Outcomes of Phase II Clinical Trials with Single-Agent Therapies in Advanced/Metastatic Non–Small Cell Lung Cancer Published between 2000 and 2009

Filip Janku1, Donald A. Berry2, Jing Gong1, Henrique A. Parsons1, David J. Stewart3, and Razelle Kurzrock1

Abstract

Purpose: We analyzed the outcomes of single-agent phase II clinical trials in non–small cell lung cancer (NSCLC) to determine trial parameters that predicted clinical activity.

Experimental Design: Data on response rate (RR), progression-free survival (PFS), and overall survival (OS) from all English language, single-agent phase II trials in advanced/metastatic NSCLC indexed by PubMed (January 2000 through December 2009) were abstracted.

Results: A total of 143 single-agent phase II trials (7,701 patients) were identified. The median RR was 10%, PFS 2.8 months, and OS 7.6 months. RR and PFS correlated with OS ($r = 0.46, P < 0.001, r = 0.52$, $P < 0.001$, respectively) and RR correlated with PFS ($r = 0.61, P < 0.001$). Treatment arms enriched for patients with molecular targets had a higher median RR (48.8% vs. 9.7%, $P = 0.005$), longer median PFS (6 vs. 2.8 months, $P = 0.005$), and OS (11.3 vs. 7.5 months, $P = 0.05$) as compared with those of unselected patients. In multivariate analysis, only studies enriched for patients with molecular targets or including drugs that eventually gained FDA/EMA approval were associated with a higher RR, and longer PFS/OS.

Conclusions: In phase II trials in NSCLC, RR and PFS correlated with OS. Studies enriched for patients with putative molecular drug targets were associated with higher therapeutic benefit as compared with those of unselected populations. Clin Cancer Res; 18(22); 6356–63. ©2012 AACR.

Introduction

There is a well-established pathway of anticancer drug development that progresses from dose-finding phase I studies through phase II efficacy trials, to phase III randomized trials to determine benefit compared with accepted standard treatment regimens (1). Phase II cancer clinical trials are designed to identify the spectrum of antitumor activity of new therapies using the dose and schedule identified in phase I trials (2). Therapies deemed to have promising activity advance to phase III testing in which they are typically compared with standard-of-care. However, only a minority of new anticancer therapies tested in early-phase trials eventually meet the standards for marketing approval by the U.S. Food and Drug Administration (FDA) or European Medicines Agency (EMA; refs. 3, 4). Indeed, phase III oncology trials are successful only approximately 34% of the time (5).

Non–small cell lung cancer (NSCLC) accounts for 85% of all lung cancers (6). NSCLC is often diagnosed at an advanced stage and has a poor prognosis. Here, we report an analysis of the outcomes of single-agent phase II studies of NSCLC published within the past decade to delineate the specific variables associated with increased antitumor activity.

Materials and Methods

Identification of phase II studies and data abstraction

We searched PubMed for all single-agent phase II trials published throughout a 10-year period from January 2000 to December 2009. Search included terms “phase II” and “cancer” or “carcinoma” and was limited to articles in English. From that broad search, we included all phase II studies in advanced/metastatic NSCLC with single-agent therapies.

The following data were extracted from each study: number of treatment arms, type of treatment, number of treating centers, trial design, enrichment for molecular targets (as defined by testing for the molecular target directly or enrichment for patient populations expected to have a high rate of the target), previous systemic therapies, primary endpoint, the year of publication, name of the journal, and its 5-year impact factor (2008 ISI Journal of Citation Reports).
Translational Relevance

Phase II studies in non–small cell lung cancer (NSCLC) and other cancers are designed to test the efficacy of cancer therapies to identify compounds with promising activity that should be considered for testing in phase III randomized trials. Despite phase II trials being used frequently in oncology, their ability to delineate promising compounds for phase III trials is not straightforward as only approximately 34% of phase III oncology trials are successful. In this study, we attempted to identify specific variables associated with better outcomes of single-agent phase II studies in NSCLC.

Statistical analysis

RR was defined as a partial or complete response. Medians for RR, PFS, and OS were calculated. Statistical dispersion was measured by the interquartile range (IQR, the distance between the 75th and the 25th percentile). Assessment of independent samples was done using non-parametric tests, such as the Wilcoxon rank-sum test and Kruskal–Wallis one-way analysis of variance by ranks. Multivariate linear regression models were fit to assess the association between several variables, RR, PFS, and OS, in which the variables included type of treatment (chemotherapy vs. targeted), year of publication (RR vs. other), the journal’s 5-year impact factor (2000–2004 vs. 2005–2009), the primary endpoint (RR vs. OS), if available. If the RR was based on number of evaluable patients, then the RR was recalculated using the intent-to-treat principle. All data were abstracted per treatment arm.

Results

Clinical trials designs and treatments

A total of 143 single-agent phase II studies in advanced or metastatic NSCLC published between January 2000 and December 2009 were identified in the PubMed database (Supplementary Appendix S1 and Table S1). These 143 studies had 163 treatment arms and enrolled a total of 7,701 patients. The median number of patients per treatment arm was 40 (range, 6–203). The details about abstracted trials are listed in Table 1.

Enrichment for molecular targets is associated with better treatment outcomes

Response rate. Of the 143 trials with 163 treatment arms, the median RR was 10% (IQR, 15.8). A RR of 0% was reported in 23 (14%) treatment arms (Fig. 1).

Twelve studies, which were enriched for the presence of molecular targets, showed a higher median RR of 48.8% (IQR, 71) as compared with 9.7% (IQR, 13.6) in the remaining 131 studies with unselected patients ($P = 0.005$; Tables 1 and 2). Figures 2 and 3 show the reported RR in trials with enriched and unselected patient populations with respect to number of patients per treatment arm (Fig. 2A), year of publication (Fig. 2B), PFS (Fig. 3A), and OS (Fig. 3B).

Because 10 of 14 arms with enriched populations used gefitinib, we compared those 10 gefitinib arms (476 patients) with 12 gefitinib arms with unselected patients ($n = 764$). Gefitinib arms with enriched populations showed a higher median RR of 58% (IQR, 41.8) as compared with 18% (IQR, 29.7) in the gefitinib arms with unselected patients ($P = 0.001$).

Treatment arms ($n = 27$) with drugs that eventually (at the time of analysis) gained FDA/EMA approval for advanced or metastatic NSCLC showed a higher median RR of 26.5% (IQR, 41.9) as compared with 9% (IQR, 13.1) in 136 treatment arms for other drugs (approved previously in NSCLC, approved previously for other indications, approved at the time of analysis for other indications, or not approved; $P < 0.001$; Table 1).

There were no differences in median RR in studies with chemotherapy versus targeted therapy; single institution versus multicenter studies; randomized versus nonrandomized studies; studies with treatment-naïve patients versus previously treated with a systemic therapy; studies published from 2000 to 2004 versus from 2005 to 2009; studies published in journals with an impact factor $\geq 5$ versus $<5$. There were also no differences in the median RR in studies with different endpoints (Table 1).

In multivariate analysis, only studies enriched for patients with molecular targets ($P < 0.001$) and studies with drugs that eventually (at the time of analysis) gained FDA/EMA approval ($P < 0.001$) were associated with a higher RR.

Progression-free survival. Of the 143 trials included in our analysis, median PFS was reported in 127 (78%) of 163 treatment arms. In these 127 treatment arms, the median PFS was 2.8 months (IQR, 2.3).

The studies that were enriched for the presence of putative molecular targets had a longer median PFS of 6 months (IQR, 6.8) as compared with 2.8 months (IQR, 1.9) in studies with unselected patients ($P = 0.005$).

In addition, gefitinib arms with enriched populations showed a longer median PFS of 7.7 months (IQR, 4.8) as...
compared with 2.8 months (IQR, 2.7) in the gefitinib arms with unselected patients ($P = 0.003$; Table 1).

Treatment arms with drugs that eventually (at the time of analysis) gained FDA and/or EMA approval for advanced or metastatic NSCLC also showed a longer median PFS of 3.7 months (IQR, 5.2) as compared with 2.8 months (IQR, 1.9) in other trials (approved previously for NSCLC, approved previously for other indications, approved at the time of analysis for other indications, or not approved; $P < 0.001$).

Treatment arms enrolling treatment-naive patients showed a prolonged median PFS of 3.5 months (IQR, 2.2) as compared with 2.7 months (IQR, 1.7) in trials enrolling patients previously treated with a systemic therapy ($P = 0.004$).

There were no differences in the median PFS among studies with chemotherapy versus targeted therapy; single institution versus multicenter studies; randomized versus nonrandomized studies; studies published from 2000 to 2004 versus from 2005 to 2009; studies published journals with impact factors $\geq 5$ versus <5. There were no differences in the median PFS in studies with different endpoints (Table 1).

In multivariate analysis, the only factors that were selected as independent predictors of longer PFS were, enrichment for patients with putative molecular targets ($P = 0.001$), studies with drugs that eventually (at the time of analysis) gained FDA/EMA approval ($P = 0.05$), and journal impact factor of less than 5 ($P = 0.01$).

Overall survival. Of the 143 trials, the median OS was reported in 148 (91%) of 163 treatment arms. In these 148 treatment arms, the median OS was 7.6 months (IQR, 3.6).

### Table 1. Characteristics of phase II studies with single agents in NSCLC and associations with outcomes

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Studies N (%)</th>
<th>Arms N (%)</th>
<th>Patients N (%)</th>
<th>Median RR % (IQR) P</th>
<th>Median PFS months (IQR) P</th>
<th>Median OS months (IQR) P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total studies</td>
<td>143 (100)</td>
<td>163 (100)</td>
<td>7,701 (100)</td>
<td></td>
<td>2.8 (2.3)</td>
<td>7.6 (3.6)</td>
</tr>
<tr>
<td>Agent</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>91 (64)</td>
<td>105 (64)</td>
<td>4,817 (61)</td>
<td>11 (14.4)</td>
<td>0.65</td>
<td>3 (2)</td>
</tr>
<tr>
<td>Targeted therapy</td>
<td>52 (36)</td>
<td>58 (36)</td>
<td>2,884 (37)</td>
<td>8.2 (23.4)</td>
<td>2.7 (3.4)</td>
<td></td>
</tr>
<tr>
<td>Number of centers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single-center</td>
<td>44 (31)</td>
<td>48 (29)</td>
<td>1,833 (24)</td>
<td>12.5 (22.5)</td>
<td>0.08</td>
<td>3.3 (2.4)</td>
</tr>
<tr>
<td>Multicenter</td>
<td>99 (69)</td>
<td>115 (71)</td>
<td>5,868 (76)</td>
<td>9 (13.2)</td>
<td>2.8 (2)</td>
<td></td>
</tr>
<tr>
<td>Planned design</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Randomized</td>
<td>31 (22)</td>
<td>49 (30)</td>
<td>2,812 (37)</td>
<td>10.9 (12.3)</td>
<td>0.94</td>
<td>2.8 (2.3)</td>
</tr>
<tr>
<td>Nonrandomized</td>
<td>112 (78)</td>
<td>114 (70)</td>
<td>4,889 (63)</td>
<td>9.9 (20.6)</td>
<td>2.8 (2.2)</td>
<td></td>
</tr>
<tr>
<td>Enrichment for molecular targets</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enriched</td>
<td>12 (8)</td>
<td>14 (9)</td>
<td>666 (9)</td>
<td>48.8 (71)</td>
<td>0.005&lt;sup&gt;a&lt;/sup&gt;</td>
<td>6 (6.8)</td>
</tr>
<tr>
<td>Not enriched</td>
<td>131 (92)</td>
<td>149 (91)</td>
<td>7,035 (91)</td>
<td>9.7 (13.6)</td>
<td>2.8 (1.9)</td>
<td></td>
</tr>
<tr>
<td>Prior therapies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment-naive</td>
<td>60 (42)</td>
<td>63 (39)</td>
<td>2,819 (37)</td>
<td>12.5 (22.2)</td>
<td>0.12</td>
<td>3.5 (2.2)</td>
</tr>
<tr>
<td>Pretreated</td>
<td>83 (58)</td>
<td>100 (61)</td>
<td>4,882 (63)</td>
<td>9.3 (13.1)</td>
<td>2.7 (1.7)</td>
<td></td>
</tr>
<tr>
<td>Endpoint</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RR</td>
<td>115 (80)</td>
<td>127 (78)</td>
<td>5,468 (71)</td>
<td>11 (17.5)</td>
<td>0.47</td>
<td>2.9 (2.3)</td>
</tr>
<tr>
<td>PFS or OS</td>
<td>21 (15)</td>
<td>25 (15)</td>
<td>1,638 (21)</td>
<td>6 (12)</td>
<td>2.5 (2.1)</td>
<td></td>
</tr>
<tr>
<td>Other (PK, PS, symptoms, toxicity)</td>
<td>7 (5)</td>
<td>11 (7)</td>
<td>595 (8)</td>
<td>8 (9.7)</td>
<td>2.8 (1)</td>
<td></td>
</tr>
<tr>
<td>Year of publication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000–2004</td>
<td>50 (35)</td>
<td>55 (34)</td>
<td>2,513 (33)</td>
<td>12 (19)</td>
<td>0.55</td>
<td>3 (2.1)</td>
</tr>
<tr>
<td>2005–2009</td>
<td>93 (65)</td>
<td>108 (66)</td>
<td>5,188 (67)</td>
<td>9.5 (13.3)</td>
<td>2.8 (2.2)</td>
<td></td>
</tr>
<tr>
<td>Journal impact factor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\geq 5$</td>
<td>49 (34)</td>
<td>60 (37)</td>
<td>3,666 (48)</td>
<td>9 (12.7)</td>
<td>0.57</td>
<td>2.7 (2.4)</td>
</tr>
<tr>
<td>$&lt;$5</td>
<td>94 (66)</td>
<td>103 (63)</td>
<td>4,035 (52)</td>
<td>11 (18.8)</td>
<td>3 (2.1)</td>
<td></td>
</tr>
<tr>
<td>FDA/EMA approval</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>26 (18)</td>
<td>27 (17)</td>
<td>1,530 (20)</td>
<td>26.5 (41.9)</td>
<td>$&lt;0.001&lt;sup&gt;a&lt;/sup&gt;$</td>
<td>3.7 (5.2)</td>
</tr>
<tr>
<td>No</td>
<td>117 (82)</td>
<td>136 (83)</td>
<td>6,171 (80)</td>
<td>9 (13.1)</td>
<td>2.8 (1.9)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: PK, pharmacokinetics; PS, performance status.

<sup>a</sup>Results from univariate analysis confirmed in multivariate analysis.

Janku et al.

Clin Cancer Res; 18(22) November 15, 2012

Clinical Cancer Research

6358

Clin Cancer Res; 18(22) November 15, 2012

Clinical Cancer Research

Published OnlineFirst September 26, 2012; DOI: 10.1158/1078-0432.CCR-12-0178

Downloaded from clincancerres.aacrjournals.org on May 2, 2017. © 2012 American Association for Cancer Research.
Figure 1. Reported RRs from 163 treatment arms. Treatment arms with therapies that gained FDA/EMA approval at the time of analysis are depicted in green (enriched for the presence of putative molecular targets) or blue (not enriched for the presence of putative molecular targets). Treatment arms with other drugs (approved previously in NSCLC, approved previously for other indications, approved at the time of analysis for other indication, or not approved) are depicted in red (enriched for the presence of putative molecular targets) or yellow (not enriched for the presence of putative molecular targets).

compared with 7.5 months (IQR, 3.2) in studies with unselected patients (P = 0.05; Table 1).

Gefitinib arms with enriched populations showed a longer median OS of 16.4 months (IQR, 6.8) as compared with 7.6 months (IQR, 4.15) in the gefitinib arms with unselected patients (P = 0.01).

Treatment arms with drugs that eventually (at the time of analysis) gained FDA/EMA approval for advanced or metastatic NSCLC also showed a longer median OS of 9.3 months (IQR, 6.7) as compared with 7.5 months (IQR, 3.1) in other trials (approved previously for NSCLC, approved previously for other indications, approved at the time of analysis for other indications, or not approved; P = 0.002).

Treatment arms enrolling treatment-naive patients showed a prolonged median OS of 8.2 months (IQR, 3.6) as compared with 7.1 months (IQR, 2.9) in trials enrolling patients previously treated with a systemic therapy (P = 0.002).

Studies with RR as a primary endpoint showed a longer median OS than studies with other primary endpoints, such as PFS/OS, and others (7.9 months (IQR, 3.6) vs. 6.7 months (IQR, 3.7) vs. 5.8 months (IQR, 3.5); P = 0.02).

There were no differences in the median OS among studies with chemotherapy versus targeted therapy; single institution versus multicenter studies; randomized versus nonrandomized studies; studies published from 2000 to 2004 versus from 2005 to 2009; studies published in journals with impact factors 0–5 versus ≥5 (Table 1).

In multivariate analysis, the only factors that were selected as independent predictors of survival were, enrichment for patients with putative molecular targets (P = 0.005), studies with drugs that eventually (at the time of analysis) gained FDA/EMA approval (P = 0.003), being previously treatment naïve (P = 0.009), and studies with RR as a primary endpoint (P = 0.02).

**Response rate correlates with PFS and OS**

We investigated whether RR predicts survival outcomes. RR was found to positively correlate with PFS (r = 0.61, P < 0.001, Fig. 3A) and OS (r = 0.46, P < 0.001, Fig. 3B). In addition, PFS was found to positively correlate with OS (r = 0.52, P < 0.001, Fig. 3C).

**Discussion**

Phase II studies are among the most common studies in cancer clinical research. The quality of phase II data may be one of the most important factors for predicting the success of a phase III trial (7, 8). However, many phase II trials prove to be of limited value in the drug development trajectory.

In our report, the number of phase II studies with single-agent therapy published and indexed in the PubMed database nearly doubled in the 5-year period from 2005 to 2009 as compared with 2000 to 2004 (93 studies with 5,188 patients enrolled vs. 50 studies with 2,513 patients). However, this increase did not translate into improvements in outcomes as measured by RR, PFS, and OS. The actual, rather than apparent, reality might be even more disappointing as trials with poor outcomes tend to not be published (9).

The reported outcomes (RR, PFS, and OS) did not consistently differ with respect to the type of therapy, number of institutions involved, presence of randomization, trial endpoints, number of prior therapies, and impact factor of the publishing journals. However, the outcomes were improved for all parameters (RR, PFS, and OS) when trials were enriched for the presence of putative molecular drug targets. The only other factor that was consistently chosen as independent in multivariate analysis for RR, PFS, and OS, was studies with drugs that eventually gained FDA and/or EMA approval. The latter finding is, of course, not unexpected.

To improve the design and scientific use of phase II clinical trials, some authors suggest replacing single-arm phase II trials with randomized phase II trials (10). Their hypothesis is that randomized phase II trials would eliminate bias associated in comparisons with historical controls and generate robust data before proceeding to phase III trials (11, 12). Other investigators view single-agent phase II trials as a valuable platform, as well as being easily executed, and they can lead to drug approval (13). Moreover, while randomized phase II studies have been useful in gaining benefit and efficacy data for drugs whose effects are characterized by prolonged stable disease, the routine use of randomized phase II trials does not seem to improve the overall success rate in drug development (12, 14).

Our data showed better outcomes in trials enriched for putative molecular targets even though not every trial in the analysis used a validated enrichment strategy (Table 2). Single-agent phase II trials are often criticized for their propensity to produce false-positive results when compared with historical controls (11). Yet, in the last decade, well-
conducted early-phase trials in populations enriched by molecular targets to define patients likely to benefit from therapy resulted in robust outcomes (e.g., BRAF inhibitors in BRAF-mutated melanoma, etc.) so that the question arises whether in such situations randomized trials are necessary or justifiable (15–19).

For many years, clinical cancer research has been tailored to low-yield therapies producing benefit in a small proportion of patients within a large clinical trial population. In these studies, the P value is driven to statistically significant levels, mostly by virtue of the large number of study participants (20). In contrast, if the treatment is given in a selective manner only to patients with tumors expressing the target of interest, the number of patients needed to adequately power the study is likely to be much smaller even if a randomized trial is conducted (21).

Identifying relevant molecular targets is of increasing importance (22, 23). Unfortunately, it is sometimes unclear whether the putative molecular target is a real target or an incidental passenger (12). For example, initial cetuximab clinical trials in colorectal cancer were designed only for patients with EGFR expression by immunohistochemistry (IHC; refs. 24–26). However, ultimately no significant correlation between EGFR expression and outcomes was identified (27, 28). Our study also identified trials in which selected targets, such as HER1-4 expression detected through IHC, proved to be flawed (Table 2).

It is known that RR does not always correlate with PFS and OS, and there has been considerable debate about the use of RR as a surrogate marker for survival, mainly because RR can be subjective, and there is a concern that it may not reliably predict PFS and OS (29). However, in a review of randomized trials in NSCLC, a strong correlation existed between RR and OS (30). Furthermore, a survey of 31 anticancer drugs approved by the FDA on the basis of RR or PFS in the absence of a randomized trial showed that these agents fared well after long-term clinical use (13). Importantly, in this study, we also showed a positive correlation between

<table>
<thead>
<tr>
<th>Study (reference in Supplementary Appendix S1)</th>
<th>Method of enrichment</th>
<th>Drug, mechanism</th>
<th>Number of patients</th>
<th>RR (%)</th>
<th>PFS (mo)</th>
<th>OS (mo)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGFR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yang and colleagues (141)</td>
<td>Enriched for EGFR mutations (East Asians, mostly nonsmokers with adenocarcinoma)</td>
<td>Gefitinib, EGFR kinase inhibitor</td>
<td>106</td>
<td>50.9</td>
<td>5.5</td>
<td>22.4</td>
</tr>
<tr>
<td>Inoue and colleagues (59)</td>
<td>EGFR mutations</td>
<td>Gefitinib, EGFR kinase inhibitor</td>
<td>30</td>
<td>66</td>
<td>6.5</td>
<td>17.8</td>
</tr>
<tr>
<td>Sequist and colleagues (111)</td>
<td>EGFR mutations</td>
<td>Gefitinib, EGFR kinase inhibitor</td>
<td>34 (3 not treated)</td>
<td>50</td>
<td>9.2</td>
<td>17.5</td>
</tr>
<tr>
<td>Tamura and colleagues (125)</td>
<td>EGFR mutations</td>
<td>Gefitinib, EGFR kinase inhibitor</td>
<td>28</td>
<td>75</td>
<td>11.5</td>
<td>Not reached</td>
</tr>
<tr>
<td>Cappuzzo and colleagues (17)</td>
<td>EGFR FISH, pAKT, nonsmokers</td>
<td>Gefitinib, EGFR kinase inhibitor</td>
<td>42</td>
<td>47.6</td>
<td>6.4</td>
<td>Not reached</td>
</tr>
<tr>
<td>Sunaga and colleagues (121)</td>
<td>EGFR mutations</td>
<td>Gefitinib, EGFR kinase inhibitor</td>
<td>21</td>
<td>76</td>
<td>12.9</td>
<td>Not reported</td>
</tr>
<tr>
<td>Sutani and colleagues (122)</td>
<td>EGFR mutations</td>
<td>Gefitinib, EGFR kinase inhibitor</td>
<td>27</td>
<td>78</td>
<td>9.4</td>
<td>15.4</td>
</tr>
<tr>
<td>Asahina and colleagues (5)</td>
<td>EGFR mutations</td>
<td>Gefitinib, EGFR kinase inhibitor</td>
<td>16</td>
<td>75</td>
<td>8.9</td>
<td>Not reached</td>
</tr>
<tr>
<td>West and colleagues (137)</td>
<td>Bronchioloalveolar carcinoma</td>
<td>Gefitinib, EGFR kinase inhibitor</td>
<td>136</td>
<td>10</td>
<td>3.5</td>
<td>13</td>
</tr>
<tr>
<td>Chen and colleagues (23)</td>
<td>East Asian</td>
<td>Gefitinib, EGFR kinase inhibitor</td>
<td>36</td>
<td>33.3</td>
<td>4.7</td>
<td>9.5</td>
</tr>
<tr>
<td>Clamon and colleagues (29)</td>
<td>HER2 expression on IHC</td>
<td>Trastuzumab, HER2 monoclonal antibody</td>
<td>24</td>
<td>4</td>
<td>2.6</td>
<td>5.3</td>
</tr>
<tr>
<td>Janne and colleagues (61)</td>
<td>Expression of at least one ERBB family protein on IHC</td>
<td>CI-1033, panHER kinase inhibitor</td>
<td>166 (3 arms)</td>
<td>2–4</td>
<td>1.9</td>
<td>6–6.6</td>
</tr>
</tbody>
</table>

Table 2. Studies enriched for patient population with putative molecular targets

Published OnlineFirst September 26, 2012; DOI: 10.1158/1078-0432.CCR-12-0178
reported RR and PFS, RR and OS, and PFS and OS (Fig. 3). On the other hand, this finding should be interpreted with caution as OS is also largely influenced by survival after progression, which frequently depends upon the administration of postprotocol therapies not captured in our analysis (31). In addition, another possible limitation is that PFS and OS were calculated together for treatment-naïve and previously treated patients.

In our study, most of the published trials enriched for putative molecular targets used an EGFR inhibitor, gefitinib. Nevertheless, the subgroup analysis of gefitinib studies showed that treatment arms enriched by patients with molecular targets outperformed treatment arms with unselected patient populations for all outcome parameters (RR, PFS, and OS). In addition, a recent clinical trial with an anaplastic lymphoma kinase (ALK) inhibitor, crizotinib, in patients with EML4-ALK-positive NSCLC showed a RR of 57% in the early-phase clinical trial (16). Because EML4-ALK positivity occurs in only about 2% to 7% of NSCLC, and patients without this aberration are unlikely to respond to crizotinib, these results further support our current observation that molecular enrichment strategies are uniquely associated with an increased RR (16). In addition, an enrichment approach is further supported by data from The Biomarker-integrated Approaches of Targeted Therapy for Lung Cancer Elimination (BATTLE) trial, which was a biomarker-based, adaptively randomized study in pretreated patients with NSCLC (32).

A second limitation of our study is the retrospective nature of the analysis. Third, there are likely to be unpublished studies or trials presented in abstract form that were not included in the analysis. Another limitation of our study is that we only included studies with single-agent therapies and not combinations. Lastly, reported studies used different criteria for response evaluation and in some of them, but not all, responses were confirmed by independent review.
The use of different criteria could introduce a bias; however, several studies previously reported an acceptable concordance, ranging from 79% to 96%, between Response Evaluation Criteria in Solid Tumors and World Health Organization or SWOG criteria (33–37). In addition, many clinical trials showed that response assessment done by investigators is highly concordant with independent review (38).

In conclusion, 143 single-agent phase II trials that enrolled 7,701 patients showed similar, relatively poor, outcomes (medians for RR, PFS, and OS were 10%, 2.8 months, and 7.6 months, respectively) regardless of treatment type. In contrast, trials of targeted therapies enriched by the presence of molecular targets had significantly higher RRs, PFS, and OS (48.8%, 6 months, and 11.3 months, respectively). Furthermore, in these trials, RR correlated with PFS and both RR and PFS correlated with OS. These observations suggest that drug development will remain slow unless relevant biomarkers are identified and used in clinical trials in NSCLC, and that either RR or PFS may be a useful surrogate endpoint for OS.

Disclosure of Potential Conflicts of Interest

D.A. Berry is a co-owner and consultant/advisory board member of Berry Consultants, LLC and has ownership interest (including patents) in Berry Consultants, LLC. D.J. Stewart has honoraria from Speakers Bureau of Pfizer and is a consultant/advisory board member of Roche and Align2Action. No potential conflicts of interest were disclosed by the other authors.

Authors’ Contributions

Conception and design: F. Janku, D.A. Berry, R. Kurzrock
Development of methodology: F. Janku, D.A. Berry, H.A. Parsons
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): F. Janku, J. Gong, H.A. Parsons
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): F. Janku, D.A. Berry, H.A. Parsons
Writing, review, and/or revision of the manuscript: F. Janku, D.A. Berry, H.A. Parsons, D.J. Stewart, R. Kurzrock
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): F. Janku, H.A. Parsons
Study supervision: F. Janku

Acknowledgments

The authors thank Ms. Joann Aaron for scientific review and editing of this article.

Grant Support

This study was supported in part by grant number R0024148 from the National Center for Research Resources, a component of the NIH Roadmap for Medical Research.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received January 18, 2012; revised July 30, 2012; accepted September 16, 2012; published OnlineFirst September 26, 2012.

References


