# CENTER FOR DRUG EVALUATION AND RESEARCH

***APPLICATION NUMBER:***

**201292Orig1s000**

# STATISTICAL REVIEW(S)

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

# STATISTICAL REVIEW AND EVALUATION

## CLINICAL STUDIES

| | |
|---|---|
| **NDA/BLA #:** | 201-292 |
| **Drug Name:** | Afatinib |
| **Indication(s):** | Treatment of patients with locally advanced or metastatic non-small cell lung cancer with epidermal growth factor receptor mutation(s) as detected by an FDA-approved test |
| **Applicant:** | Boehringer Ingelheim |
| **Date(s):** | Receipt Date: 11/15/2012 |
| | PDUFA Date: 7/15/2013 |
| **Review Priority:** | Priority |
| **Biometrics Division:** | V |
| **Statistical Reviewer:** | Jonathan Norton, PhD |
| **Concurring Reviewers:** | Kun He, PhD |
| | Rajeshwari Sridhara, PhD |
| **Medical Division:** | Division of Oncology Products 2 |
| **Clinical Team:** | Shakun Malik, MD |
| | Gideon Blumenthal, MD |
| | Patrician Keegan, MD |
| **Project Manager:** | Deanne Varney |

**Keywords:** Log-rank test, Cox regression model

# Table of Contents

# 1 EXECUTIVE SUMMARY

Boehringer-Ingelheim has submitted an original NDA for afatinib, a new molecular entity. The proposed indication is "for the treatment of patients with locally advanced or metastatic non-small cell lung cancer (NSCLC) with epidermal growth factor receptor (EGFR) mutation(s) as detected by an FDA-approved test."

Study 1200.32 was a randomized, open-label, parallel-arm Phase 3 trial comparing afatinib with pemetrexed/cisplatin. Eligible patients had Stage IIIB or IV adenocarcinoma, were EGFR-TKI (tyrosine-kinase inhibitor) naïve, and had an EGFR mutation. Screening was done using a laboratory developed test. Patients were randomized in a 2:1 ratio to afatinib or chemotherapy.

Patients on the afatinib arm received 40 mg daily in 21-day cycles. The dose could be adjusted between 20 mg and 50 mg, but all dose reductions were permanent. Patients in the chemotherapy arm received intravenous (IV) pemetrexed 500 mg/m$^2$ and IV cisplatin 75 mg/m$^2$ IV at the beginning of the 21-day cycle. Patients on either arm could switch to other anti-cancer therapy, including EGFR TKI drugs, after progression.

The primary endpoint of the study was progression-free survival (PFS) as determined by central assessment. Afatinib was statistically superior to chemotherapy by stratified log-rank test with p = 0.0003. The median PFS was 11.14 months (95% confidence interval (CI) = [9.63, 13.63]) in the afatinib arm compared to 6.90 months (95% CI = [5.39, 8.25]) in the chemotherapy arm. The estimated hazard ratio was 0.577 favoring afatinib, with a 95% CI of [0.425, 0.784]. The statistical results for investigator-determined PFS were quite similar, and there was an acceptable level of agreement. Other sensitivity analyses also supported the finding of superiority for afatinib on PFS.

There was no statistical difference between afatinib and chemotherapy for overall survival (OS, p = 0.55). The median OS on the afatinib arm was 28.06 months (95% CI = [24.64, 32.95]), compared to 28.16 (95% CI = [20.73, 33.22]) on the chemotherapy arm. The hazard ratio was 0.907 in favor with afatinib, with a 95% CI of [0.660, 1.1246].

Patients were stratified by the type of EGFR mutation: Del19, L858R, and Other. While subgroup analyses were planned for this and other factors, there was no control for multiplicity. This lack of control makes it difficult to interpret any subgroup results after the fact. With this caveat in mind, it is noteworthy that in the Other subgroup there was a nominally-significant finding (p = 0.03) that afatinib was worse than chemotherapy for OS. There was also a negative trend for on PFS within this subgroup, but it was not even nominally significant (p = 0.11).

Given the grave importance of the OS endpoint, an argument might be made to exclude patients in the Other subgroup from the labeled indication based on the nominally-significant evidence of increased mortality for patients on afatinib. Aside from the statistical concern based on a lack of type-I error control, such as an exclusion would arguably risk a logical contradiction: it may not make sense to ***distrust*** the inconclusive OS results in the study population of 345 patients but

***trust*** the same endpoint in a subgroup of 37 patients. Another point to consider is that the Other subgroup is not homogeneous. It is quite plausible that some patients within this group have mutations favorable to treatment with afatinib and others do not.

If PFS is deemed an acceptable primary endpoint then the study supports a finding of efficacy in the randomized population. While there is some evidence of harm in patients in the Other EGFR subgroup, it is not statistically persuasive. Given sufficient external evidence, it may be still be appropriate to restrict the indication. It is also appropriate to consider the availability of alternative treatments for patients with a mutation in the Other subgroup.

In conclusion, Study 1200.32 demonstrated a statistically-significant improvement in PFS for patients randomized to afatinib compared to those assigned to chemotherapy. The study did not show a statistically-significant difference in OS between the two treatment arms. Whether the results from Study 1200.32 provide evidence for a favorable benefit-risk ratio and support approval of afatinib will be determined by the clinical review team. The clinical team will also need to carefully weigh the evidence to determine the appropriate indication.

# 2 INTRODUCTION

## 2.1 Overview

On November 15, 2012, Boehringer-Ingelheim submitted an original NDA for afatinib, a new molecular entity. The proposed indication is "for the treatment of patients with locally advanced or metastatic non-small cell lung cancer (NSCLC) with epidermal growth factor receptor (EGFR) mutation(s) as detected by an FDA-approved test." The applicant intends to market 20, 30, 40, (b) (4) mg tablets that would be taken daily. A marketing application has also been submitted to the Center for Devices and Radiological Health for a proposed companion diagnostic.

The Division of Drug Oncology Products (DDOP) held a pre-Phase III meeting with the applicant on October 16, 2008. In response to a question about whether progression-free survival (PFS) would be acceptable as primary endpoint for Study 1200.32, DDOP stated, "In general, a substantial, robust improvement in PFS that is clinically meaningful and statistically persuasive, and has an acceptable risk-benefit profile may be considered for regulatory decision-making." However, DDOP also stated that this endpoint is subject to ascertainment bias and that progression events should be confirmed by blinded independent review.

On April 16, 2009, DDOP sent a Special Protocol Assessment – No Agreement letter for Study 1200.32. The letter noted that the proposed primary endpoint was PFS, whereas previous approved drugs for first-line NSCLC had used overall survival (OS). Further, the letter noted that patients will be discontinued based on investigator-assessed progression whereas the primary endpoint would be based on centrally-assessed progression. A sensitivity analysis was requested in which patients in the experimental arm are deemed to progress at the earlier date (between investigator and central assessors) and patients in the control arm progress at the later date. The applicant submitted the results of such an analysis in the NDA.

Table 1 shows the controlled studies submitted in this NDA for NSCLC. This review will focus on Study 1200.32, the Phase III study for first-line treatment of NSCLC in patients with an EGFR-activating mutation. For Study 1200.42, only the single-arm data has been submitted. Study 1200.23 failed on its primary endpoint of overall survival ($p = 0.74$) and hence will be considered a failed study. For the current submission, the Division of Oncology Products 2 (DOP2) will only consider an indication aligned with the patient population in Study 1200.32. (Due to a re-organization, this NDA now falls within the purview of DOP2 rather than the defunct DDOP.)

**Table 1: Submitted Controlled Studies of Afatinib for NSCLC**

| Study | Phase and Design | Treatment Arms with # of Subjects | Study Population |
|---|---|---|---|
| 1200.32<br>LUX-Lung 3 | Phase III<br>Randomized, open-label, parallel arm | Afatinib: 230<br>Chemotherapy: 115 | EGFR-TKI naïve patients harboring an EGFR-activating mutation |
| 1200.42<br>LUX-Lung 5<br><br>(Incomplete) | Phase III<br>Randomized, open-label, parallel arm with active run-in | *Part A*<br>Afatinib: 1100<br><br>*Part B (not submitted)*<br>Afatinib: 234<br>Chemotherapy: 117 | Patients who previously failed erlotinib or gefitinib |
| 1200.23<br>LUX-Lung 1 | Phase IIb/III<br>Randomized, double-blind, parallel-arm | Afatinib + BSC : 390 treated, 244 in primary analysis<br>Placebo + BSC: 195 treated, 114 in primary analysis | Patients who previously failed erlotinib or gefitinib |

Note: BSC = best supportive care

## 2.2 Data Sources

The applicant submitted analysis data, tabulation data in SDTM format, and source data. The application also included SAS program code for key analyses and to derive the analysis files from the source files. The data files can be found in the EDR at \\cdsesub1\EVSPROD\NDA201292\0000\m5\datasets\1200-0032 and \\cdsesub1\EVSPROD\NDA201292\0000\m5\datasets\1200-0032-source.

# 3 STATISTICAL EVALUATION

## 3.1 Data and Analysis Quality

The overall quality of the submission is acceptable.

The applicant sent updated survival data on January 28, 2013. On February 4, DOP2 sent the following information request on behalf of the reviewer:

> The protocol for study 1200.32 states that patients will be followed every 60 days (+/ - 15 days) until death. In the updated overall survival data sent on 1/28/2013, there were 33 patients who had a censoring date prior to 11/1/2012. In order to assess the benefit - risk profile of this product, FDA needs updated data on the vital status of these censored patients. We request that you use all practical methods, including checking appropriate public records, to provide either a death date or more recent censoring date (i.e., date last known alive) for each of these patients. We acknowledge that your ability to follow up on some patients may be limited by withdrawal of consent.

The applicant sent the requested information, but the effect on the study results was quite small. See Section 3.2.4 for further discussion.

## 3.2 Evaluation of Efficacy

### 3.2.1 Study Design and Endpoints

Study 1200.32 was a randomized, open-label, parallel-arm Phase 3 trial comparing afatinib with pemetrexed/cisplatin as a first-line treatment for patients with NSCLC. Eligible patients had Stage IIIB or IV adenocarcinoma, were EGFR-TKI (tyrosine-kinase inhibitor) naïve, and had an EGFR mutation. Screening was done using a laboratory developed test. (A sub-sample of patients was later tested using the proposed companion diagnostic.) Patients were randomized in a 2:1 ratio to afatinib or chemotherapy.

Patients on the afatinib arm received 40 mg daily in 21-day cycles. The dose could be adjusted between 20 mg and 50 mg, but all dose reductions were permanent. Treatment was continued until documented progression, intolerable toxicity, or patient/investigator request for permanent discontinuation. Patients in the chemotherapy arm received intravenous (IV) pemetrexed 500 $mg/m^2$ and IV cisplatin 75 $mg/m^2$ IV at the beginning of the 21-day cycle. These patients were to receive 6 cycles of treatment unless it was discontinued for one of the reason previously listed. The trial was open-label due to the different routes of administration in the two arms.

The primary endpoint was PFS. The "key" secondary endpoints were objective response (OR, defined as complete or partial response), disease control (defined as OR or stable disease) and OS.

PFS was assessed according to the RECIST 1.1 criteria using central, blinded, independent review. All scans were reviewed by two independent radiologists, with a third radiologist acting as an adjudicator if they disagreed. An oncologist then reviewed the clinical information and made the final determination of the date of the progression.

### 3.2.2 Statistical Methodologies

Patients were randomized using a central interactive voice/web response system. The randomization was stratified by EGFR mutation (L858R vs. Del19 vs. other mutation) and race (Asian vs. non-Asian).

The primary analysis for PFS was a log-rank test stratified by the randomization factors. The applicant provided the following code for the test:

```
proc lifetest data = surv_tst&tnum. method=km;
time eptn*eptcen(1);
test trt;
strata egfrstdc asiayndc/test=none;
ods output logunichisq=lrtest&tnum(keep=probchisq);
run;
```

The statistical analysis plan (SAP) stated that the SAS 9.2 software would be used for the analysis. According to the 9.2 manual, the above code calculates the ***stratified linear rank test***. The manual provides different syntax to compute the ***stratified log-rank test***. Compared to the code above, the correct test can be performed by removing the line "*test [arm_variable]*" *and* replacing "*test=none*" with "*group= [arm_variable]*" in the *strata* command. The reviewer replicated the applicant's results and also re-ran the primarily analysis using the appropriate syntax. The results were similar but not identical. See Section 3.2.4 for results.

Following a pre-specified hierarchy, the secondary endpoints were tested in the following order: OR, disease control, OS. Logistic regression was used to compare the rates of OR and disease control across treatment arms, with the analysis adjusted by the stratification factors. OS was analyzed using a stratified log-rank test, as with PFS.

The primary analysis was to occur after approximately 217 PFS events. Assuming a hazard ratio of .64 favoring afatinib, this was expected to provide 90% power. An interim OS analysis was to occur at the time of the primary analysis, using a stopping boundary of $p < 0.0001$. The final OS analysis was to occur after 209 deaths. The OS data had not matured at the time of NDA submission. At the request of DOP2, the applicant submitted an updated OS analysis in January 2013, at which point 175 deaths had occurred.

Despite the interest in the EGFR subgroups, there was no plan for controlling multiplicity across the three subgroups. The SAP stated that the efficacy would be "explored" in these subgroups along with numerous other factors.

Table 2, which was provided by the applicant, shows the rules for determining the PFS or censoring date. The applicant also conducted the following sensitivity analyses:

1) Investigator, rather than central, assessment
2) “Extreme worst case”: If investigator and central assessment disagree, choose the earlier date of progression for patients in the afatinib arm and choose the later date for patients in the chemotherapy arm.
3) “Symmetric worst case”: If investigator and central assessment disagree, choose the earlier date of progression
4) If investigator detects progressive disease but it is not verified by central assessment, then consider progression to occur at the next scheduled tumor assessment

**Table 2: Rules for determining PFS or censoring date**
(Source: Table 7.4:1, Trial Statistical Analysis Plan)

| Rule # | Situation | Outcome (event or censored) | Date of PFS event or censoring |
|---|---|---|---|
| 1 | No baseline tumour assessment (no death before second scheduled assessment) | censored | Date of randomisation |
| 2 | Progressed from central imaging (no missed radiologic assessments) | event | Date of PD |
| 3a | Non-PD from central imaging[1], death before next scheduled assessment | event | Date of death |
| 3b | Non-PD from central imaging[1], one missed assessment, death or progression after date of missed assessment, but before a second scheduled assessment | event | Date of PD or death |
| 3c | Non-PD from central imaging[1], more than one consecutive missed assessment, death or progression after date of second missed assessment | censored | Date of last imaging before missed assessment |
| 3d | Non-PD from central imaging[1], more than one consecutive missed assessment, non-PD according to imaging after missed assessments | censored | Date of last non-PD imaging |
| 4 | New anti-cancer medication before progression or death | censored | Date of last imaging before new anti-cancer medication |
| 5 | Death before the scheduled date of first imaging | event | Date of death |
| 6a | No imaging performed post-baseline, patient dies between first and second scheduled assessments | event | Date of death |
| 6b | No imaging performed post-baseline, patient dies after second scheduled assessment | censored | Date of randomisation |
| 6c | No imaging performed post-baseline, vital status is unknown or patient known to be alive | censored | Date of randomisation |
| 7 | Alive and not progressed from central imaging (no missed assessments) | censored | Date of last imaging |

[1] - From the last assessment at which CR, PR or SD was assessed.

### 3.2.3 Patient Disposition, Demographic and Baseline Characteristics

A total of 345 patients were randomized, 230 to afatinib and 115 to chemotherapy. Among the randomized patients, five were not treated, including one in the afatinib arm and four in the chemotherapy arm. The disposition of the treated patients is shown in Table 3.

**Table 3: Disposition of Treated Patients**

| | Afatinib N (%) | Chemotherapy N (%) | Total N (%) |
|---|---|---|---|
| Patients treated | 229 (100%) | 111 (100%) | 340 (100%) |
| No longer on treatment at cut-off | 164 (72%) | 111 (100%) | 275 (81%) |
| Completed chemotherapy | N.A. | 60 (54%) | N.A. |
| Progressive disease | 133 (58%) | 19 (17%) | 152 (45%) |
| Other AE | 23 (10%) | 17 (15%) | 40 (12%) |
| Non-compliance | 1 (0.4%) | 4 (4%) | 5 (1%) |
| Lost to follow-up | 0 | 0 (0%) | 0 (0%) |
| Refused to continue study medication | 6 (3%) | 11 (10%) | 17 (5%) |
| Other - Patient died | 1 (0.4%) | 0 (0%) | 1 (0.3%) |
| On treatment at cut-off | 65 (28%) | 0 (0%) | 65 (19%) |

Table 4 shows demographics and other baseline characteristics of the randomized patients.

**Table 4: Baseline Characteristics of Randomized Patients**

| Variable | Afatinib (N=230) | Chemotherapy (N=115) | Total (N=345) |
|---|---|---|---|
| Gender – N (%) | | | |
| Male | 83 (36%) | 38 (33%) | 121 (35%) |
| Female | 147 (64%) | 77 (67%) | 224 (65%) |
| Age – Mean (SD) years | 61 (10) | 60 (10) | 60 (10) |
| Age Group – N (%) | | | |
| < 65 years | 140 (61%) | 71 (62%) | 211 (61%) |
| >= 65 years | 90 (39%) | 44 (38%) | 134 (39%) |
| Race Group – N (%) | | | |
| Caucasian | 61 (27%) | 30 (26%) | 91 (26%) |
| East Asian | 165 (72%) | 83 (72%) | 248 (72%) |
| Other Asian | 1 (< 1%) | 0 (0%) | 1 (< 1%) |
| Other | 3 (1%) | 2 (2%) | 5 (1%) |
| Region – N (%) | | | |
| Asia | 160 (70%) | 83 (72%) | 243 (70%) |
| North America | 2 (1%) | 0 (0%) | 2 (< 1%) |
| Other | 68 (30%) | 32 (28%) | 100 (29%) |
| ECOG at screening – N (%) | | | |
| 0 | 92 (40%) | 42 (37%) | 134 (39%) |
| 1 | 138 (60%) | 73 (63%) | 211 (61%) |

### 3.2.4 Results and Conclusions

Table 5 shows the results for primary endpoint in the randomized population. It was provided by the applicant, but the reviewer reproduced the values using the source data files. Unless otherwise noted, the reader should assume that the results in tables included from the NDA submission were reproduced by the reviewer from submitted data files. As discussed previously, the applicant used slightly incorrect code for the stratified log-rank test. The correct p-value is .0003. The other values in the table were not affected. Unless otherwise indicated, the reader may assume that the two test methods yielded identical results within the precision shown.

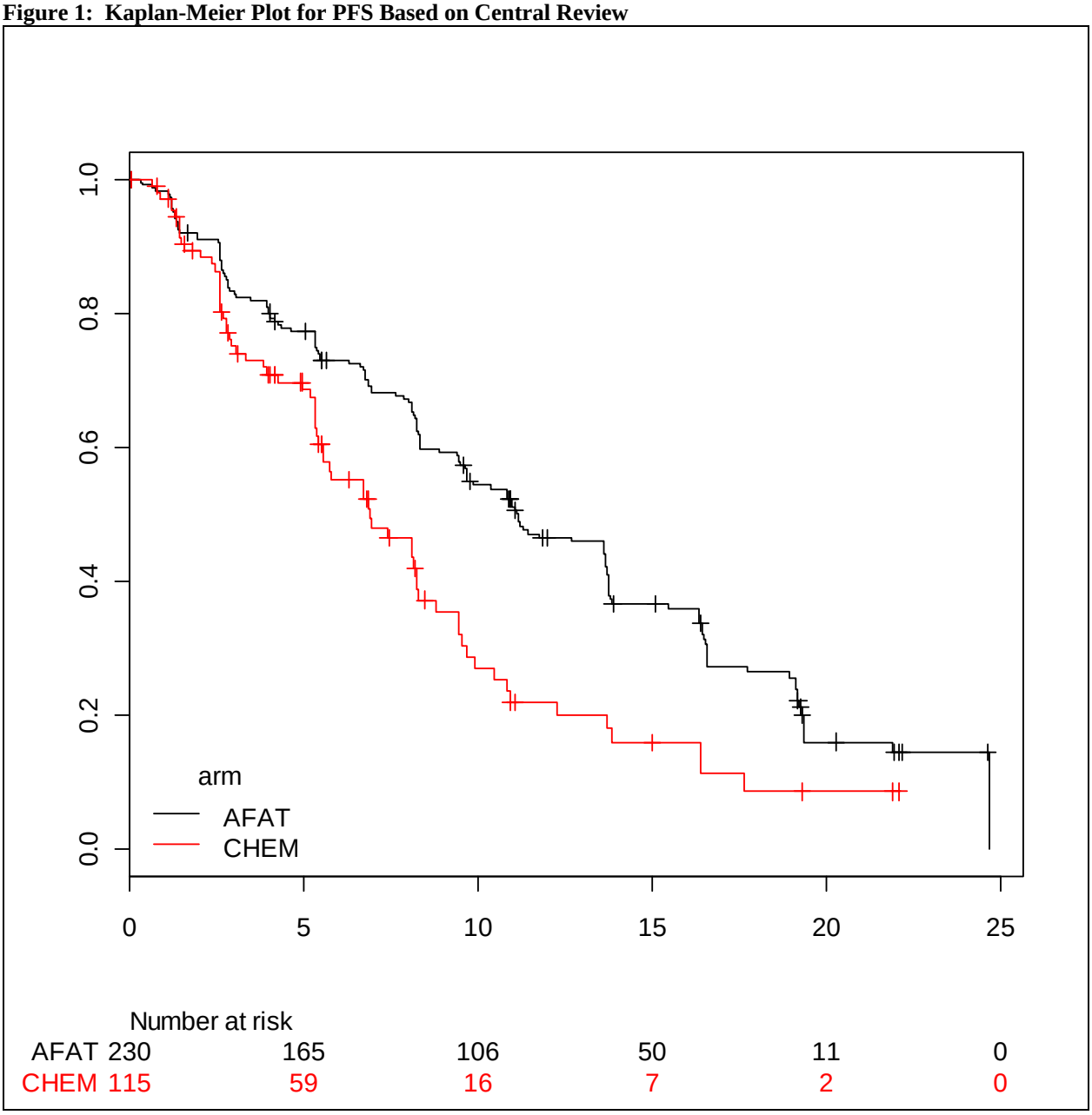
**Table 5: Primary Endpoint — PFS Based on Central Review**
**(Source: Table 11.4.1.1.1:1, Clinical Study Report)**

| | Afatinib | Chemotherapy |
|---|---|---|
| Patients [N (%)] | 230 (100.0) | 115 (100.0) |
| Patients with PFS event [N (%)] | 152 (66.1) | 69 (60.0) |
| PFS time [months] | | |
| 25th percentile (95% CI) | 5.32 (3.98, 6.87) | 3.06 (2.56, 5.32) |
| Median (95% CI) | 11.14 (9.63, 13.63) | 6.90 (5.39, 8.25) |
| 75th percentile (95% CI) | 19.12 (16.49, 19.35) | 10.84 (8.77, 16.39) |
| Hazard ratio vs. chemotherapy[1] | 0.577 | |
| 95% CI | (0.425, 0.784) | |
| p-value (2-sided)[2] | 0.0004 | |

Abbreviations: CI = confidence interval.
[1] Hazard ratio derived from a Cox proportional hazard model stratified by EGFR mutation category and race.
[2] Derived from a log-rank test stratified by EGFR mutation category and race.

Figure 1 shows the Kaplan-Meier plot for the primary endpoint.

**Figure 1: Kaplan-Meier Plot for PFS Based on Central Review**



The applicant also conducted the sensitivity analyses described in Section 3.2.2. Table 6 shows the results for PFS as determined by the investigator. The stated p-value of < 0.0001 applied to both the linear rank and log-rank tests. Note that the summary statistics are similar to those from the primary analysis.

**Table 6: PFS Based on Investigator Assessment**
(Source: Table 11.4.1.1.2:1, Clinical Study Report)

| | Afatinib | Chemotherapy |
|---|---|---|
| Patients [N (%)] | 230 (100.0) | 115 (100.0) |
| Patients with PFS event [N (%)] | 155 (67.4) | 83 (72.2) |
| PFS time [months] | | |
| Median (95% CI) | 11.07 (9.66, 13.60) | 6.70 (5.42, 8.11) |
| Hazard ratio vs. chemotherapy[1] | 0.488 | |
| 95% CI | (0.367, 0.649) | |
| p-value (2-sided)[2] | <0.0001 | |

Abbreviations: CI = confidence interval.
[1] Hazard ratio derived from a Cox proportional hazard model stratified by EGFR mutation category and race.
[2] Derived from a log-rank test stratified by EGFR mutation category and race.

Table 7 shows the degree of concordance between the two methods of assessing PFS.

**Table 7: Concordance between investigator and central assessment**
(Source: Table 11.4.1.1.2:2, Clinical Study Report)

| | | PFS event based on investigator assessment | | | |
|---|---|---|---|---|---|
| | | Afatinib N (%) | | Chemotherapy N (%) | |
| | | No | Yes | No | Yes |
| PFS event based on central independent review | No | 54 (23.5) | 24 (10.4) | 23 (20.0) | 23 (20.0) |
| | Yes | 21 (9.1) | 131 (57.0) | 9 (7.8) | 60 (52.2) |

The denominator used for this table is the total of patients randomised to the respective treatment arm, i.e., 230 patients for the afatinib arm and 115 patients for the chemotherapy arm.

Another planned sensitivity analysis for PFS was the so-called “extreme worst case” analysis in which disagreements between the investigator and the central assessment process were resolved by choosing the earlier date in the afatinib arm and the later date in the chemotherapy arm. This analysis showed a trend favoring afatinib with estimated HR = 0.85 (95% CI = [0.63, 1.16]) and nominal p = 0.31. The “symmetric worst case”, in which disagreements were resolved by choosing the earlier progression date, also significantly favored afatinib with nominal p < 0.0001. The estimated HR was 0.51 (95% CI = [0.39, 0.67]). Finally, the applicant conducted an analysis in which progression was adjudicated to occur at the next scheduled assessment if the investigator detected progression and the central assessment did not. This analysis also favored afatinib with nominal p < 0.0001 and HR = 0.52 (95% CI = [0.40, 0.68]).

Figures 2 and 3, which were provided by the applicant, show the timing of the imaging assessments for the afatinib and chemotherapy arms, respectively.

**Figure 2: Timing of imaging assessments for patients randomized to afatinib**
(Source: Figure 11.4.1.1.2:1, Clinical Study Report)

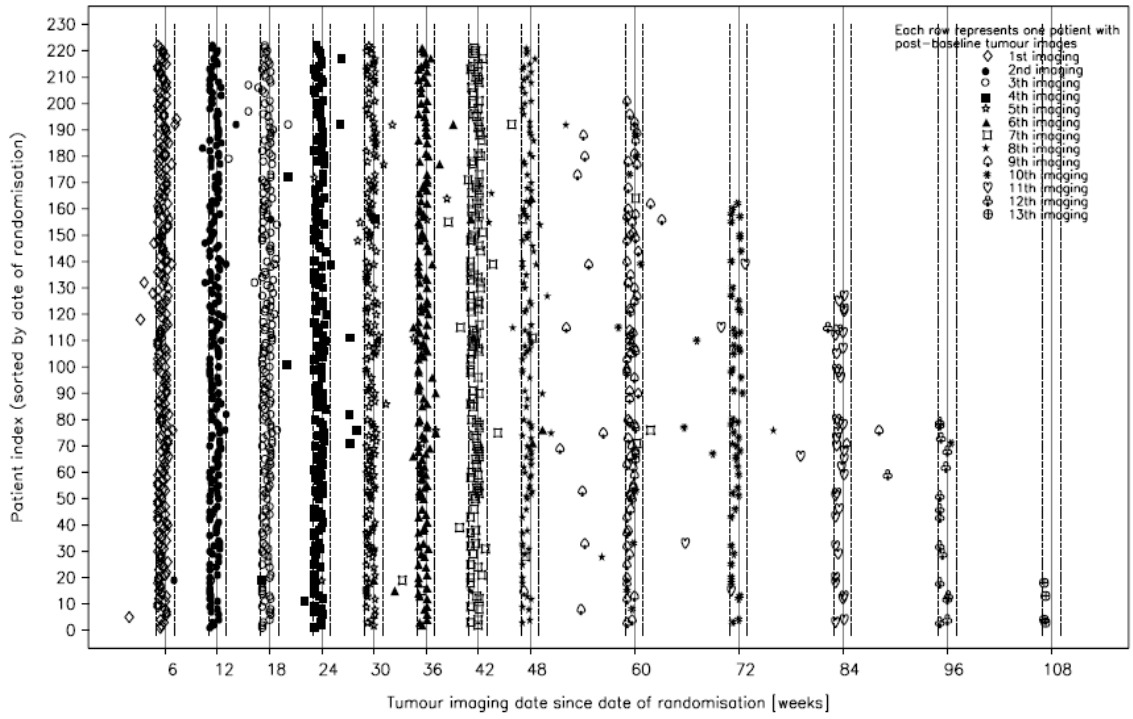
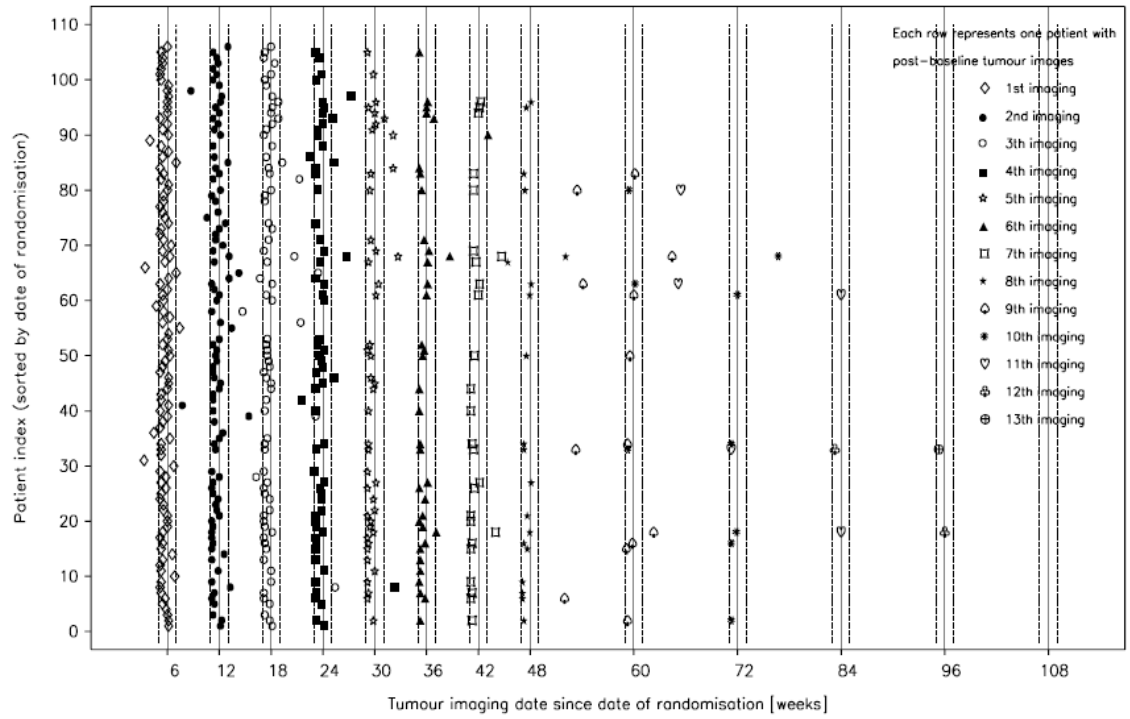



**Figure 3: Timing of assessments for patients randomized to chemotherapy**
(Source: Figure 11.4.1.1.2:2, Clinical Study Report)



Reviewer's Comments:

- Figures 2 and 3 show that the timing of the assessments was well-controlled. There does not appear to be a systematic difference between the arms. Furthermore, the median PFS for the two treatment arms differed by approximately 3 scans (48 weeks vs. 29 weeks). These findings suggest that the apparent superiority of afatinib over control on PFS is not an artifact of the timing of the scans.
- The reported sensitivity results and the level of agreement between the investigator and central assessments also suggest a robust finding of superiority on PFS. The fact that the "extreme worst case" analysis did not show a significant difference between the treatment arms is to be expected given the degree to which the investigational treatment was disfavored by the analysis method.

Table 8 shows the results of the planned interim OS analysis that was conducted at the time of the primary analysis. Since the final OS analysis is scheduled to take place after 209 deaths, this interim analysis was based on 47% information. The reviewer confirmed the results in the table, but the also conducted a stratified log-rank test which yielded p=0.6045.

**Table 8: OS at Time of Primary Analysis**
(Source: Table 11.4.1.2.2:1, Clinical Study Report)

| | Afatinib | Chemotherapy |
|---|---|---|
| Patients [N (%)] | 230 (100.0) | 115 (100.0) |
| Deaths [N (%)] | 67 (29.1) | 31 (27.0) |
| Survival time [months] | | |
| 25th percentile (95% CI) | 16.23 (13.24, 17.94) | 14.82 (13.04, 21.62) |
| Median (95% CI) | NE (22.64, NE) | NE (21.62, NE) |
| 75th percentile (95% CI) | NE (NE, NE) | NE (NE, NE) |
| Hazard ratio vs. chemotherapy[1] | 1.121 | |
| 95% CI | (0.727, 1.728) | |
| p-value (2-sided)[2] | 0.6046 | |

Abbreviations: CI = confidence interval; NE = not estimable.
[1] Hazard ratio derived from a Cox proportional hazard model stratified by EGFR mutation category and race.
[2] Derived from a log-rank test stratified by EGFR mutation category and race.

After the NDA was submitted, DOP2 requested updated survival data which the applicant submitted on January 28, 2013. Table 9 shows the updated OS results. The stratified log-rank test yielded p=0.5456. These results are based on 84% information.

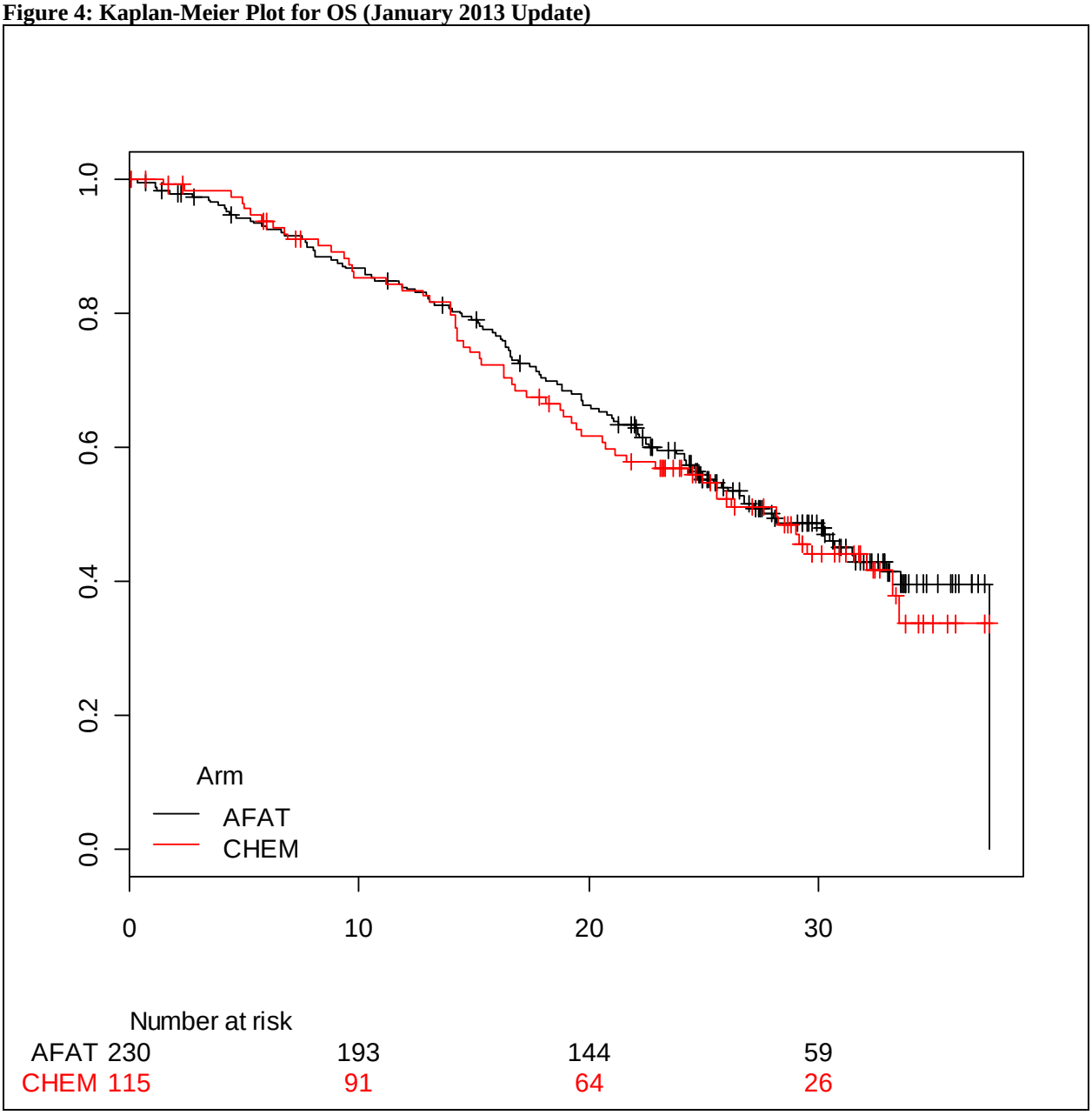
**Table 9: OS Results at January 2013 Update**
(Table 15.2.3.1:1 in January 2013 Update)

| | Afatinib 40 | Pe500+Cis75 |
|---|---|---|
| Total randomised [N(%)] | 230 ( 100.0) | 115 ( 100.0) |
| Patients died [N(%)] | 116 ( 50.4) | 59 ( 51.3) |
| Survival time [months] | | |
| 25th percentile (95% CI) | 16.39 (13.93, 18.60) | 14.82 (12.75, 18.14) |
| Median (95% CI) | 28.06 (24.64, 32.95) | 28.16 (20.73, 33.22) |
| 75th percentile (95% CI) | 37.45 ( NE, NE) | NE (33.51, NE) |
| Afatinib 40 vs. Pe500+Cis75: | | |
| Hazard ratio* | 0.907 | |
| (95% CI) | (0.660, 1.246) | |
| p-value# | 0.5457 | |

* Hazard ratio from Cox proportional hazard model stratified by EGFR mutation group and race
\# P-value from log-rank stratified by EFGR mutation group and race (two-sided)
NE = not estimable

Figure 4 shows the Kaplan-Meier plot from the updated survival data.

**Figure 4: Kaplan-Meier Plot for OS (January 2013 Update)**




As discussed in Section 3.1, DOP2 requested further follow-up data on 33 patients who were censored prior to November 1, 2012. The applicant submitted updated results on February 19, 2013. At that point, there were no additional deaths, but 10 patients had updated censoring dates. Incorporating this new data, the median survival was 28.02 months (95% CI = [24.61, 32.92]) in the afatinib arm and 28.12 months (95% CI = [20.70, 33.18]) in the control arm. The updated HR estimate was 0.900 (95% CI = [0.655, 1.237]) and the stratified log-rank test yielded a p-value of 0.5163.

Reviewer's Comment: The data from January will be taken as the key OS data in this review.

The proportions of subjects with OR, according to central assessment, were 56.1% (129/230) in the afatinib arm and 22.6% (26/115) in the chemotherapy arm. This difference was significant with $p < 0.0001$. The odds ratio was 4.66 in favor of afatinib, with a 95% CI of [2.77, 7.83].

Review's Comment: The clinical benefit of disease control is unclear, so that endpoint will not be discussed in this review.

### 3.3 Evaluation of Safety

The safety review was primarily conducted by Shakun Malik, MD. At Dr. Malik's request, the statistical reviewer reproduced the applicant's summary of adverse reactions in Study 1200.32 and also combined the rash and acne categories. See Dr. Malik's review for further details.

## 4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

### 4.1 Gender, Race, Age, and Geographic Region

Table 10 shows PFS and OS for Study 1200.32 in the subgroups of gender, race, and age. The analyses were unstratified.

**Table 10: PFS and OS by Demographic Subgroups**

| **Subgroup (N)** | **PFS HR (95% CI)** | **OS HR (95% CI)** |
|---|---|---|
| **Gender** | | |
| Female (224) | 0.54 (0.38, 0.78) | 0.92 (0.61, 1.39) |
| Male (121) | 0.61 (0.37, 1.01) | 0.94 (0.58, 1.54) |
| **Age** | | |
| >= 65 years (211) | 0.64 (0.39, 1.03) | 0.94 (0.57, 1.57) |
| < 65 years (134) | 0.53 (0.36, 0.76) | 0.92 (0.62, 1.37) |
| **Race** | | |
| Asian (249) | 0.54 (0.38, 0.78) | 1.00 (0.68, 1.46) |
| Non-Asian (96) | 0.68 (0.39, 1.19) | 0.80 (0.46, 1.40) |

The reviewer did not conduct a subgroup analysis by geographic region because it was strongly associated with race. Only two patients were in North America.

Reviewer's Comment: These subgroup analyses should be taken as exploratory.

### 4.2 In Other Special/Subgroup Populations

The type of EGFR mutation was a stratification factor in Study 1200.32 and is relevant to labeling. Table 11 shows the reviewer's results for PFS, with patients grouped according to the EGFR stratification factor. Patients with a "Common" mutation, Del19 or L858R, are also shown as a group. These results are based on a patient's actual EGFR mutation in cases where the

wrong stratum was used at randomization. The analyses are not stratified by race. Note that there is a non-significant trend (unadjusted p = 0.11) for patients with "Other" mutations (besides Del19 or L858R) to have shorter PFS on afatinib compared to chemotherapy. Table 12 shows the corresponding results for OS. Note that in the Other subgroup the patients on afatinib had worse results, with an unadjusted p= 0.03. These results are discussed further in Section 5.1.

**Table 11: Progression Free Survival by EGFR Subgroup**

| | **Afatinib** | **Pemetrexed/Cisplatin** |
|---|---|---|
| **Patients with Common Mutation (Del19 or L858R)** | | |
| Number of Death or Progression, N (%) | 130/204 (64%) | 61/104 (59%) |
| Median (months) | 13.6 | 6.9 |
| 95% CI | (10.8, 13.8) | (5.4, 8.2) |
| HR (95% CI) | 0.47 (0.34, 0.65) | |
| Log-Rank Test P-value* | < 0.0001 | |
| **Patients with Del19 Mutation** | | |
| Number of Death or Progression, N (%) | 67/113 (59%) | 35/57 (61%) |
| Median (months) | 13.7 | 5.6 |
| 95% CI | (11.1, 16.4) | (3.1, 8.1) |
| HR (95% CI) | 0.28 (0.18, 0.44) | |
| Log-Rank Test P-value* | < 0.0001 | |
| **Patients with L858R Mutation** | | |
| Number of Death or Progression, N (%) | 63/91 (69%) | 26/47 (55%) |
| Median (months) | 10.8 | 8.1 |
| 95% CI | (8.2, 13.8) | (5.7, 9.7) |
| HR (95% CI) | 0.73 (0.46, 1.16) | |
| Log-Rank Test P-value* | 0.18 | |
| **Patients with Other Mutation (Not Del19 or L858R)** | | |
| Number of Death or Progression, N (%) | 22/26 (85%) | 8/11 (73%) |
| Median (months) | 2.8 | 9.9 |
| 95% CI | (2.6, 6.7) | (3.8, 13.8) |
| HR (95% CI) | 1.89 (0.84, 4.28) | |
| Log-Rank Test P-value* | 0.11 | |

*Based on unstratified log-rank test, without control for multiplicity.

**Table 12: Overall Survival by EGFR subtype (January 2013 Update)**

| | Afatinib | Pemetrexed/Cisplatin |
|---|---|---|
| **Patients with Common Mutation (Del19 or L858R)** | | |
| Number Died, N (%) | 97/204 (48%) | 55/104 (53%) |
| Median (months) | 30.3 | 26.2 |
| 95% CI | (26.1, 37.5) | (20.6, 33.2) |
| HR (95% CI) | 0.81 (0.58, 1.13) | |
| Log-Rank Test P-value* | 0.22 | |
| **Patients with Del19 Mutation** | | |
| Number Died, N (%) | 51/113 (45%) | 36/57 (63%) |
| Median (months) | 31.6 | 21.1 |
| 95% CI | (26.7, 37.5) | (16.3, 29.1) |
| HR (95% CI) | 0.55 (0.36, 0.85) | |
| Log-Rank Test P-value* | 0.006 | |
| **Patients with L858R Mutation** | | |
| Number Died, N (%) | 46/91 (51%) | 19/47 (40%) |
| Median (months) | 27.2 | NE |
| 95% CI | (19.8, NE) | (13.0, NE) |
| HR (95% CI) | 1.30 (0.76, 2.22) | |
| Log-Rank Test P-value* | 0.33 | |
| **Patients with Other Mutation (Not Del19 or L858R)** | | |
| Number Died, N (%) | 19/26 (73%) | 4/11 (36%) |
| Median (months) | 15.9 | NE |
| 95% CI | (7.5, 24.6) | (6.8, NE) |
| HR (95% CI) | 3.08 (1.04, 9.15) | |
| Log-Rank Test P-value* | 0.03 | |

*Based on unstratified log-rank test, without control for multiplicity. NE = Not Estimable.

# 5 SUMMARY AND CONCLUSIONS

## 5.1 Statistical Issues

The statistical interpretation of the primary endpoint, PFS assessed by blinded central review, is relatively straightforward. The magnitude of the treatment effect, about four months at the median, and the close timing of the imaging assessments suggest that the apparent superiority of afatinib is unlikely to be an artifact of scheduling. There was also an acceptable level of agreement between the investigator and central review. The planned sensitivity analyses also support a positive finding. In summary, the data demonstrate that afatinib was superior to chemotherapy for PFS at the standard significance level in the randomized population.

In contrast, there is no statistical evidence of a treatment effect on OS in the randomized population. Although the survival data are immature, the trend as of the January 2013 update slightly favors afatinib (HR = 0.907, CI = [0.660, 1.246]). The lack of a difference between the afatinib and chemotherapy arms for OS may be explainable by differences in post-progression treatment. In particular, patients randomized to the chemotherapy arm were more likely to take an EGFR TKI-containing medication after progression than those on the afatinib arm. However, the causal effect of such non-randomized treatment is inherently difficult to assess.

The randomization was stratified by genetic subgroup, with three levels: Del19, L858R, and Other. A subgroup analysis was planned for this factor, but there was no control for multiplicity. For OS, there was a nominally-significant finding ($p = 0.03$) that afatinib was worse than chemotherapy in the Other subgroup which consisted of 37 patients and had 23 deaths. There was also a non-significant adverse finding ($p = 0.11$) for PFS in this subgroup. The lack of multiplicity control makes these exploratory analyses difficult to interpret, however. If one uses a post-hoc Bonferroni correction for the three subgroups, then the OS result is no longer significant. Furthermore, this correction may be inadequate in light of the fact that 13 other exploratory subgroup analyses were specified in the statistical analysis plan, including one additional stratification factor (race). Finally, note that there was no treatment effect on OS in the ITT population and one could therefore argue that the subgroup analyses may not be conducted for this endpoint.

There are at least two ways to consider the subgroup results within the realm of statistics. One way is to strictly control type-I error at the .05 (or even lower) level. In this case, the adverse finding for OS in the Other subgroup should clearly be disregarded. A more subtle approach considers the consequences of making an error in either direction and incorporates external evidence. See Section 5.2 for further discussion.

## 5.2 Collective Evidence

The results from Study 1200.32 show that there was an anti-tumor (PFS) effect in the randomized population. The study did not show an effect on OS in the randomized population.

Interpretation of the results from the EGFR subgroups is difficult. Since there was no pre-specified plan to correct for multiple comparisons, the subgroup analyses would usually be considered exploratory. Even within this context, however, the finding of worse OS (unadjusted p = 0.03) on afatinib for patients in the Other subgroup should not be taken lightly as it potentially represents excess mortality.

It is worth considering that if one deems afatinib to have a positive benefit-risk profile, then one must necessarily discount the lack of an effect on OS in the overall study population. A recommendation to approve afatinib with Other patients excluded from the indication would arguably be self-contradictory: it may not make sense to ***distrust*** the OS results in the study population of 345 patients but ***trust*** the same endpoint in a subgroup of 37 patients. Although PFS also showed an unfavorable trend in the Other subgroup, the p-value was non-significant. Another point to consider is that the Other subgroup is not homogeneous. It is quite plausible that some patients within this group have mutations favorable to treatment with afatinib and others do not.

In conclusion, if PFS is deemed an acceptable primary endpoint then the study supports a finding of efficacy in the randomized population. While there is some evidence of harm in patients in the Other EGFR subgroup, it is not statistically persuasive. Given sufficient external evidence, it may be still be appropriate to restrict the indication. It is also appropriate to consider the availability of alternative treatments for patients with a mutation in the Other subgroup.

### 5.3 Conclusions and Recommendations

The submitted study in first-line patients (1200.32) demonstrated that patients had a statistically significant improvement in PFS when treated with afatinib compared to those treated with chemotherapy. The study did not show a statistically-significant difference in the overall survival between the two treatment arms with 84% information relative to the 209 deaths required for the final OS analysis. Whether the results from Study 1200.32 provide evidence for a favorable benefit-risk ratio and support approval of afatinib will be determined by the clinical review team. The clinical team will also need to carefully weigh the evidence to determine the appropriate indication.

---

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

---

/s/

---

JONATHAN D NORTON
04/19/2013

KUN HE
04/19/2013
Accepted as a complete review

RAJESHWARI SRIDHARA
04/22/2013

# STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

**NDA Number:** 201-292 **Applicant:** Boehringer Ingelheim **Stamp Date:** 11-15-2012

**Drug Name:** Afatinib

On **initial** overview of the NDA/BLA application for RTF:

| | Content Parameter | Yes | No | NA | Comments |
|---|---|---|---|---|---|
| 1 | Index is sufficient to locate necessary reports, tables, data, etc. | X | | | |
| 2 | ISS, ISE, and complete study reports are available (including original protocols, subsequent amendments, etc.) | X | | | |
| 3 | Safety and efficacy were investigated for gender, racial, and geriatric subgroups investigated (if applicable). | X | | | |
| 4 | Data sets in EDR are accessible and do they conform to applicable guidances (e.g., existence of define.pdf file for data sets). | X | | | |

Note: Checked pivotal study (1200.32), which will be focus of review.

**IS THE STATISTICAL SECTION OF THE APPLICATION FILEABLE?** Yes, but defer to clinical team on issues with safety data.

If the NDA/BLA is not fileable from the statistical perspective, state the reasons and provide comments to be sent to the Applicant.

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

| Content Parameter (possible review concerns for 74-day letter) | Yes | No | NA | Comment |
|---|---|---|---|---|
| Designs utilized are appropriate for the indications requested. | X | | | Primary endpoint is PFS |
| Endpoints and methods of analysis are specified in the protocols/statistical analysis plans. | X | | | |
| Interim analyses (if present) were pre-specified in the protocol and appropriate adjustments in significance level made. DSMB meeting minutes and data are available. | | | X | No interim analysis for PFS |
| Appropriate references for novel statistical methodology (if present) are included. | | | X | |
| Safety data organized to permit analyses across clinical trials in the NDA/BLA. | | | X | Defer to clinical team |
| Investigation of effect of dropouts on statistical analyses as described by applicant appears adequate. | X | | | Will conduct own analyses |

# STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

**Additional requests for the 74-day letter (non-filing issues)**

For study 1200.32, submit the individual results from the radiologists and oncologist comprising the independent review panel.

For study 1200.32, submit any macros and formats needed to run the submitted SAS code.

---

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

---

/s/

---

JONATHAN D NORTON
12/14/2012

KUN HE
12/14/2012