</i>	Definitely true = 1 Mostly true = 2 Don't know = 3 Mostly false = 4 Definitely false = 5		
	<i>Items 11b and 11d</i>	Definitely true = 5 Mostly true = 4 Don't know = 3 Mostly false = 2 Definitely false = 1		
Vitality	<i>Items 9a and 9e</i>	All of the time = 5 Most of the time = 4 Some of the time = 3 A little of the time = 2 None of the time = 1	4-20	(S-4)/16x100
	<i>Items 9g and 9i</i>	All of the time = 1 Most of the time = 2 Some of the time = 3 A little of the time = 4 None of the time = 5		
Social functioning	<i>Item 6</i>	Not at all = 5 Slightly = 4 Moderately = 3 Quite a bit = 2	2-10	(S-2)/8x100

Scale	Item	Coding (makes use of QSRCCD)	Range of raw score	Normalization*
		Extremely = 1		
	<i>Item 10</i>	All of the time = 1 Most of the time = 2 Some of the time = 3 A little of the time = 4 None of the time = 5		
Role-emotional	<i>Items 5a to 5c</i>	All of the time = 1 Most of the time = 2 Some of the time = 3 A little of the time = 4 None of the time = 5	3-15	(S-3)/12x100
Mental health	<i>Items 9b, 9c and 9f</i>	All of the time = 1 Most of the time = 2 Some of the time = 3 A little of the time = 4 None of the time = 5	5-25	(S-5)/20x100
	<i>Items 9d and 9h</i>	All of the time = 5 Most of the time = 4 Some of the time = 3 A little of the time = 2 None of the time = 1		
*S = raw score = sum of item scores after coding				

For the 8 raw scale scores calculations, if at least 50% of the items comprising a scale have non-missing responses, impute the average score calculated for the non-missing scale items for the missing items of the same scale. If more than 50% of scale items comprising a scale are missing a response, set the scale to missing.

Then each scale score needs to be standardized (z-score) as below:

**Standardization of the SF-36 (Version 2) Scale Scores**

Item	Standardization (Z-score)
Physical functioning	$PF\_Z = (PF-83.29094)/23.75883$
Role-physical	$RP\_Z = (RP-82.50964)/25.52028$
Bodily pain	$BP\_Z = (BP-71.32527)/23.66224$
General health	$GH\_Z = (GH-70.84570)/20.97821$
Vitality	$VT\_Z = (VT-58.31411)/20.01923$
Social functioning	$SF\_Z = (SF-84.30250)/22.91921$
Role-emotional	$RE\_Z = (RE-87.39733)/21.43778$
Mental health	$MH\_Z = (MH-74.98685)/17.75604$

Note: Based on the 1998 general United States population.

Finally, the SF-36 physical and mental component transformed scores are calculated using the below conventions.

**SF-36 Aggregate Component and Transformed Scores**

Score	Calculation
<b>Aggregate Standardized Score</b>	
Physical component	$AGG\_PHYS = (PF\_Z * 0.42402) + (RP\_Z * 0.35119) +$ $(BP\_Z * 0.31754) + (GH\_Z * 0.24954) + (VT\_Z * 0.02877) +$ $(SF\_Z * -0.00753) + (RE\_Z * -0.19206) + (MH\_Z * -0.22069)$
Mental component	$AGG\_MENT = (PF\_Z * -0.22999) + (RP\_Z * -0.12329) +$ $(BP\_Z * -0.09731) + (GH\_Z * -0.01571) + (VT\_Z * 0.23534) +$ $(SF\_Z * 0.26876) + (RE\_Z * 0.43407) + (MH\_Z * 0.48581)$
<b>Transformed Summary</b>	
Transformed physical	$PCS = 50 + (AGG\_PHYS * 10)$
Transformed mental	$MCS = 50 + (AGG\_MENT * 10)$

In addition, the 8 scale scores will be reported in norm-based transformation. Norm-based transformation will be achieved by multiplying each z-score by 10 and adding 50 to it:

<b>Item</b>	<b>Norm-based transformation of the z-scores</b>
Physical functioning	Norm-based PF = 50 + (PF_Z*10)
Role-physical	Norm-based RP = 50 + (RP_Z*10)
Bodily pain	Norm-based BP = 50 + (BP_Z*10)
General health	Norm-based GH = 50 + (GH_Z*10)
Vitality	Norm-based VT = 50 + (VT_Z*10)
Social functioning	Norm-based SF = 50 + (SF_Z*10)
Role-emotional	Norm-based RE = 50 + (RE_Z*10)
Mental health	Norm-based MH = 50 + (MH_Z*10)

Derivation of these summary scores, including appropriate handling of missing items, will be done using the scoring software provided by the owner of the scale (QualityMetrics).

Please refer to the SF-36 v2 manual for additional information.

The analysis will present the 2 summary component measures and the 8 domains:

- Transformed physical component scale PCS and mental component scale MCS
- Norm-based physical functioning (PF), role physical (RF), bodily pain (BP), general health (GH), vitality (VT), social functioning (SF), role emotional (RE) and mental health (MH)

The analysis will present actual values, changes from baseline and percent changes from baseline in the above-mentioned scores.

Note that the percentage change from baseline remains missing if the baseline value itself equals zero.

### **8.10.2. Tables over the whole 24 weeks period**

Same analysis as for the DAS28(CRP). Population: ITT. Note that the SF-36 is only scored at baseline, Week 4, Week 12, and Week 24/EDV. LOCF and observed case (descriptive only at baseline and weeks 12 + 24). Tables will present the 8 domain scores and the 2 main scores.

#### **Table 14.2.1.10.1:SF-36: Descriptive statistics of the actual values**

#### **Table 14.2.1.10.2:SF-36: Descriptive statistics of the change from baseline**

#### **Table 14.2.1.10.3:SF-36: Descriptive statistics of the percent change from baseline**

(descriptive statistics at each time point; frequency table at weeks 12 + 24)

#### **Table 14.2.1.10.4:SF-36: Summary with overall p-value and pairwise comparisons versus placebo**

**Table 14.2.1.10.5: SF-36: Exploratory pairwise comparisons among the dose groups within each regimen****Table 14.2.1.10.6: SF-36: Exploratory pairwise comparisons between regimens within each daily dose level****8.10.3. Figures over the whole 24 weeks period**

Line plots showing baseline, Week 4, 12 and 24. Same analysis as for the DAS28(CRP). Population: ITT. LOCF only. Both main scores as well as the 8 subdomains will be presented.

**Figure 14.2.1.10.1: SF-36: Plot of the actual values over time****Figure 14.2.1.10.2: SF-36: Plot of the changes from baseline over time****Figure 14.2.1.10.3: SF-36: Plot of the percent changes from baseline over time****8.10.4. Tables after Week 12**

Tables will present Weeks 4, 12 and 24, unless specified otherwise, for the 15 arms of “efficacy chapter 3”. The same descriptive analysis as DAS28(CRP) will be performed. Analysis population: ITT. LOCF and observed case (weeks 12 + 24). Tables will present the 8 domain scores and the 2 main scores.

**Table 14.2.1.10.7: SF-36 after Week 12: Descriptive statistics of the actual values****Table 14.2.1.10.8: SF-36 after Week 12: Descriptive statistics of the change from baseline****Table 14.2.1.10.9: SF-36 after Week 12: Descriptive statistics of the percent change from baseline****8.10.5. Figures after Week 12**

No figures planned.

**8.10.6. Listings****Listing 16.2.6.10: SF-36: Derived summary scores**

Listing per treatment group, per subject and per time point of all derived summary scores (the 2 component scores and the 8 domain scores): actual values, changes from baseline and percent changes from baseline. The original 36 items will not be listed.

Population: ITT.

## 9. DEFINITIONS OF PHARMACOKINETICS TABLES, LISTINGS AND FIGURES

These analyses will be described in a separate SAP. A separate report will be written based on the analysis performed by [REDACTED]

## 10. DEFINITIONS OF PHARMACODYNAMICS TABLES, LISTINGS AND FIGURES

Mock tables, listings and figures are provided in a separate document.

All analyses will be run on the ITT population, unless specified otherwise.

### 10.1. PARAMETERS

Assessment: marker YKL-40 (also known as Chitinase 3 L 1, CHI3L1) in serum.

Visits to present:

- Baseline
- Week 12: this is either the Week 12 visit itself, OR the EDV visit in case the subject dropped out of the study prior to Week 12, OR the last post-baseline visit prior to Week 12.

The actual values, changes from baseline and percent changes from baseline will be derived and analyzed.

Note that the percentage change from baseline remains missing if the baseline value itself equals zero.

### 10.2. TABLES

Tables will present Baseline and Week 12, unless specified otherwise. Tables will present the 7 treatment groups.

Analysis population: ITT.

#### **Table 14.2.3.1: Pharmacodynamics: YKL-40: Descriptive statistics**

Descriptive statistics of actual values, changes from baseline and percent changes from baseline.

#### **Table 14.2.3.2: Pharmacodynamics: YKL-40: Summary with overall p-value and pairwise comparisons versus placebo**

Analysis model is similar to DAS28(CRP). Only one time point: Week 12.

#### **Table 14.2.3.3: Pharmacodynamics: YKL-40: Exploratory pairwise comparisons among the dose groups within each regimen**

**Table 14.2.3.4: Pharmacodynamics: YKL-40: Exploratory pairwise comparisons between regimens within each daily dose level**

**Table 14.2.3.5: Pharmacodynamics: YKL-40: Descriptive statistics for Week 12 ACR20 responders versus non-responders**

Same as previous table, but split up for Week 12 ACR20 responders and non-responders.

**Table 14.2.3.6: Pharmacodynamics: YKL-40: Descriptive statistics for Week 12 ACR50 responders versus non-responders**

**Table 14.2.3.7: Pharmacodynamics: YKL-40: Descriptive statistics for Week 12 ACR70 responders versus non-responders**

**Table 14.2.3.8: Pharmacodynamics: YKL-40: Correlation with ACR20/50/70**

Relationship between the change from baseline in PD marker and ACR20 will be investigated via a logistic regression model on the ACR20 responses, with the following factors in the model: treatment groups, baseline YKL-40, change from baseline in YKL-40. ACR20 and YKL-40 will be taken from the same visit.

Same for the ACR50 and ACR70.

**Table 14.2.3.9: Pharmacodynamics: YKL-40: Correlation with continuous efficacy parameters**

Relationship between the change from baseline in PD marker and the change from baseline in efficacy parameters will be investigated via Pearson's correlation coefficient as provided by SAS PROC CORR. Efficacy parameters and YKL-40 will all be taken from the same visit.

The following efficacy parameters will be tested:

- ACR20 components: TJC68, SJC66, the 3 VAS scores, CRP and HAQ-DI.
- Summary scores: DAS28(CRP), CDAI and SDAI.

### 10.3. FIGURES

**Figure 14.2.3.1: Pharmacodynamics: YKL-40: Mean plot over time**

Line plot baseline-W12, like for the efficacy parameters. Plot will contain the 7 groups.

Analysis population: ITT.

**Figure 14.2.3.2: Pharmacodynamics: YKL-40: Correlation plots versus efficacy**

Correlation scatterplots of the percent changes from baseline in YKL-40 versus:

- Percent change from baseline in: TJC68, SJC66, the 3 VAS scores, CRP and HAQ-DI
- Percent change from baseline in DAS28(CRP), CDAI and SDAI

Efficacy parameters and YKL-40 will all be taken from the same visit. Each plot on a new page, using different plot symbols per group. Because of the large number of groups, put the placebo with the qd dose groups on one plot, and the placebo with the bid dose groups on a second plot.



Analysis population: ITT.

## 10.4. LISTINGS

### Listing 16.2.8: Pharmacodynamics: YKL-40

Listing per treatment group, per subject and per time point of the YKL-40 results: actual values, changes from baseline and percent changes from baseline.

Population: ITT.

## 11. DEFINITIONS OF PK/PD TABLES, LISTINGS AND FIGURES

These analyses will be described in a separate SAP. A separate report will be written based on the analysis performed by [REDACTED].

## 12. DEFINITIONS OF SAFETY TABLES, LISTINGS AND FIGURES

Mock tables, listings and figures are provided in a separate document.

All analyses will be run on the Safety population, unless specified otherwise.

### 12.1. ADVERSE EVENTS

#### 12.1.1. Treatment-Emergent Principle

All adverse events with onset on or after first dosing are considered treatment-emergent adverse events (TEAE).

Adverse events will be placed into analysis periods according to their start date. Analysis periods: see section 3.2. The AE will only be presented in the analysis period during which it started. Rule: period start date  $\leq$  AE start date  $\leq$  period stop date.

In case the AE start date is incomplete, a worst-case allocation will be done according to the available parts of the AE start date. When too much of the AE start date is missing to apply the above rules, the AE will be allocated to:

- the first treatment period, in case the doubt is between the screening and first treatment analysis period;
- both treatment periods, in case the doubt is between both treatment periods.

All adverse events emerging during the screening period will only be listed, not presented in any of the tables. These events are not TEAEs. All tables will present TEAEs only.

### 12.1.2. Treatment Relatedness

Following (ICH-E3), the investigator's assessment of drug relatedness will be dichotomized as follows:

Drug related: at least possibly drug related, OR with missing drug relatedness (= worst-case)

Not drug related: less than possibly drug related.

In tabulations this dichotomized parameter will be used, but in the listings the original parameter will be presented. Note that there is only the drug-relatedness assessment by the investigator.

### 12.1.3. Worst-Case Principle

When cross-tabulating AE preferred terms versus an AE attribute (e.g., intensity), the worst-case is always applied within each analysis period. I.e., when a subject has two times the same AE preferred term in the same analysis period, then the subject is reported only once: only with the worst intensity. If this happens in two different analysis periods, the AE is reported twice: once in each analysis period.

### 12.1.4. Adverse Event Onset Day and Duration

#### AE onset day in the study

= (AE start date) – (date of first study drug administration) + 1, when the AE start date is completely known;

= missing when the AE start date is incomplete or unknown.

#### AE onset day in the period

= (AE start date) – (start date of the period into which the AE was allocated) + 1, when the AE start date is completely known;

= missing when the AE start date is incomplete or unknown.

#### AE duration

= (AE stop date) – (AE start date) + 1, when both dates are completely known;

= (trial termination date) – (AE start date) + 1, when the AE start date is fully known but the AE is not resolved at the end of the study; in this case the duration will be presented as ">x days" in the listing to identify it as a censored result;

= missing when the AE start date is incomplete or unknown, or when the AE has resolved but with an incomplete or unknown end date.

### 12.1.5. Calculation of Percentages

All percentages will be calculated against the total number of subjects that are still in the study in that particular analysis period.

### 12.1.6. Tables up to Week 12

No formal inferential statistics (p-values) will be derived.

An as-treated analysis will be performed. All tabulations and listings will present treatments rather than analysis periods.

#### **Table 14.3.1.1: Treatment-emergent adverse events up to Week 12: Summary table**

Tabulation per treatment of the number and percentage of subjects with the following:

- Subjects with at least one treatment-emergent adverse event (TEAE)
- Subjects with at least one serious TEAE
- Subjects who died (treatment-emergent)
- Subjects with at least one mild TEAE as worst intensity
- Subjects with at least one moderate TEAE as worst intensity
- Subjects with at least one severe TEAE
- Subjects with at least one TEAE that was considered treatment-related
- Subjects with at least one treatment-emergent SAE that was considered treatment-related
- Subjects with at least one TEAE for which the study treatment was temporarily stopped
- Subjects with at least one TEAE for which the study treatment was permanently stopped
- Subjects with at least one TEAE flagged as an individual stopping criterion by the investigator.

Population: safety.

Table to be created for the Week 12 interim analysis as well.

#### **Table 14.3.1.2: Treatment-emergent adverse events up to Week 12: Summary table by region**

Same table as the preceding one, but by geographical region.

Population: safety.

#### **Table 14.3.1.3: Treatment-emergent adverse events up to Week 12: Tabulation of all adverse events**

Tabulation of TEAE preferred terms per body class, per treatment group.

Population: safety.

Table to be created for the Week 12 interim analysis as well.

**Table 14.3.1.4: Treatment-emergent adverse events up to Week 12: Tabulation of all adverse events by region**

Same as previous table, but split up by geographical region.

Population: safety.

**Table 14.3.1.5: Treatment-emergent adverse events up to Week 12: Tabulation of all adverse events by gender**

Same as previous table, but split up for males and females.

Population: safety.

**Table 14.3.1.6: Treatment-emergent adverse events up to Week 12: Tabulation of all adverse events by age category**

Same as previous table, but split up for age categories:

< 45 ;

[45,65[ ;

[65,75[ ;

≥ 75 years.

Population: safety.

**Table 14.3.1.7: Treatment-emergent adverse events up to Week 12: Tabulation per intensity**

Cross-tabulation of TEAE preferred terms versus their intensity. Use the worst-case intensity per TEAE per subject. Per treatment group.

Population: safety.

Table to be created for the Week 12 interim analysis as well.

**Table 14.3.1.8: Treatment-emergent adverse events up to Week 12: Tabulation of all treatment-related events**

Tabulation of TEAE preferred terms per body class, per treatment group. Selecting only the TEAEs that were treatment-related (see section 12.1.2).

Population: safety.

Table to be created for the Week 12 interim analysis as well.

**Table 14.3.1.9: Treatment-emergent adverse events up to Week 12: Tabulation of all treatment-related events by gender****Table 14.3.1.10: Treatment-emergent adverse events up to Week 12: Tabulation of all treatment-related events by age category**

**Table 14.3.1.11: Treatment-emergent adverse events up to Week 12: Tabulation of the intensity of treatment-related events**

Cross-tabulation of TEAE preferred terms versus their intensity; per body class and per treatment group. Use the worst-case intensity per TEAE per subject, selecting only the TEAEs that were treatment-related (see section 12.1.2).

Population: safety.

Table to be created for the Week 12 interim analysis as well.

**Table 14.3.1.12: Treatment-emergent adverse events up to Week 12: Tabulation of the drug-relatedness of serious events**

Cross-tabulation of TEAE preferred terms versus their drug-relatedness; per body class and per treatment group. Selecting only the treatment-emergent serious adverse events.

Population: safety.

Table to be created for the Week 12 interim analysis as well.

**Table 14.3.1.13: Treatment-emergent adverse events up to Week 12: Tabulation of the drug-relatedness of events leading to discontinuation**

Cross-tabulation of TEAE preferred terms versus their drug-relatedness; per body class and per treatment group. Selecting only the TEAEs for which the study treatment was permanently discontinued (AE page), or for which the study was discontinued (trial termination page).

Population: safety.

Table to be created for the Week 12 interim analysis as well.

**Table 14.3.1.14: Treatment-emergent adverse events up to Week 12: Tabulation of the drug-relatedness of events leading to a temporary stop**

Cross-tabulation of TEAE preferred terms versus their drug-relatedness; per body class and per treatment group. Selecting only the TEAEs for which the study treatment was temporarily stopped (AE page).

Population: safety.

**Table 14.3.1.15: Treatment-emergent adverse events up to Week 12: Tabulation of the drug-relatedness of individual stopping rules**

Cross-tabulation of TEAE preferred terms versus their drug-relatedness; per body class and per treatment group. Selecting only the TEAEs that were flagged as “stopping criterion” in the eCRF by the investigator.

Population: safety.

### 12.1.7. Tables over the whole 24 weeks period

The same tables will be created for the full 24 weeks period of the study, using the 11 groups as specified in section 2.4.

All tables will be created for 3 analysis periods: “first 12 weeks”, “second 12 weeks”, and “full 24 weeks” as defined in section 3.2.

Population: Safety.

#### **Table 14.3.1.16: Treatment-emergent adverse events over the whole study: Summary table**

Tabulation per treatment of the number and percentage of subjects with the following:

- Subjects with at least one treatment-emergent adverse event (TEAE)
- Subjects with at least one serious TEAE
- Subjects who died (treatment-emergent)
- Subjects with at least one mild TEAE as worst intensity
- Subjects with at least one moderate TEAE as worst intensity
- Subjects with at least one severe TEAE
- Subjects with at least one TEAE that was considered treatment-related
- Subjects with at least one treatment-emergent SAE that was considered treatment-related
- Subjects with at least one TEAE for which the study treatment was temporarily stopped
- Subjects with at least one TEAE for which the study treatment was permanently stopped
- Subjects with at least one TEAE flagged as an individual stopping criterion by the investigator
- Subjects with at least one TEAE that has not resolved.

Population: safety.

#### **Table 14.3.1.17: Treatment-emergent adverse events over the whole study: Summary table by region**

(only for the “full 24 weeks”, not for the two periods separately)

#### **Table 14.3.1.18: Treatment-emergent adverse events over the whole study: Tabulation of all adverse events**

#### **Table 14.3.1.19: Treatment-emergent adverse events over the whole study: Tabulation of all adverse events by region**

(only for the “full 24 weeks”, not for the two periods separately)

#### **Table 14.3.1.20: Treatment-emergent adverse events over the whole study: Tabulation of all adverse events by gender**

(only for the “full 24 weeks”, not for the two periods separately)

**Table 14.3.1.21: Treatment-emergent adverse events over the whole study: Tabulation of all adverse events by age category**

(only for the “full 24 weeks”, not for the two periods separately)

**Table 14.3.1.22: Treatment-emergent adverse events over the whole study: Tabulation per intensity****Table 14.3.1.23: Treatment-emergent adverse events over the whole study: Tabulation of all treatment-related events****Table 14.3.1.24: Treatment-emergent adverse events over the whole study: Tabulation of all treatment-related events by gender**

(only for the “full 24 weeks”, not for the two periods separately)

**Table 14.3.1.25: Treatment-emergent adverse events over the whole study: Tabulation of all treatment-related events by age category**

(only for the “full 24 weeks”, not for the two periods separately)

**Table 14.3.1.26: Treatment-emergent adverse events over the whole study: Tabulation of the intensity of treatment-related events****Table 14.3.1.27: Treatment-emergent adverse events over the whole study: Tabulation of the drug-relatedness of serious events****Table 14.3.1.28: Treatment-emergent adverse events over the whole study: Tabulation of the drug-relatedness of events leading to discontinuation****Table 14.3.1.29: Treatment-emergent adverse events over the whole study: Tabulation of the drug-relatedness of events leading to a temporary stop****Table 14.3.1.30: Treatment-emergent adverse events over the whole study: Tabulation of the drug-relatedness of individual stopping rules****12.1.8. Listings****Listing 16.2.10.1: Treatment-emergent adverse events: Summary listing**

Listing per treatment group, per subject and per analysis period (excluding screening) of the following:

- Period start and end date
- AE preferred term (flagging serious TEAEs with an asterisk \*)
- AE start and end date
- AE onset day in study
- AE onset day in period
- AE duration
- AE intensity
- AE drug relatedness
- AE outcome
- AE action taken
- Concomitant therapy started (yes/no)
- Individual stopping criterion (yes/no)

In such a way that all information fits on one line for each AE.

Periods without an AE will be included as “no AE” records.

Population: safety.

#### **Listing 16.2.10.2: Pre-treatment adverse events: Summary listing**

Listing per treatment group and per subject of the screening analysis period of all AE details, like in the previous listing. Subjects without an AE can be omitted from this listing.

Population: safety.

#### **Listing 16.2.10.3: Adverse events: Full listing of the serious adverse events**

Listing per treatment group, per subject and per analysis period (excluding screening) of all AE details, including MedDRA coding. Only actual serious adverse events will be presented here (irrespective their treatment-emergence, so also showing any SAEs during the screening analysis period).

Population: safety.

#### **Listing 16.2.10.4: Treatment-emergent adverse events: Full listing of the events leading to discontinuation**

Same as the previous listing, but only selecting TEAEs that lead to a permanent stop of study medication, or of the study itself. This is including the “stopping rules”.

Population: safety.

#### **Listing 16.2.10.5: Treatment-emergent adverse events: Full listing of the events leading to a temporary stop**

Same as the previous listing, but only selecting TEAEs that lead to a temporary stop of study medication.

Population: safety.



**Listing 16.2.10.6: Adverse events: Coding information**

Listing of all available coding steps between AE verbatim and AE system organ class, mentioning also the subjects who had this AE.

Population: safety.

**Listing 16.2.10.7: Adverse events: Listing of subjects for whom a narrative is required**

Listing and flagging subjects who:

- Died
- Had an SAE
- Had an AE leading to study discontinuation
- Had an AE leading to a permanent stop of study drug
- Had an AE flagged as stopping criterion by the investigator in the eCRF.

Population: all screened subjects.

The clinical study report will add subject case narratives in the body of the report and/or in section 14.3.1. No standard programming is planned for this.

**12.2. LABORATORY SAFETY****12.2.1. Laboratory Units**

The statistical analysis will only present results in Standard International (SI) units. Other units will not be presented.

Lab tests with only a very low sample size ( $N < 3$  in the study total column) will not be presented in the tables, but only in the listings. Lab tests that are not part of the planned test panels according to the protocol will only be listed.

Endocrine tests will only be presented for male subjects. If any results would be available for females (not planned in the protocol), then these will be left out of the tables but still presented in the listings.

Urinalysis tests will be presented as part of the descriptive statistics and/or shift tables.

The serum CRP will not be presented in this section (is considered an efficacy parameter).

**12.2.2. Derived lab tests**

The following two “tests” will be derived from the data provided by the central lab: the ratios TC/HDL and LDL/HDL, where TC=total cholesterol, and the ratio between the CD4 and CD8 absolute counts (CD4/CD8).

Note that if the denominator is zero, the ratios will not be calculated.

The following “normal range” will be applied:

Lab test label in analysis	Normal range, males	Normal range, females
Derived TC/HDL ratio	0.0 - 5.0	0.0 - 4.5
Derived LDL/HDL ratio	0.0 - 3.5	0.0 - 3.0

No normal range is defined for the CD4/CD8 ratio.

These derived “tests” will be analyzed together with the tests provided by the lab.

### 12.2.3. Number of Significant Digits

The original results and normal limits will be rounded to present only a relevant number of digits. The Mock TLFs contains a table on the expected number of significant digits per lab test. This rounding will be done prior to any parameter derivation. Note that this table also contains the classification of lab tests into categories.

### 12.2.4. Baseline and Change from Baseline

The baseline is defined as the last sample prior to first dosing, so should normally be from the Day -1 visit. Baseline will be determined per lab test individually. It is recognized that baseline tests may thus come from more than one lab sample and not just from the “baseline visit” sample.

The change from baseline will be calculated for all post-baseline time points as:

Change from baseline at time point t = value at time point t – baseline value.

The percentage change from baseline will be calculated for all post-baseline time points as:

Percent change from baseline at time point t = 100 x (change from baseline at time point t / baseline value).

Note that the percent change from baseline will be left missing in case the baseline value is zero.

### 12.2.5. Scoring of Laboratory Values

#### 12.2.5.1. Scoring according to normal ranges

All values will be compared to their matching normal ranges. The normal ranges provided by the laboratory will be used for this, as available in the database. Values will be scored as abnormally low (L), normal (N) or abnormally high (H).

A value is classified as abnormally low (L) when the value < lower limit of the normal range. A value is classified as abnormally high (H) when the value > upper limit of the normal range. Values equal to the boundaries are still considered normal (N).

An original value like “<X” where X equals the lower limit of the normal range will be classified as abnormally low (L). An original value like “>X” where X equals the upper limit of the normal range will be classified as abnormally high (H).

A normal range of “<10” will be interpreted as: lower limit = 0; upper limit = 10.

A normal range of “>10” will be interpreted as: lower limit = 10; upper limit = +INF.

Tests without normal ranges will not be scored.

Any already existing low/normal/high flags in the raw database will be ignored. The newly derived flag will be used for the analysis TLFs. Any clinical significance flags will be used in the listings.

### 12.2.5.2. CTCAE V3.0 gradings for biochemistry tests

CTCAE version 3.0 classification is only for the below tests. For other tests, only the classification according to normal ranges will be done.

Lab test	Grade 0	Grade 1	Grade 2	Grade 3	Grade 4
Albumin (g/L)	≥ LLN	[30,LLN[	[20,30[	<20	(NAP)
ALT, AST, alkaline phosphatase and γGT	≤ ULN	]1.0,2.5] x ULN	]2.5,5.0] x ULN	]5.0,20.0] x ULN	> 20.0 x ULN
Amylase and lipase	≤ ULN	]1.0,1.5] x ULN	]1.5,2.0] x ULN	]2.0,5.0] x ULN	> 5.0 x ULN
Total bilirubin	≤ ULN	]1.0,1.5] x ULN	]1.5,3.0] x ULN	]3.0,10.0] x ULN	> 10.0 x ULN
Calcium low (mmol/L)	≥ LLN	[2.00,LLN[	[1.75,2.00[	[1.50,1.75[	<1.50
Calcium high (mmol/L)	≤ ULN	]ULN,2.9]	]2.9,3.1]	]3.1,3.4]	>3.4
Total cholesterol (mmol/L)	≤ ULN	]ULN,7.75]	]7.75,10.34]	]10.34,12.92]	>12.92
Creatinine	≤ ULN	]1.0,1.5] x ULN	]1.5,3.0] x ULN	]3.0,6.0] x ULN	> 6.0 x ULN
Glucose low (mmol/L)	≥ LLN	[3.0,LLN[	[2.2,3.0[	[1.7,2.2[	<1.7
Glucose high (mmol/L)	≤ ULN	]ULN,8.9]	]8.9,13.9]	]13.9,27.8]	>27.8
Phosphate (mmol/L)	≥ LLN	[0.8,LLN[	[0.6,0.8[	[0.3,0.6[	<0.3
Potassium low (mmol/L)	≥ LLN	[3.0,LLN[	(NAP)	[2.5,3.0[	<2.5
Potassium high (mmol/L)	≤ ULN	]ULN,5.5]	]5.5,6.0]	]6.0,7.0]	>7.0
Sodium low (mmol/L)	≥ LLN	[130,LLN[	(NAP)	[120,130[	<120
Sodium high (mmol/L)	≤ ULN	]ULN,150]	]150,155]	]155,160]	>160
Triglycerides	≤ ULN	]1.0,2.5] x ULN	]2.5,5.0] x ULN	]5.0,10.0] x ULN	> 10.0 x ULN

Note:

– LLN = lower normal limit

- ULN = upper normal limit
- (NAP) = grade does not exist.

### 12.2.5.3. CTCAE v3.0 gradings for hematology tests

CTCAE version 3.0 classification is only for the below tests. For other tests, only the classification according to normal ranges will be done.

Lab test	Grade 0	Grade 1	Grade 2	Grade 3	Grade 4
Hemoglobin (g/L)	≥ LLN	[100,LLN[	[80,100[	[65,80[	< 65
WBC (giga/L)	≥ LLN	[3.0,LLN[	[2.0,3.0[	[1.0,2.0[	<1.0
Lymphocytes (giga/L)	≥ LLN	[0.8,LLN[	[0.5,0.8[	[0.2,0.5[	< 0.2
Neutrophils (giga/L)	≥ LLN	[1.5,LLN[	[1.0,1.5[	[0.5,1.0[	< 0.5
Platelet count (giga/L)	≥ LLN	[75,LLN[	[50,75[	[25,50[	< 25
INR and partial thromboplastin time	≤ ULN	]1.0,1.5] x ULN	]1.5,2.0] x ULN	>2.0 x ULN	(NAP)

Note:

- LLN = lower normal limit
- ULN = upper normal limit
- (NAP) = grade does not exist.

### 12.2.5.4. ATP III classifications

Lab test	Categories				
LDL (mg/dL)	< 100	[100,130[	[130,160[	[160,190[	≥ 190
HDL (mg/dL)	< 40	[40,60[	≥ 60		
Total cholesterol (mg/dL)	< 200	[200,240[	≥ 240		
Triglycerides (mg/dL)	< 150	[150,200[	[200,500[	≥ 500	

### 12.2.6. Handling of Urinalysis Results

Urinalysis tests for which a normal range is available will also be categorized as L, N, or H. If there is doubt on whether the result is within normal range or not, the worst-case will be taken. E.g., when the test result equals “4 to 6” and the normal range is 0-5, then the test result will be classified as abnormally high (H) if it is a post-baseline visit, and will be classified as normal (N) if it is a screening or baseline visit. The original urinalysis test result will not be used for the tables, but will only be listed.

Urinalysis tests for which no normal range is available will be listed only.

### 12.2.7. Last on-treatment result

Derived per parameter separately, for values as well as abnormalities.

Three last on-treatment determinations will be done, depending on the chapter of analysis:

1. Including all nonmissing post-baseline values, including “use=No” and unscheduled measurements, up to and including the Week 12 value.
2. Including all nonmissing post-Week 12 values, including “use=No” and unscheduled measurements, up to and including the Week 24 value.
3. Including all nonmissing post-baseline values, including “use=No” and unscheduled measurements, up to and including the Week 24 value.

Early discontinuation visits are considered as on-treatment visits as well.

### **12.2.8. Worst-Case Abnormality**

Derived per parameter separately, for abnormalities only.

Three worst-case determinations will be done, depending on the chapter of analysis:

1. Including all nonmissing post-baseline values, including “use=No” and unscheduled measurements, up to and including the Week 12 value. Also including the follow-up value in case the subject did not participate in the second period of the study. If the subject participated in the second period of the study, then the follow-up value will be included in the worst-case of the second period of the study, not of the first period.
2. Including all nonmissing post-Week 12 values, including “use=No” and unscheduled measurements, up to and including the follow-up value.
3. Including all nonmissing post-baseline values, including “use=No” and unscheduled measurements, up to and including the follow-up value.

#### **12.2.8.1. According to normal ranges**

Possible categories:

- H = abnormally high: at least one postdose measurement is above the normal range, and there are no values below the normal range.
- L = abnormally low: at least one postdose measurement is below the normal range, and there are no values above the normal range.
- H+L = abnormally high and low: at least one postdose measurement is above the normal range, and at least one other postdose measurement is below the normal range.
- N = normal: all postdose measurements are within normal limits.

#### **12.2.8.2. According to CTCAE tox grades**

The worst-case is the highest post-baseline tox grade.

#### **12.2.8.3. According to the ATP III classifications**

The worst-case is the highest post-baseline class for LDL, total cholesterol and triglycerides.  
The worst-case is the lowest post-baseline class for HDL.

## 12.2.9. Tables up to Week 12

No formal inferential statistics (p-values) will be derived. Time points to present: screening, baseline, Weeks 1-2-4-8-12, last on-treatment, worst case.

If the subject did not participate in the second period of the study, then the follow-up visit will be presented in the table of the first period of the study. If the subject participated in the second period of the study, then the follow-up value will be presented in the table of the second period of the study, not of the first period.

### **Table 14.3.2.1: Laboratory data up to Week 12: Descriptive statistics of the actual values per time point**

Descriptive statistics of actual values per lab test category (hematology, biochemistry, endocrine, urinalysis), lab test and unit, treatment group and time point.

Population: safety.

Population: safety, selecting subjects with a baseline value within normal range (for endocrine tests only).

Table to be created for the Week 12 interim analysis as well.

### **Table 14.3.2.2: Laboratory data up to Week 12: Descriptive statistics of the changes from baseline per time point**

Descriptive statistics of changes from baseline per lab test category (hematology, biochemistry, endocrine, urinalysis), lab test and unit, treatment group and time point. The actual value at baseline will also be presented in this table.

Population: safety.

Population: safety, selecting subjects with a baseline value within normal range (for endocrine tests only).

Table to be created for the Week 12 interim analysis as well.

### **Table 14.3.2.3: Laboratory data up to Week 12: Descriptive statistics of the percent changes from baseline per time point**

Descriptive statistics of percent changes from baseline per lab test category (hematology, biochemistry, endocrine, urinalysis), lab test and unit, treatment group and time point. The actual value at baseline will also be presented in this table.

Population: safety.

Population: safety, selecting subjects with a baseline value within normal range (for endocrine tests only).

Table to be created for the Week 12 interim analysis as well.

**Table 14.3.2.4: Laboratory data up to Week 12: Shift table of the lab abnormalities per time point**

Shift table per lab test category (hematology, biochemistry, endocrine, urinalysis), lab test, treatment group and time point. The table will present the shift in abnormality (L/N/...) at worst-case and last on-treatment (other time points: not needed) versus the baseline abnormality (L/N/H). Tests without normal ranges will not be presented in this table.

Population: safety.

**Table 14.3.2.5: Laboratory data up to Week 12: Treatment-emergent lab abnormalities per time point**

Frequency table of the treatment-emergent lab abnormalities per lab test category (hematology, biochemistry, endocrine, urinalysis), lab test, treatment group and time point (including the worst-case). A post-baseline abnormality is considered treatment-emergent if it differs from the baseline result. Tests without normal ranges will not be presented in this table.

Population: safety.

Table to be created for the Week 12 interim analysis as well.

**Table 14.3.2.6: Laboratory data up to Week 12: Frequency table of the CTCAE toxicity grades per time point**

Frequency table (with cumulative percentages) per lab test category (hematology, biochemistry, endocrine, urinalysis), lab test, treatment group and time point.

Table for selected tests only.

Population: safety.

Table to be created for the Week 12 interim analysis as well.

**Table 14.3.2.7: Laboratory data up to Week 12: Shift table of the CTCAE toxicity grades per time point**

Shift table per lab test category (hematology, biochemistry, endocrine, urinalysis), lab test, treatment group and time point. The table will present the shift in CTCAE tox grade (0/1/2/3/4) at worst-case and last on-treatment (other time points: not needed) versus the baseline grade (0/1/2/3/4).

Table for selected tests only.

Population: safety.

**Table 14.3.2.8: Laboratory data up to Week 12: Treatment-emergent CTCAE toxicity grades per time point**

Frequency table of the treatment-emergent lab CTCAE tox grades per lab test category (hematology, biochemistry), lab test, treatment group and time point (including worst-case). A

post-baseline grade is considered treatment-emergent if it is higher than the baseline tox grade.

Table for selected tests only.

Population: safety.

Table to be created for the Week 12 interim analysis as well.

**Table 14.3.2.9: Laboratory data up to Week 12: Frequency table of the ATP III classes per time point**

Frequency table (with cumulative percentages) per lab test category (hematology, biochemistry, endocrine, urinalysis), lab test, treatment group and time point.

Table for selected tests only.

Population: safety.

Table to be created for the Week 12 interim analysis as well.

**Table 14.3.2.10: Laboratory data up to Week 12: Shift table of the ATP III classes per time point**

Shift table per lab test category (hematology, biochemistry, endocrine, urinalysis), lab test, treatment group and time point. The table will present the shift in ATP III class at worst-case and last on-treatment (other time points: not needed) versus the baseline class.

Table for selected tests only.

Population: safety.

**Table 14.3.2.11: Laboratory data up to Week 12: Treatment-emergent ATP III classes per time point**

Frequency table of the treatment-emergent lab ATP III classes per lab test category (hematology, biochemistry), lab test, treatment group and time point (including worst-case). A post-baseline class is considered treatment-emergent if it differs from the baseline result and is of higher category.

Table for selected tests only.

Population: safety.

Table to be created for the Week 12 interim analysis as well.

**Table 14.3.2.12: Laboratory data up to Week 12: Frequency table of percent changes from baseline per time point in inhibin B and testosterone**

Frequency table of the percent changes from baseline per lab test category (endocrine), lab test, treatment group and time point. The percent changes from baseline will be categorized



as: <-50%, [-50%,+50%], >+50%. A worst-case will also be shown selected as the lowest post-baseline value.

Table only for inhibin B, free testosterone, total testosterone.

Population: safety – males.

Population: safety, selecting male subjects with a baseline value within normal range.

Table to be created for the Week 12 interim analysis as well.

**Table 14.3.2.13: Laboratory data up to Week 12: Shift table of the categorical lab data per time point**

Shift table per lab test category (hematology, biochemistry, endocrine, urinalysis), lab test, treatment group and time point. Selecting only the tests with categorical data. The table will present the shift in value at each post-baseline time point versus the baseline value.

Population: safety.

**12.2.10. Figures up to Week 12**

**Figure 14.3.2.1: Laboratory data up to Week 12: Mean plots of the actual values over time**

Mean plots of the actual values over time, with each lab test on a new page but with all treatment groups on the same plot using different plot symbols. Unscheduled or follow-up results will not be part of this plot.

Population: safety.

Figure to be created for the Week 12 interim analysis as well.

**Figure 14.3.2.2: Laboratory data up to Week 12: Mean plots of the changes from baseline over time**

Mean plots of the changes from baseline over time, with each lab test on a new page but with all treatment groups on the same plot using different plot symbols. Unscheduled or follow-up results will not be part of this plot. With a horizontal reference line at zero, indicating no change. Plots will start with a zero mean at baseline.

Population: safety.

Figure to be created for the Week 12 interim analysis as well.

**Figure 14.3.2.3: Laboratory data up to Week 12: Mean plots of the percent changes from baseline over time**

Mean plots of the percent changes from baseline over time, with each lab test on a new page but with all treatment groups on the same plot using different plot symbols. Unscheduled or follow-up results will not be part of this plot. With a horizontal reference line at zero, indicating no change. Plots will start with a zero mean at baseline.

Population: safety.

Figure to be created for the Week 12 interim analysis as well.

### **12.2.11. Tables over the whole 24 weeks period**

The same tables will be made as for the first period of the study, but with the 11 adapted treatment groups. Time points to present: screening, baseline, Week 1-2-4-8-12, last on-treatment during period 1, worst case during period 1, Weeks 16-20-24, last on-treatment during period 2, follow-up (all follow-up visits in the study), worst case during period 2, last on-treatment during the whole study, worst case during the whole study. Population: Safety.

**Table 14.3.2.14: Laboratory data over the whole study: Descriptive statistics of the actual values per time point**

**Table 14.3.2.15: Laboratory data over the whole study: Descriptive statistics of the changes from baseline per time point**

**Table 14.3.2.16: Laboratory data over the whole study: Descriptive statistics of the percent changes from baseline per time point**

**Table 14.3.2.17: Laboratory data over the whole study: Shift table of the lab abnormalities per time point**

**Table 14.3.2.18: Laboratory data over the whole study: Treatment-emergent lab abnormalities per time point**

**Table 14.3.2.19: Laboratory data over the whole study: Frequency table of the CTCAE toxicity grades per time point**

**Table 14.3.2.20: Laboratory data over the whole study: Shift table of the CTCAE toxicity grades per time point**

**Table 14.3.2.21: Laboratory data over the whole study: Treatment-emergent CTCAE toxicity grades per time point**

**Table 14.3.2.22: Laboratory data over the whole study: Frequency table of the ATP III classes per time point**

**Table 14.3.2.23: Laboratory data over the whole study: Shift table of the ATP III classes per time point**

**Table 14.3.2.24: Laboratory data over the whole study: Treatment-emergent ATP III classes per time point**

**Table 14.3.2.25: Laboratory data over the whole study: Frequency table of percent changes from baseline per time point in inhibin B and testosterone**

**Table 14.3.2.26: Laboratory data over the whole study: Shift table of the categorical lab data per time point**

### 12.2.12. Figures over the whole 24 weeks period

The follow-up visit will not be shown on the mean plots because it is expected that it will happen only for a limited number of subjects, as a large number will roll over into study 205 without a follow-up visit.

The same plots again, but with the 11 adapted treatment groups. Plots will show data from screening up to the end of the study.

**Figure 14.3.2.4: Laboratory data over the whole study: Subject profile plots over time – subjects with a treatment-emergent abnormality**

**Figure 14.3.2.5: Laboratory data over the whole study: Mean plots of the actual values over time**

**Figure 14.3.2.6: Laboratory data over the whole study: Mean plots of the changes from baseline over time**

**Figure 14.3.2.7: Laboratory data over the whole study: Mean plots of the percent changes from baseline over time**

### 12.2.13. Listings

All listings will contain next to the actual data (raw and changes) the following parameters:

- a fasted Y/N flag
- an abnormality L/H flag
- CTCAE toxicity grade, ATP III classification
- the test's normal range
- a clinical relevance flag

#### **Listing 16.2.11.1: Laboratory data: Full listing**

Listing per treatment group, per subject and per time point of all recorded data.

Population: safety.

#### **Listing 16.2.11.2: Laboratory data: Abnormalities listing**

Listing per treatment group, per subject and per time point of all post-baseline data scored as out-of-normal-range, plus also the baseline reference time point.

Population: safety.

#### **Listing 16.2.11.3: Laboratory data: Clinically significant abnormalities**

Listing per treatment group, per subject and per time point of the eCRF's verbatim text field on clinically significant abnormalities.

Population: safety.

#### **Listing 16.2.11.4: Laboratory data: Urinalysis results**

Listing per treatment group, per subject and per time point of the abnormal urinalysis results only (including the baseline result as reference).

Population: safety.

#### **Listing 16.2.11.5: Laboratory data: Positive pregnancy tests**

Listing per treatment group, per subject and per time point of all pregnancy test results (urine and serum) for female subjects who had a positive test at some point.

Population: safety.

### **12.3. ECG**

#### **12.3.1. Available Data**

Available parameters: HR, QRS, QT, and an interpretation. The RR interval, when available, will only be listed.

#### **12.3.2. Calculated Parameters**

The QTc will always be derived during the statistical analysis, even when already available in the database. The value in the database will not be used in the analysis.

The QTc will be calculated using the following formula: Fridericia's cube-root corrected QT (Fridericia, 1920):

$$QTcF \text{ (ms)} = QT \text{ (ms)} \times \sqrt[3]{\frac{HR(bpm)}{60}}$$

The QTcF will be rounded up to the nearest integer prior to the derivation of changes and categorizations.

#### **12.3.3. Baseline and Change from Baseline**

The baseline is defined as the last nonmissing value prior to dosing. Baseline will be determined per ECG parameter individually. It is recognized that baseline parameters may thus come from more than one ECG reading and not just from the "baseline visit" reading.

The change from baseline will be calculated for all post-baseline time points as:

Change from baseline at time point t = value at time point t – baseline value.

#### **12.3.4. Visits for analysis**

An ECG is only taken at Screening (baseline), Week 12, Week 24/EDV and Follow-up. Therefore the analysis will not use any time windowing, but will just present the visits as they are: Baseline, Week 12 and Week 24, and Follow-up. See also mocks.

In case the subject prematurely discontinues the study during the first period of the study, the EDV ECG will be used as “Week 12”. In case the subject prematurely discontinues the study during the second period of the study, the EDV ECG will be used as “Week 24”.

Because of this, there will not be a “last on-treatment” nor a “worst-case”.

### 12.3.5. Normal Ranges

For the QT and QTcF parameters, the following categorizations will be done:

- of the actual values:
  - $\leq 450$  ms,
  - ]450,480] ms,
  - ]480,500] ms,
  - $>500$  ms;
- of the changes from baseline:
  - $\leq 30$  ms (including all decreases in QT),
  - ]30,60] ms,
  - $>60$  ms.

### 12.3.6. Analysis tables at Week 12

No formal inferential statistics (p-values) are planned in this SAP.

If the subject did not participate in the second period of the study, then the follow-up visit will be presented in the table of the first period of the study. If the subject participated in the second period of the study, then the follow-up value will be presented in the table of the second period of the study, not of the first period.

#### **Table 14.3.3.1: ECG up to Week 12: Descriptive statistics of the actual values per time point**

Descriptive statistics of the actual values per test and unit, and per treatment group.

Population: safety.

#### **Table 14.3.3.2: ECG at Week 12: Descriptive statistics of the changes from baseline**

Descriptive statistics of the changes from baseline per test and unit, and per treatment group. The actual value at baseline will also be presented in this table.

Population: safety.

#### **Table 14.3.3.3: ECG at Week 12: Shift table of the QT abnormalities**

Shift table per test, treatment group and time point. The table will present the shift in abnormality at each post-baseline time point (including the worst-case) versus the baseline abnormality. Parameters: QT and QTcF.

Population: safety.

**Table 14.3.3.4: ECG at Week 12: Treatment-emergent QT abnormalities**

Frequency table of the treatment-emergent abnormalities per parameter and treatment group. A post-baseline abnormality is considered treatment-emergent if it is higher than the baseline result. Parameters: QT and QTcF.

Population: safety.

**Table 14.3.3.5: ECG at Week 12: Treatment-emergent abnormal change in QT**

Frequency table per test and treatment group. The table will present the abnormality of the changes. Parameters: QT and QTcF.

Population: safety.

**Table 14.3.3.6: ECG up to Week 12: Frequency table of the eCRF ECG interpretation per time point**

Frequency table per treatment group of the ECG interpretation scores as recorded in the eCRF. Interpretation scores are:

- Normal ECG in sinus rhythm;
- Cardiac conduction disorders;
- Rate and rhythm disorders;
- Supraventricular arrhythmias;
- Ventricular arrhythmias.

Note that more than one of these items can be ticked.

Population: safety.

**12.3.7. Analysis tables for the whole 24 weeks period**

Repeat tables, population=safety, with the adapted treatment groups focusing on subjects who used GLPG0634 during both study periods:

**Table 14.3.3.7: ECG over the whole study: Descriptive statistics of the actual values per time point****Table 14.3.3.8: ECG over the whole study: Descriptive statistics of the changes from baseline****Table 14.3.3.9: ECG over the whole study: Shift table of the QT abnormalities****Table 14.3.3.10: ECG over the whole study: Treatment-emergent QT abnormalities****Table 14.3.3.11: ECG over the whole study: Treatment-emergent abnormal change in QT****Table 14.3.3.12: ECG over the whole study: Frequency table of the eCRF ECG interpretation per time point**

### 12.3.8. Figures

No figures planned.

### 12.3.9. Listings

#### Listing 16.2.12.1: ECG: Full listing

Listing per treatment group, per subject and per time point of the ECG parameters (raw values as well as changes from baseline), flagging abnormal results. Includes the ECG interpretation and morphology findings.

Population: safety.

#### Listing 16.2.12.2: ECG: Abnormalities listing

Listing per treatment group, per subject and per time point of all post-baseline data scored as out-of-normal-range or clinically significant, plus also the baseline reference time point.

Population: safety.

## 12.4. VITAL SIGNS

### 12.4.1. Available Data

Available parameters: pulse rate, diastolic and systolic blood pressure, temperature.

### 12.4.2. Baseline and Change from Baseline

The baseline value will be the last nonmissing value prior to first dosing. Baseline will be determined per vital signs parameter individually. It is recognized that baseline parameters may thus come from more than one vital signs reading and not just from the “baseline visit” reading.

The change from baseline will be calculated for all post-baseline time points as:

Change from baseline at time point t = value at time point t – baseline value.

### 12.4.3. Normal Ranges

Parameter, unit	Normal range	
	Lower limit	Upper limit
Pulse, bpm	40	100
DBP, mmHg	45	90
SBP, mmHg	90	140
Temperature, °C	35.5	37.5

Values equal to the boundaries are still considered normal (N). A value is classified as abnormally low (L) when the value < lower limit of the normal range. A value is classified as abnormally high (H) when the value > upper limit of the normal range.

#### 12.4.4. Last on-treatment result

Derived per parameter separately, for values as well as abnormalities.

Three last on-treatment determinations will be done, depending on the chapter of analysis:

1. Including all nonmissing post-baseline values, including “use=No” and unscheduled measurements, up to and including the Week 12 value.
2. Including all nonmissing post-Week 12 values, including “use=No” and unscheduled measurements, up to and including the Week 24 value.
3. Including all nonmissing post-baseline values, including “use=No” and unscheduled measurements, up to and including the Week 24 value.

Early discontinuation visits are considered as on-treatment visits as well.

#### 12.4.5. Worst-Case Abnormality

Derived per parameter separately, for abnormalities only.

Three worst-case determinations will be done, depending on the chapter of analysis:

1. Including all nonmissing post-baseline values, including “use=No” and unscheduled measurements, up to and including the Week 12 value. Also including the follow-up value in case the subject did not participate in the second period of the study. If the subject participated in the second period of the study, then the follow-up value will be included in the worst-case of the second period of the study, not of the first period.
2. Including all nonmissing post-Week 12 values, including “use=No” and unscheduled measurements, up to and including the follow-up value.
3. Including all nonmissing post-baseline values, including “use=No” and unscheduled measurements, up to and including the follow-up value.

Possible categories:

- H = abnormally high: at least one postdose measurement is above the normal range, and there are no values below the normal range.
- L = abnormally low: at least one postdose measurement is below the normal range, and there are no values above the normal range.
- H+L = abnormally high and low: at least one postdose measurement is above the normal range, and at least one other postdose measurement is below the normal range.
- N = normal: all postdose measurements are within normal limits.

#### 12.4.6. Tables up to Week 12

No formal inferential statistics (p-values) will be derived. Time points to present: screening, baseline, Weeks 1-2-4-8-12, last on-treatment, worst case.

If the subject did not participate in the second period of the study, then the follow-up visit will be presented in the table of the first period of the study. If the subject participated in the



second period of the study, then the follow-up value will be presented in the table of the second period of the study, not of the first period.

**Table 14.3.4.1: Vital signs up to Week 12: Descriptive statistics of the actual values**

Descriptive statistics of the actual values per test and unit, treatment group and time point.

Population: safety.

**Table 14.3.4.2: Vital signs up to Week 12: Descriptive statistics of the changes from baseline**

Descriptive statistics of the changes from baseline per test and unit, treatment group and time point. The actual value at baseline will also be presented in this table.

Population: safety.

**Table 14.3.4.3: Vital signs up to Week 12: Summary table of means**

Summary table per test and unit, treatment group and time point, showing only the mean value of actual values and changes from baseline.

Population: safety.

Table to be created for the Week 12 interim analysis as well.

**Table 14.3.4.4: Vital signs up to Week 12: Shift table of the abnormalities per time point**

Shift table per test, treatment group and time point. The table will present the shift in abnormality (L/N/...) for “last” and “worst-case” versus the baseline abnormality (L/N/H).

Population: safety.

**Table 14.3.4.5: Vital signs up to Week 12: Treatment-emergent abnormalities per time point**

Frequency table of the treatment-emergent abnormalities per parameter, treatment group and time point (including the worst-case). A post-baseline abnormality is considered treatment-emergent if it is different from the baseline result.

Population: safety.

## **12.4.7. Tables over the whole 24 weeks period**

The same tables again, but with adapted treatment groups focusing on the subjects who used GLPG0634 during both study periods. Time points to present: screening, baseline, Week 1-2-4-8-12, last on-treatment during period 1, worst case during period 1, Weeks 16-20-24, last on-treatment during period 2, last on-treatment during the whole study, follow-up (all follow-up visits in the study), worst case during period 2, worst case during the whole study.

Population: Safety.

**Table 14.3.4.6: Vital signs over the whole study: Descriptive statistics of the actual values**

**Table 14.3.4.7: Vital signs over the whole study: Descriptive statistics of the changes from baseline**

**Table 14.3.4.8: Vital signs over the whole study: Summary table of means**

**Table 14.3.4.9: Vital signs over the whole study: Shift table of the abnormalities per time point**

**Table 14.3.4.10: Vital signs over the whole study: Treatment-emergent abnormalities per time point**

## 12.4.8. Figures

No figures are planned.

## 12.4.9. Listings

### Listing 16.2.13.1: Vital signs: Full listing

Listing per treatment, per subject and per time point of all parameters: raw values, changes from baseline, and flagging abnormal results.

Population: safety.

### Listing 16.2.13.2: Vital signs: Abnormalities listing

Listing per treatment group, per subject and per time point of all post-baseline data scored as out-of-normal-range or clinically significant, plus also the baseline reference time point.

Population: safety.

## 12.5. PHYSICAL EXAMINATIONS

### Listing 16.2.14: Physical examinations: Abnormal findings

Listing per treatment group, per subject and per time point of the selection of all abnormal findings (i.e., “abnormal” or “changed”).

Population: safety.

## 13. REFERENCES

Fridericia, L. (1920). Die systolendauer im elektrokardiogramm bei normalen menschen und bei herzkranken. *Acta Med Scand*, 53:469-486.

ICH-E3. (December 1995). *Structure and content of clinical study reports. Step 4 Guideline*.

ICH-E6. (17 July 1996). *Guideline for good clinical practice. Step 5 Guideline.*

ICH-E9. (5 February 1998). *Statistical principles for clinical trials. Step 4 guideline.*



Galapagos

---

## STATISTICAL ANALYSIS PLAN –ADDENDUM

---

**Project Number:** GLPG0634

**Study Number:** GLPG0634-CL-203 (DARWIN1)

**Study Title:** Randomized, double-blind, placebo-controlled, multicenter, phase IIb dose finding study of GLPG0634 administered for 24 weeks in combination with methotrexate to subjects with moderately to severely active rheumatoid arthritis who have an inadequate response to methotrexate alone.

**Development Phase:** IIb

**Status:** Final

**Version:** 2.00

**Date:** 20-Jul-2016

**EudraCT:** 2012-003635-31

**IND:** 115,510

**Sponsor:** Galapagos NV  
Industriepark Mechelen Noord  
Generaal De Wittelaan L11 A3  
2800 Mechelen, Belgium

**Biostatistician:**

Tel: [REDACTED]

GSM: [REDACTED]

**Clinical Study Lead:** [REDACTED]

---

### CONFIDENTIALITY STATEMENT

The information contained in this document is privileged and confidential. It is the property of Galapagos NV and may not be used, disclosed, reproduced or otherwise disseminated within your organization or communicated to any third parties without the express written authorization of Galapagos NV.

---

## STATISTICAL ANALYSIS PLAN HISTORY

<b>SAP / Amendment #</b>	<b>Date</b>	<b>Version</b>
Statistical Analysis Plan-Addendum	13-Jun-2016	Version 1.0
Statistical Analysis Plan-Addendum v2.0	20-Jul-2016	Version 2.0: In section “Step 3: Correlation of each individual biomarker versus week 12 efficacy and safety”, supporting tables were added to the plots to display the Pearson or point-biserial correlation coefficients. In section “Step 4: Biomarkers predicting the Week 12 Efficacy and Safety”, baseline value of the efficacy or safety parameters will be added as covariate in the ANCOVA model expect for ACR-N parameters.

## TABLE OF CONTENTS

<b>List of Tables .....</b>	<b>4</b>
<b>List of Figures .....</b>	<b>4</b>
<b>List of Listings .....</b>	<b>5</b>
<b>1. Introduction .....</b>	<b>6</b>
<b>2. PD Populations .....</b>	<b>6</b>
<b>3. Treatment Groups.....</b>	<b>6</b>
3.1. Randomized versus Actual Treatment .....	6
3.2. Treatment Group Labels.....	6
<b>4. Analysis Periods and Analysis Time Points .....</b>	<b>6</b>
4.1. Relative Number of Days .....	6
4.2. Algorithm of Allocating Visits to Time Windows .....	7
4.3. Selection of Visits .....	7
<b>5. Handling of Data .....</b>	<b>8</b>
5.1. Calculation of Descriptive Statistics .....	8
5.2. Calculation of Percentages .....	8
5.3. Handling of Values below (or above) a Threshold .....	8
5.4. Handling of Missing Data .....	8
5.4.1. Observed Cases (OC) .....	8
5.4.1. Last Observation Carried Forward (LOCF) .....	8
5.4.2. Handling of Missing Date(Time) or Partially Known Date(Time).....	9
5.5. Handling of Outliers.....	9
5.6. Stratification factors.....	9
<b>6. Software and Procedures .....</b>	<b>9</b>
6.1. Software.....	9
6.2. Procedures .....	9
6.3. Formats.....	10
<b>7. Statistical Methods.....</b>	<b>10</b>
7.1. Planned Analyses, Protocol Amendments Included .....	10
7.2. Changes to the Planned Analyses, Not Covered by Protocol Amendments .....	10
<b>8. Definitions of Pharmacodynamics Analysis Tables, Listings and Figures.....</b>	<b>10</b>
8.1. PD Parameters .....	11
8.2. Step 1: Overall analysis of the biomarkers.....	11
8.3. Step 2: Analysis of the biomarkers, split up for W12 responders and non-responders.....	15
8.4. Step 3: Correlation of each individual biomarker versus Week 12 efficacy and safety.....	16
8.5. Step 4: Biomarkers predicting the Week 12 Efficacy and Safety .....	17
<b>9. References .....</b>	<b>20</b>

**LIST OF TABLES**

Table PD 1: Descriptive statistics of the actual values ..... 11

Table PD 2: Descriptive statistics of the changes from baseline..... 11

Table PD 3: Descriptive statistics of the percent changes from baseline..... 12

Table PD 4: Summary with overall p-value and pairwise comparisons versus placebo ..... 12

Table PD 5: Exploratory pairwise comparisons between regimens within each daily dose level..... 13

Table PD 6: Descriptive statistics of the actual values, split up for Week 12 responders vs non-responders ..... 15

Table PD 7: Descriptive statistics of the changes from baseline, split up for Week 12 responders vs non-responders ..... 16

Table PD 8: Descriptive statistics of the percent changes from baseline, split up for Week 12 responders vs non-responders ..... 16

Table PD 9: Summary on changes from baseline with overall p-value and p-value for Week 12 response/non-response ..... 16

Table PD 10: Exploratory model to assess the relation between baseline PD markers and Week 12 response ..... 17

Table PD 11: Exploratory model to assess the relation between the Week 4 change from baseline in PD markers and Week 12 response ..... 18

Table PD 12: Exploratory model to assess the relation between the Week 4 %change from baseline in PD markers and Week 12 response ..... 18

Table PD 13: Exploratory model to assess the relation between baseline PD markers and Week 12 changes from baseline in efficacy..... 18

Table PD 14: Exploratory model to assess the relation between the Week 4 change from baseline in PD markers and Week 12 changes from baseline in efficacy ..... 18

Table PD 15: Exploratory model to assess the relation between the Week 4 percent change from baseline in PD markers and Week 12 changes from baseline in efficacy..... 18

Table PD 16: Exploratory model to assess the relation between baseline PD markers and Week 12 changes from baseline in safety ..... 19

Table PD 17: Exploratory model to assess the relation between the Week 4 change from baseline in PD markers and Week 12 changes from baseline in safety ..... 19

Table PD 18: Exploratory model to assess the relation between the Week 4 percent change from baseline in PD markers and Week 12 changes from baseline in safety..... 19

**LIST OF FIGURES**

Figure PD 1: Mean(SE) plot of the actual values ..... 14

Figure PD 2: Mean(SE) plot of the changes from baseline .....15

Figure PD 3: Mean(SE) plot of the percent changes from baseline.....15

Figure PD 4: Box plot of the actual values .....15

Figure PD 5: Box plot of the changes from baseline .....15

Figure PD 6: Box plot of the percent changes from baseline .....15

Figure PD 7: Scatterplots of the PD markers.....15

Figure PD 8: Mean(SE) plot of the actual values: Week 12 responders vs non-responders .....16

Figure PD 9: Mean(SE) plot of the changes from baseline: Week 12 responders vs non-responders.....16

Figure PD 10: Mean(SE) plot of the percent changes from baseline: Week 12 responders vs non-responders .....16

Figure PD 11: Box plot of the actual values: Week 12 responders vs non-responders .....16

Figure PD 12: Box plot of the changes from baseline: Week 12 responders vs non-responders .....16

Figure PD 13: Box plot of the percent changes from baseline: Week 12 responders vs non-responders .....16

Figure PD 14: Scatterplots of PD markers versus efficacy .....16

Figure PD 15: Scatterplots of PD markers versus safety .....17

Figure PD 16: Scatterplots of baseline PD markers versus efficacy.....18

Figure PD 17: Scatterplots of Week 4 change from baseline PD markers versus efficacy.....18

Figure PD 18: Scatterplots of Week 4 %change from baseline PD markers versus efficacy.....19

Figure PD 19: Scatterplots of baseline PD markers versus safety .....19

Figure PD 20: Scatterplots of Week 4 change from baseline PD markers versus safety.....19

Figure PD 21: Scatterplots of Week 4 percent change from baseline PD markers versus safety .....19

**LIST OF LISTINGS**

No listings planned.



## 1. INTRODUCTION

This document is an addendum to the main study SAP version 5.00. It is intended to define the statistical analysis of the additional pharmacodynamics measurements.

This Addendum to the final SAP was written after database unblinding and after viewing the main study results.

## 2. PD POPULATIONS

The PD population will be a subpopulation of the ITT from the main study analysis:

- All randomized subjects who received at least one dose of study medication and have at least one post-Baseline efficacy assessment. This is the ITT population in the main study analysis, available in ADaM.ADSL variable ITTFL.
- The 50 mg/day dose groups will be excluded (no PD samples analyzed for these groups; and still excluded from PD population if by accident there would be a result from a 50 mg/day subject).
- Also subjects who have no PD results (eg not allowed to analyze as per subject ICF version) will be excluded from this analysis.

The analysis population will always be indicated in a subtitle in the table, listing or figure.

## 3. TREATMENT GROUPS

### 3.1. RANDOMIZED VERSUS ACTUAL TREATMENT

For efficacy and PD parameters, the treatment group as assigned by the randomization will be used in the analysis (i.e., as-randomized analysis).

### 3.2. TREATMENT GROUP LABELS

The following treatment group labels will be used in the tables, listings and figures:

- Placebo
- GLPG0634 100 mg q.d.
- GLPG0634 200 mg q.d.
- GLPG0634 50 mg b.i.d.
- GLPG0634 100 mg b.i.d.

The 50 mg q.d. and 25 mg b.i.d. groups will not be used (because PD samples are not analysed).

## 4. ANALYSIS PERIODS AND ANALYSIS TIME POINTS

### 4.1. RELATIVE NUMBER OF DAYS

The relative day (DY) is calculated as follows:

= Visit date – reference date + 1 day, when the visit date is on or after the reference date

= Visit date – reference date, when the visit date is before the reference date.

The reference date in the study is the first study medication intake date, which by definition has DY=1. There is no DY=0.

## 4.2. ALGORITHM OF ALLOCATING VISITS TO TIME WINDOWS

The main study analysis used the following time windows:

All visits (including early discontinuation visits and unscheduled visits but excluding visits without data) will be placed into time windows according to their relative day (DY) in the study, according to the following allocation table:

Time point label	Target day	Interval lower bound	Interval upper bound
Screening (3)	NA	NA	NA
Baseline (1)	NA	NA	NA
Week 1	8	2	11
Week 2	15	12	21
Week 4	29	22	42
Week 8	57	43	70
Week 12	85	71	98
Week 16	113	99	126
Week 20	141	127	154
Week 24	169	155	182
>Week 24 (2)	NA	183	+∞
Follow-up (3)	7-10 days after last visit	NA	NA

(NA = not applicable).

### Notes:

- (1) The actual baseline reference value will be determined per parameter as the last available pre-dosing data point, so might differ from this “baseline” visit interval. There will be only one baseline time point in this study: the last non-missing value prior to first dosing of study drug.
- (2) Visits falling really late in the study (relative day > 182; excluding the follow-up visit) will be allocated to a “>Week 24” interval, and will not be shown in tables nor figures, but will only be listed.
- (3) For the screening and follow-up visits, no time window is defined. The visit as recorded in the eCRF will be used. Follow-up visits of subjects who discontinue the study prior to the Week 12 re-randomization will be shown separately as well in the tables that present the first part of the study only. For tables that present the study as a whole, all follow-up visits will be shown.

Tables, figures and listings will present the time points, not the visits.

For this additional PD analysis, only Baseline, Week 4 and Week 12 samples will be analyzed. These will be placed into the above time windows to allow a matching with the main analysis. If a PD result ends up in an analysis visit different from “baseline”, “week 4” and “week 12”, then it will not be used in the tables and figures.

## 4.3. SELECTION OF VISITS

It is possible that more than one visit gets allocated into the same time window. In that case, only one visit will be selected for tables and figures. The non-selected visit(s) will only be listed and flagged as use=no.

The visit with a relative day (DY) closest to the target day will be selected. If there are multiple visits at the same distance of the scheduled visit day (meaning: equal  $ABS(DY - \text{target day})$ ), then the one latest in time is selected.

In case more than one parameter is measured per time point, the selection is performed per parameter and per time point, not per “sample” and per time point. Missing values are removed before the selection is made.

## **5. HANDLING OF DATA**

### **5.1. CALCULATION OF DESCRIPTIVE STATISTICS**

For continuous parameters, descriptive statistics will be presented when  $N \geq 2$ . When  $N=1$ , the observation will not be shown in the table but only in the listing.

Descriptive statistics will include at least the following:

- the number of non-missing data points (N)
- the number of values above the lower limit of quantification and below the upper limit of quantification (only applicable to tables of actual values)
- the arithmetic mean
- the geometric mean (only applicable to tables of actual values)
- the standard error (SE) and the standard deviation (SD)
- the median, minimum and maximum
- 95% confidence interval of the mean (only when requested).

### **5.2. CALCULATION OF PERCENTAGES**

Missing values will not be included in the denominator count when computing percentages.

### **5.3. HANDLING OF VALUES BELOW (OR ABOVE) A THRESHOLD**

Values below the detection limit (“BLQ”, “<x” or “OOR <”) will be imputed by a value of “0.1”. Note that the central lab was unable to provide a value for the LLOQ.

Values flagged as extrapolated (ie, a result flagged with an asterisk) in the transfer from the bioanalysis lab will be handled as if they are BLQ, and will be imputed by “0.1” as well.

### **5.4. HANDLING OF MISSING DATA**

#### **5.4.1. Observed Cases (OC)**

No imputation is done of missing values.

#### **5.4.1. Last Observation Carried Forward (LOCF)**

A missing W4 analysis visit will be imputed by the last preceding non-missing data point. This can be a baseline value.

A missing W12 analysis visit will be imputed by the last preceding non-missing data point. Again this could be a baseline value. This will only be done for the main descriptive statistics and comparisons vs placebo. Correlation analyses will be performed on OC only.

#### **5.4.2. Handling of Missing Date(Time) or Partially Known Date(Time)**

No imputation will be done of missing date(time) fields, nor of the missing parts of partially known date(time) fields. When needed, a worst-case selection will be made.

### **5.5. HANDLING OF OUTLIERS**

All measured values will be included in the analyses. However, analyses may be repeated after exclusion of outliers.

### **5.6. STRATIFICATION FACTORS**

In case there is a discrepancy in randomization stratification stratum between IWRS and the actual eCRF data, then the actual eCRF stratum will be used in analyses and analysis models.

These are available in the final ADaM.ADSL dataset: region and previous use of biologics.

## **6. SOFTWARE AND PROCEDURES**

### **6.1. SOFTWARE**

SAS version 9.2 will be used for programming.

Each TLF will be presented in a separate PDF file. Image files (e.g. PNG or in Excel) should be available for plots.

SAS line size and page size:

- SAS LSIZE = 144
- SAS PSIZE = 48

Page layout:

- landscape
- paper size: A4
- margins in MS Word:
  - top, bottom, left, right: 2.5 cm
  - gutter: 0 cm
  - header, footer, from edge: 1.3 cm

### **6.2. PROCEDURES**

Analyses will comply with ICH regulations, in particular: (ICH-E3) , (ICH-E6) and (ICH-E9).

The following Galapagos SOPs will be followed:

- SOP-CLI-001: Developing Clinical Study Documents (version 1.0)
- SOP-CLI-003: Managing Data Processing Activities (version 1.0)

The following [REDACTED] STAT SOPs will be followed:

- [REDACTED]-SOP-ST-003\_1\_0: Statistical Analysis (Date effective: 30 September 2015)
- [REDACTED]-SOP-ST-004\_1\_0: SAS Programming for Statistical Analysis (Date effective: 30 September 2015)

### **6.3. FORMATS**

Derived data will be following CDISC ADaM 2.1 format. Since this extra analysis only contains PD, only one derived ADaM dataset ADPD2 will be created, containing the new PD parameters. A “DEFINE” file is not needed.

## **7. STATISTICAL METHODS**

### **7.1. PLANNED ANALYSES, PROTOCOL AMENDMENTS INCLUDED**

A descriptive analysis will present the actual values, changes from baseline and percent changes from baseline.

Treatment arms will be compared (similarly to efficacy) via an ANCOVA model on the changes from baseline, with explanatory factors: treatment, baseline value, stratification factors (geographical region, previous use of biologics). Pairwise comparisons versus placebo will be presented, corrected for multiplicity using Hommel’s method. Exploratory comparisons between the two regimens within each dose level will be presented as well.

Correlations versus efficacy and safety parameters will be investigated via logistic regression or ANCOVA models.

Figures will be added to depict the data.

Early dropouts and missing analysis visits will be imputed via the LOCF algorithm for the descriptive analysis and comparisons vs placebo. Correlation analyses will be performed on OC only.

### **7.2. CHANGES TO THE PLANNED ANALYSES, NOT COVERED BY PROTOCOL AMENDMENTS**

Not applicable – all these analyses are post-lock.

## **8. DEFINITIONS OF PHARMACODYNAMICS ANALYSIS TABLES, LISTINGS AND FIGURES**

All analyses will be run on the PD population, unless specified otherwise.

### 8.1. PD PARAMETERS

The raw PD markers will be provided by the bioanalysis lab: GM-CSF, IFN $\gamma$ , IL-1 $\beta$ , IL-2, IL-4, IL-5, IL-6, IL-7, IL-8, IL-10, IL-12 (p70), IL-13, IL-17A, IL-21, IL-23, MIP-1 $\alpha$ , MIP-1 $\beta$  and TNF- $\alpha$ . All results are expressed in pg/mL. The raw values (with their unit) will be used in the analysis without any transformation or manipulation (except for the imputation of values BLOQ or missing values).

Analysis will show actual values, changes from baseline and percent changes from baseline.

%change from baseline = 100 x (change from baseline) / baseline, and is only defined when the baseline  $\neq$  0.

### 8.2. STEP 1: OVERALL ANALYSIS OF THE BIOMARKERS

**Table PD 1: Descriptive statistics of the actual values**

Descriptive statistics of the actual values per time point, and per treatment group.

The table will show:

- LOCF
- observed-case

Mock:

		GLPG0634 q.d. dose groups			GLPG0634 b.i.d. dose groups	
Week	Statistic	Placebo (N=XXX)	100 mg q.d. (N=XXX)	200 mg q.d. (N=XXX)	50 mg b.i.d. (N=XXX)	100 mg b.i.d. (N=XXX)
<b>Baseline</b>						
	N	XXX	XXX	XXX	XXX	XXX
	N_det[1]	XXX	XXX	XXX	XXX	XXX
	Mean	XX.X	XX.X	XX.X	XX.X	XX.X
	SE	XX.XX	XX.XX	XX.XX	XX.XX	XX.XX
	SD	XX.XX	XX.XX	XX.XX	XX.XX	XX.XX
	Median	XX.X	XX.X	XX.X	XX.X	XX.X
	Min,Max	XX,XX	XX,XX	XX,XX	XX,XX	XX,XX
	GeoMean	XX.X	XX.X	XX.X	XX.X	XX.X
<b>Week 4 LOCF</b>						
	N	XXX	XXX	XXX	XXX	XXX
	N_det[1]	XXX	XXX	XXX	XXX	XXX
	Mean	XX.X	XX.X	XX.X	XX.X	XX.X
	SE	XX.XX	XX.XX	XX.XX	XX.XX	XX.XX
	SD	XX.XX	XX.XX	XX.XX	XX.XX	XX.XX
	Median	XX.X	XX.X	XX.X	XX.X	XX.X
	Min,Max	XX,XX	XX,XX	XX,XX	XX,XX	XX,XX
	GeoMean	XX.X	XX.X	XX.X	XX.X	XX.X
<b>Week 12 LOCF</b>						
	N	XXX	XXX	XXX	XXX	XXX
	N_det[1]	XXX	XXX	XXX	XXX	XXX
	Mean	XX.X	XX.X	XX.X	XX.X	XX.X
	SE	XX.XX	XX.XX	XX.XX	XX.XX	XX.XX
	SD	XX.XX	XX.XX	XX.XX	XX.XX	XX.XX
	Median	XX.X	XX.X	XX.X	XX.X	XX.X
	Min,Max	XX,XX	XX,XX	XX,XX	XX,XX	XX,XX
	GeoMean	XX.X	XX.X	XX.X	XX.X	XX.X

[1] Number of results >LLOQ and <ULOQ.

Note that this table contains geometric means, but none of the other tables need to.

**Table PD 2: Descriptive statistics of the changes from baseline**

Descriptive statistics of the changes from baseline per time point, and per treatment group.

The table will show:

- LOCF
- observed-case

Mock:

GLPG0634-CL-203 (DARWIN1) SAS date and time  
Page X of Y

Table PD 2: Descriptive statistics of the changes from baseline

---

Add-on therapy to methotrexate. Population: PD. Imputation method: LOCF (last observation carried forward)  
Parameter: PD marker (pg/mL)

---

Week	Statistic	Placebo (N=XXX)	GLPG0634 q.d. dose groups		GLPG0634 b.i.d. dose groups	
			100 mg q.d. (N=XXX)	200 mg q.d. (N=XXX)	50 mg b.i.d. (N=XXX)	100 mg b.i.d. (N=XXX)
Week 4 LOCF						
	N	XXX	XXX	XXX	XXX	XXX
	Mean	XX.X	XX.X	XX.X	XX.X	XX.X
	SE	XX.XX	XX.XX	XX.XX	XX.XX	XX.XX
	SD	XX.XX	XX.XX	XX.XX	XX.XX	XX.XX
	Median	XX.X	XX.X	XX.X	XX.X	XX.X
	Min,Max	XX,XX	XX,XX	XX,XX	XX,XX	XX,XX
Week 12 LOCF						
	N	XXX	XXX	XXX	XXX	XXX
	Mean	XX.X	XX.X	XX.X	XX.X	XX.X
	SE	XX.XX	XX.XX	XX.XX	XX.XX	XX.XX
	SD	XX.XX	XX.XX	XX.XX	XX.XX	XX.XX
	Median	XX.X	XX.X	XX.X	XX.X	XX.X
	Min,Max	XX,XX	XX,XX	XX,XX	XX,XX	XX,XX

**Table PD 3: Descriptive statistics of the percent changes from baseline**

Repeat table for the percent changes from baseline.

**Table PD 4: Summary with overall p-value and pairwise comparisons versus placebo**

ANCOVA model on the LOCF change from baseline biomarker at each time point. Factors in the model are: treatment, baseline value of the PD marker, region, and previous use of biologics.

Per time point, the table will present:

- Overall type-III p-value for all explanatory factors in the model.
- Per group: difference in LS-means versus the placebo group (GLPG0634 minus placebo), together with a 95% confidence interval of this difference.
- Per group: p-value for pairwise comparison versus the placebo group, derived from contrasts within the overall model. These p-values will be presented both as uncorrected p-values as well as corrected for multiplicity according to Hommel’s closed-testing method (SAS PROC MULTTEST).

Mock:

GLPG0634-CL-203 (DARWIN1)

SAS date and time  
Page X of Y

Table PD 4: Summary with overall p-value and pairwise comparisons versus placebo (LOCF)

Add-on therapy to methotrexate. Population: PD. Imputation method: LOCF (last observation carried forward)  
Parameter: PD marker (pg/mL)

Week and statistic	Overall	GLPG0634 q.d. dose groups		GLPG0634 b.i.d. dose groups		
		Placebo (N=XXX)	100 mg q.d. (N=XXX)	200 mg q.d. (N=XXX)	50 mg b.i.d. (N=XXX)	100 mg b.i.d. (N=XXX)
<b>Week 4 LOCF</b>						
Overall p-values: [1]						
- Treatment	X.XXXX					
- Baseline value	X.XXXX					
- Region	X.XXXX					
- Prior biologics	X.XXXX					
Pairwise comparisons versus Placebo: [2]						
- Sample mean (SE)		XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)
- LS-mean (SE)		XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)
- LS-mean difference (SE)			XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)
- 95% CI of the difference			XXX.X, XXX.X	XXX.X, XXX.X	XXX.X, XXX.X	XXX.X, XXX.X
- Uncorrected p-value			X.XXXX	X.XXXX	X.XXXX	X.XXXX
- Corrected p-value			X.XXXX	X.XXXX	X.XXXX	X.XXXX
<b>Week 12 LOCF</b>						
Overall p-values: [1]						
- Treatment	X.XXXX					
- Baseline value	X.XXXX					
- Region	X.XXXX					
- Prior biologics	X.XXXX					
Pairwise comparisons versus Placebo: [2]						
- Sample mean (SE)		XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)
- LS-mean (SE)		XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)
- LS-mean difference (SE)			XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)
- 95% CI of the difference			XXX.X, XXX.X	XXX.X, XXX.X	XXX.X, XXX.X	XXX.X, XXX.X
- Uncorrected p-value			X.XXXX	X.XXXX	X.XXXX	X.XXXX
- Corrected p-value			X.XXXX	X.XXXX	X.XXXX	X.XXXX

[1] Type III overall p-values from an ANCOVA model on the change from baseline per time point, with factors treatment, baseline value, region and previous use of biologics.  
[2] Pairwise comparisons of each group versus the placebo group (derived as contrasts within the same overall model): difference in LS-means (GLPG0634 - placebo) with 95% confidence interval of this difference, uncorrected p-value and Hommel-corrected p-value.

The table will show:

- LOCF
- observed-case

**Table PD 5: Exploratory pairwise comparisons between regimens within each daily dose level**

ANCOVA model on the change from baseline at each time point. Factors in the model are: baseline value, regimen, dose, regimen\*dose interaction, region, and previous use of biologics. Placebo subjects will not be used in this model (as there is no q.d./b.i.d. regimen available for placebo). The regimen\*dose interaction will be removed from the final model when not significant at level 0.10 (two-sided type-III p-value). Regimen will have 2 categories: q.d. or b.i.d. Dose will have 2 categories, corresponding to the total daily dose: 100 mg/day or 200 mg/day.

Per time point, the table will present:

- Per dose: difference in LS-means of q.d. versus b.i.d. within the same dose level (100 or 200 mg/day), together with a 95% confidence interval of this difference.
- Per group: p-value for pairwise comparison of q.d. versus b.i.d. These p-values will be presented both as uncorrected p-values, and are derived as contrasts within the overall model. As this is a purely exploratory analysis, no adjustment for multiplicity will be performed.

Mock:



GLPG0634-CL-203 (DARWIN1)

SAS date and time  
Page X of Y

Table PD 5: Exploratory pairwise comparisons between regimens within each daily dose level (LOCF)

Add-on therapy to methotrexate. Population: PD. Imputation method: LOCF (last observation carried forward)  
Parameter: PD marker (pg/mL)

Week	Statistic	Overall	GLPG0634 100 mg/day	GLPG0634 200 mg/day
Week 4 LOCF	Overall model p-values: [1]			
	- Baseline value	X.XXXXX		
	- Dose	X.XXXXX		
	- Regimen	X.XXXXX		
	- Dose*regimen	X.XXXXX		
	- Region	X.XXXXX		
	- Previous biologics	X.XXXXX		
	Pairwise comparisons of q.d. versus b.i.d.: [2]			
	- LS-mean difference (SE) [3]		XX.X (XX.XX)	XX.X (XX.XX)
	- 95% CI of the difference		XXX.X,XXX.X	XXX.X,XXX.X
- Uncorrected p-value		X.XXXXX	X.XXXXX	
Week 12 LOCF	Overall model p-values: [1]			
	- Baseline value	X.XXXXX		
	- Dose	X.XXXXX		
	- Regimen	X.XXXXX		
	- Dose*regimen	X.XXXXX		
	- Region	X.XXXXX		
	- Previous biologics	X.XXXXX		
	Pairwise comparisons of q.d. versus b.i.d.: [2]			
	- LS-mean difference (SE) [3]		XX.X (XX.XX)	XX.X (XX.XX)
	- 95% CI of the difference		XXX.X,XXX.X	XXX.X,XXX.X
- Uncorrected p-value		X.XXXXX	X.XXXXX	

[1]: Type III p-values from an ANCOVA model on the change from baseline per time point, with factors baseline value, dose (100/200 mg/day excluding placebo), regimen (q.d./b.i.d.), dose\*regimen interaction, region and previous use of biologics. The dose\*regimen interaction is only kept in the model when p<0.10.  
[2]: Comparisons within the same dose level of the q.d. versus b.i.d. regimens, derived as contrast within the same overall model.  
[3]: Difference: (q.d. group) - (b.i.d. group).

The table will show:

- LOCF
- observed-case

**Figure PD 1: Mean(SE) plot of the actual values**

Mean(SE) plot over time of the actual values. Use a time-proportional horizontal axis.

Mock:

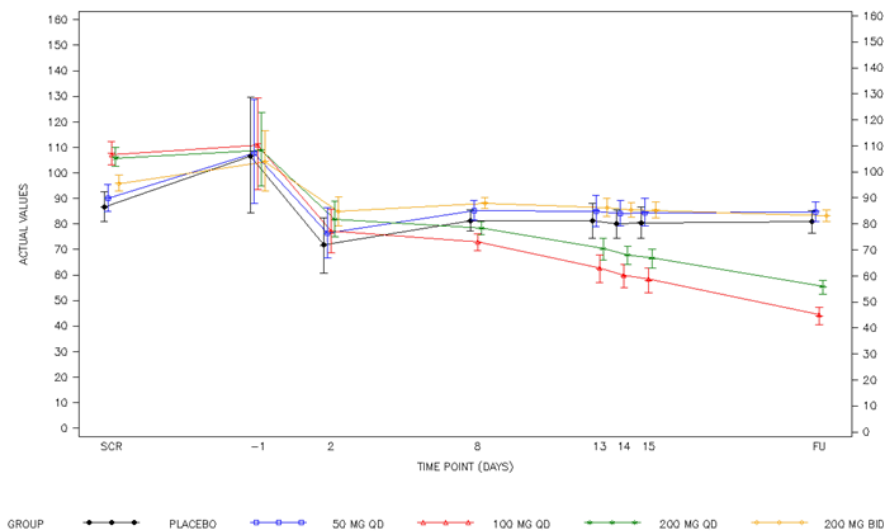
GLPGXXXX-CL-XXXX

SAS date and time  
Page X of Y

Figure 14.3.2.2: Laboratory data: Mean plots of the actual values over time

Population: Safety

Clinical chemistry: Lab test name (unit)  
Normal range: XX.X - XX.X



The plot will show:

- LOCF
- observed-case

Note that this is a generic mock plot. The “normal range” line does not apply to this PD analysis.

The following standardized colors, line types and plot symbols will be used consistently throughout the entire analysis:

<b>Treatment group label</b>	<b>plot symbol</b>	<b>line type</b>	<b>color</b>
Placebo	Dot	Full line	Black
GLPG0634 100 mg q.d.	Square	Full line	Blue
GLPG0634 200 mg q.d.	Diamond	Dotted line	Green
GLPG0634 50 mg b.i.d.	Star	Full line	Red
GLPG0634 100 mg b.i.d.	Plus	Dotted line	Orange

**Figure PD 2: Mean(SE) plot of the changes from baseline**

**Figure PD 3: Mean(SE) plot of the percent changes from baseline**

**Figure PD 4: Box plot of the actual values**

**Figure PD 5: Box plot of the changes from baseline**

**Figure PD 6: Box plot of the percent changes from baseline**

**Figure PD 7: Scatterplots of the PD markers**

Individual scatterplots for each PD marker versus the others. A separate plot per time point.

- A) Actual values
- B) Change from baseline
- C) %change from baseline

### **8.3. STEP 2: ANALYSIS OF THE BIOMARKERS, SPLIT UP FOR W12 RESPONDERS AND NON-RESPONDERS**

All parameters shown are OC only.

**Table PD 6: Descriptive statistics of the actual values, split up for Week 12 responders vs non-responders**

Repeats of the analysis of previous section, splitting up each page for Week 12 responders and non-responders on a separate page:

- A) Subjects who had ACR20 response at W12
- B) Subjects who did not have ACR20 response at W12
- C) Subjects who had ACR50 response at W12
- D) Subjects who did not have ACR50 response at W12
- E) Subjects who had DAS28 remission at W12
- F) Subjects who did not have DAS28 remission at W12
- G) Subjects who had DAS28 remission or low disease activity at W12

H) Subjects who did not have DAS28 remission or low disease activity at W12.

**Table PD 7: Descriptive statistics of the changes from baseline, split up for Week 12 responders vs non-responders**

**Table PD 8: Descriptive statistics of the percent changes from baseline, split up for Week 12 responders vs non-responders**

**Table PD 9: Summary on changes from baseline with overall p-value and p-value for Week 12 response/non-response**

ANCOVA model on the change from baseline in biomarker at each time point. Factors in the model are: treatment, baseline value of the PD marker, the response (W12 OC yes/no result), region, and previous use of biologics.

The p-value for the “response” parameter will indicate whether the average change in biomarker is different for responders vs non-responders.

Plain unformatted SAS output is acceptable.

For each of the responder/non-responder combinations of above tables.

**Figure PD 8: Mean(SE) plot of the actual values: Week 12 responders vs non-responders**

Mean(SE) plots over time of the actual values like the above, eg split over 2 panels: subjects who reached response at W12 vs subjects who did not reach response at W12 (OC). For each of the responder/non-responder combinations of above tables.

**Figure PD 9: Mean(SE) plot of the changes from baseline: Week 12 responders vs non-responders**

**Figure PD 10: Mean(SE) plot of the percent changes from baseline: Week 12 responders vs non-responders**

**Figure PD 11: Box plot of the actual values: Week 12 responders vs non-responders**

**Figure PD 12: Box plot of the changes from baseline: Week 12 responders vs non-responders**

**Figure PD 13: Box plot of the percent changes from baseline: Week 12 responders vs non-responders**

#### **8.4. STEP 3: CORRELATION OF EACH INDIVIDUAL BIOMARKER VERSUS WEEK 12 EFFICACY AND SAFETY**

All parameters are analyzed as OC only.

**Figure PD 14: Scatterplots of PD markers versus efficacy**

Scatterplots of each PD marker vs the efficacy variables. Separate plot per time point.

To depict the correlation numerically, a supporting table will be added to show the Pearson correlation coefficient at each time point and for each group. Versus binary responses (ACR20/50 and DAS28(CRP) remission and remission/LDA), a point-biserial correlation coefficient will be presented. Correlations will be flagged as: strong ( $>0.7$  in absolute value), moderate ( $>0.5$  in absolute value), or weak ( $>0.3$  in absolute value).

Efficacy variables: ACR-N, DAS28(CRP), TJC68, SJC66, TJC28, SJC28, HAQ-DI, Pain VAS, Investigator's global disease activity VAS, Subject's global disease activity VAS, FACIT-F and CRP.

- A) Actual values
- B) Change from baseline
- C) Percent changes from baseline.

#### **Figure PD 15: Scatterplots of PD markers versus safety**

Scatterplots of each PD marker vs the safety variables. Separate plot per time point.

To depict the correlation numerically, a supporting table will be added to show the Pearson correlation coefficient at each time point and for each group. Correlations will be flagged as: strong ( $>0.7$  in absolute value), moderate ( $>0.5$  in absolute value), or weak ( $>0.3$  in absolute value).

Safety variables: hemoglobin, lymphocyte absolute count and subsets (CD3, CD4, CD8, CD4/CD8 ratio, CD16 CD56, CD19), neutrophils absolute count, and platelet count.

- A) Actual values
- B) Change from baseline
- C) Percent changes from baseline.

### **8.5. STEP 4: BIOMARKERS PREDICTING THE WEEK 12 EFFICACY AND SAFETY**

All parameters are analyzed as OC only.

#### **Table PD 10: Exploratory model to assess the relation between baseline PD markers and Week 12 response**

Logistic regression model, modeling the responders.

Factors in the model are: treatment, region, previous use of biologics, and the baseline values of all the PD markers. Apply a backwards elimination process to retain only relevant parameters (significance level 0.05). Treatment must always be kept in the model.

Plain SAS procedure output is acceptable. Only the final model will be shown.

For responses:

- A) ACR20
- B) ACR50
- C) DAS28(CRP) remission
- D) DAS28(CRP) remission and low disease activity

**Table PD 11:** Exploratory model to assess the relation between the Week 4 change from baseline in PD markers and Week 12 response

**Table PD 12:** Exploratory model to assess the relation between the Week 4 %change from baseline in PD markers and Week 12 response

**Table PD 13:** Exploratory model to assess the relation between baseline PD markers and Week 12 changes from baseline in efficacy

ANCOVA model, modeling: the change from baseline in efficacy at W12.

Factors in the model are: treatment, region, previous use of biologics, baseline value (except for ACR-N parameter), and the baseline values of all the PD markers. Apply a backwards elimination process (significance level 0.05). Treatment must always be kept in the model.

Plain SAS procedure output is acceptable. Only the final model will be shown.

For efficacy parameters (Week 12 OC):

- A) ACR-N
- B) DAS28(CRP)
- C) TJC68
- D) SJC66
- E) TJC28
- F) SJC28
- G) HAQ-DI
- H) Pain VAS
- I) Investigator's global disease activity VAS
- J) Subject's global disease activity VAS
- K) FACIT-F
- L) CRP

**Table PD 14:** Exploratory model to assess the relation between the Week 4 change from baseline in PD markers and Week 12 changes from baseline in efficacy

**Table PD 15:** Exploratory model to assess the relation between the Week 4 percent change from baseline in PD markers and Week 12 changes from baseline in efficacy

**Figure PD 16:** Scatterplots of baseline PD markers versus efficacy

Scatterplots of each PD marker (baseline value) vs the efficacy variables (W12 change from baseline).

Efficacy variables: ACR-N, DAS28(CRP), TJC68, SJC66, TJC28, SJC28, HAQ-DI, Pain VAS, Investigator's global disease activity VAS, Subject's global disease activity VAS, FACIT-F and CRP.

**Figure PD 17:** Scatterplots of Week 4 change from baseline PD markers versus efficacy

**Figure PD 18: Scatterplots of Week 4 %change from baseline PD markers versus efficacy**

**Table PD 16: Exploratory model to assess the relation between baseline PD markers and Week 12 changes from baseline in safety**

ANCOVA model, modeling: the change from baseline in efficacy at W12.

Factors in the model are: treatment, region, previous use of biologics, baseline value, and the baseline values of all the PD markers. Apply a backwards elimination process (significance level 0.05). Treatment must always be kept in the model.

Plain SAS procedure output is acceptable. Only the final model will be shown.

For safety parameters (Week 12 OC):

- A) Hemoglobin
- B) Lymphocyte absolute count
- C) CD3 absolute count
- D) CD4 absolute count
- E) CD8 absolute count
- F) CD4/CD8 ratio
- G) CD16 CD56 absolute count
- H) CD19 absolute count
- I) Neutrophils absolute count
- J) Platelet count

**Table PD 17: Exploratory model to assess the relation between the Week 4 change from baseline in PD markers and Week 12 changes from baseline in safety**

**Table PD 18: Exploratory model to assess the relation between the Week 4 percent change from baseline in PD markers and Week 12 changes from baseline in safety**

**Figure PD 19: Scatterplots of baseline PD markers versus safety**

Scatterplots of each PD marker (baseline value) vs the safety variables (W12 change from baseline).

Safety variables: hemoglobin, lymphocyte absolute count, neutrophils absolute count, and platelet count.

**Figure PD 20: Scatterplots of Week 4 change from baseline PD markers versus safety**

**Figure PD 21: Scatterplots of Week 4 percent change from baseline PD markers versus safety**

## 9. REFERENCES

ICH-E3. (December 1995). *Structure and content of clinical study reports. Step 4 Guideline.*

ICH-E6. (17 July 1996). *Guideline for good clinical practice. Step 5 Guideline.*

ICH-E9. (5 February 1998). *Statistical principles for clinical trials. Step 4 guideline.*