Perspective

Developing Safety Criteria for Introducing New Agents into Neoadjuvant Trials

Angela DeMichele1, Donald A. Berry2, JoAnne Zujewski3, Sally Hunsberger3, Larry Rubinstein3, Joseph E. Tomaszewski3, Gary Kellogg4, Jane Perlmutter, Meredith Buxton5, Julia Lyandres4, Kathy S. Albain7, Chris Benz4, A. Jo Chien4, Paul Haluska9, Brian Leyland-Jones11, Minetta C. Liu9, Pamela Munster4, Olufunmilayo Olopade8, John W. Park4, Barbara A. Parker5, Lajos Pusztai12, Debu Tripathy6, Hope Rugo4, Douglas Yee10, and Laura Esserman4

Abstract

New approaches to drug development are critically needed to lessen the time, cost, and resources necessary to identify and optimize active agents. Strategies to accelerate drug development include testing drugs earlier in the disease process, such as the neoadjuvant setting. The U.S. Food and Drug Administration (FDA) has issued guidance designed to accelerate drug approval through the use of neoadjuvant studies in which the surrogate short-term endpoint, pathologic response, can be used to identify active agents and shorten the time to approval of both efficacious drugs and biomarkers identifying patients most likely to respond. However, this approach has unique challenges. In particular, issues of patient safety are paramount, given the exposure of potentially curable patients to investigational agents with limited safety experience. Key components to safe drug development in the neoadjuvant setting include defining a study population at sufficiently poor prognosis with standard therapy to justify exposure to investigational agents, defining the extent and adequacy of safety data from phase I, detecting potentially harmful interactions between investigational and standard therapies, improving study designs, such as adaptive strategies, that limit patient exposure to ineffective agents, and intensifying safety monitoring in the course of the trial. The I-SPY2 trial is an example of a phase II neoadjuvant trial of novel agents for breast cancer in which these issues have been addressed, both in the design and conduct of the trial. These adaptations of phase II design enable acceleration of drug development by reducing time and cost to screen novel therapies for activity without compromising safety. Clin Cancer Res; 19(11); 2817–23. ©2013 AACR.

Introduction

Improvements in understanding cancer biology and the human genome have led to the development of new classes of “targeted therapies” designed to interfere with critical molecules and pathways driving tumor growth and hold promise to improve outcomes while reducing toxicity. However, the process of clinical testing for investigational agents has remained relatively stagnant for decades. Drug development has followed a series of expensive and time-consuming steps, from preclinical testing through phase I, II, and III, requiring thousands of patients, hundreds of millions of dollars, and usually more than a decade for each successful compound to come to market (1). New approaches to drug development are critically needed to lessen the time, cost, and resources necessary to identify and optimize active agents. The American Society of Clinical Oncology’s recently published ‘Blueprint for Transforming Clinical and Translational Cancer Research’ calls for an overhaul in the design and conduct of clinical trials and new strategies to accelerate drug development (2).

Recognizing this need, the U.S. Food and Drug Administration (FDA) has recently issued guidance outlining a new pathway to accelerate drug approval through neoadjuvant trials in breast cancer (3). Because outcomes for neoadjuvant and standard adjuvant chemotherapy are equivalent (4), such trials in early or locally advanced cancer provide a unique and powerful opportunity to examine the efficacy of investigational agents and identify the patients and tumors in which they are most effective without compromising curability (5). Among the numerous benefits to this approach include in vivo assessment to provide real-time examination of tumor response...
and pharmacodynamics while the tumor remains in its microenvironment. The primary tumor is accessible for concurrent biomarker assessment and development of companion diagnostics needed to provide predictive markers of response. Most importantly, several measures, including pathologic response, residual cancer burden, and changes in proliferation as measured by Ki67, provide proximate surrogate end points that reflect later outcomes (4) and, in breast cancer, are specific for particular subtypes of the disease (6).

However, this approach also has unique challenges. In particular, issues of patient safety are paramount, given the exposure of potentially curable patients to investigational agents with limited safety experience. The FDA has appropriately warned that although "promising investigational agents should be incorporated into standard treatment for early-stage breast cancer as rapidly as possible, this goal must be weighed against the limited safety data available for new drugs when they are used in patients with curable cancer" (3). Phase II studies in incurable patients with advanced disease have been the standard way to gather drug safety data following phase I determination of the recommended phase II dose in which a modest expansion cohort has been treated. By the time new agents are tested in patients with curable disease, data are likely to be available from numerous phase II studies in tens to hundreds of advanced-disease patients, providing a robust safety profile. This approach, although more time consuming and expensive, has provided a cushion of comfort and safety to investigators, clinicians, and patients in subsequent testing of new agents in the early-disease setting.

New approaches and study designs addressing the issue of patient safety are necessary to enable a new agent to move directly from phase I to the potentially curative neoadjuvant setting to realize the benefits of accelerated drug development. An example of one such trial is the I-SPY2 Trial, designed to assess the benefits of adding novel agents to standard chemotherapy for breast cancer in the neoadjuvant setting. Key considerations in selecting investigational agents for phase II neoadjuvant testing were developed by a group of experts in the design phase of I-SPY2, as outlined in Table 1 and below. Addressing these concerns during the selection of agents and study design optimizes the potential success of this approach while mitigating risk to patients and minimizing the likelihood of prematurely discarding a potentially efficacious therapy.

Balancing Risk by Limiting Enrollment to a Study Population with Sufficiently High Risk for Poor Outcome with Standard Therapy

Traditionally, toxicity risk tolerance in oncology largely hinges on the underlying risk of disease progression and death; the higher the risk of poor outcome, the higher the tolerance for toxicity risk. In the case of neoadjuvant trials, the ability to identify appropriate patients for study inclusion requires the ability to accurately identify patients at sufficiently high risk that they are unlikely to be cured by standard therapy alone and to establish a level of risk that is justified ethically in exposing patients to investigational therapy. Risk assessment tools such as molecular prognostic and predictive markers are now available for many diseases to aid in this identification, including MammaPrint (Agendia) and OncotypeDx (Genomic Health). However, none of these tools have yet been validated in the neoadjuvant setting. Thus, if used for treatment assignment, these profiles should be conducted under an investigational device exemption (IDE).

Equally important, when assessing new agents in the neoadjuvant setting, is that the proximate end point, pathologic response at surgery, be a valid surrogate for ultimate survival outcomes, and that this surrogate be valid for the tumor subtype/molecular subtype being studied. In breast cancer, pathologic complete response is a strong surrogate for outcome in estrogen receptor-negative (ER−)/Her2+ and ER+/Her2− breast cancer but is not as useful for patients with ER+ disease, many of whom have little response to chemotherapy but have good prognosis due to the therapeutic effects of adjuvant hormonal therapy (6). The evolving understanding of breast cancer biology based on 4 distinct "subtypes" suggests that, in fact, some ER+ tumors (namely those that are of luminal B type) may have greater chemosensitivity than others (7). Endpoints for ER− neoadjuvant trials may instead require modification of the patient population (high-risk molecular profile) or endpoint [preoperative endocrine prognostic index (PEPI) score; ref. 8]. Without a valid response surrogate, assessment of an incremental benefit of a novel agent in the neoadjuvant setting is not possible.

Finally, fully informed consent of patients about their prognosis, known risks and benefits of standard therapy, and risks and benefits of participating in a trial of investigational therapy is paramount in this endeavor. Given that potential trial participants are at high risk of recurrence with...
standard treatment, patients themselves may be willing to consider a relatively high tolerance for toxicity from investigational therapy. Studies examining patient preferences for treatment show that a majority would accept treatment with experimental therapy for the modest gains in survival (9). Engagement of patient advocates in developing approaches to the consent process is essential to optimize patient understanding of the risks and benefits of neoadjuvant therapies and assure that consent is truly “informed.”

The Drug Selection Process: Defining Extent And Adequacy of Safety Data from Single-Agent and Combination Studies

Another critical consideration is whether sufficient toxicity data exist for a candidate investigational agent to provide a reasonable estimate of the risks to patients. Phase I studies in the advanced-disease setting provide the screen for severe safety signals if they exist, and many include an expansion cohort at the planned phase II dose to improve the ability to rule out unexpected and serious toxicity. However, typical phase I dose-expansion cohorts are small (8–10 patients), limiting the ability to detect rare toxicities. How do we determine the minimal number of patients who must be exposed to the agent (either alone or in combination with cytotoxic therapies) at the recommended phase II dose before the agent is deemed safe enough to be given to potentially curable patients? A threshold number of patients treated in the single-agent setting should be determined by the cumulative past experience with the investigational agent in single-agent studies, even if the agent will ultimately go to phase II testing in a combination treatment. Toxicities that have emerged in the single-agent setting should be specifically quantified and their prevalence determined by the cumulative past experience with the investigational agent as a single agent in a phase I or II setting.

Tables and figures are not provided in the natural text. The text continues with detailed discussion on the drug selection process, the role of patient advocates, and the importance of toxicology data in the context of different doses, schedules, and diseases. For most toxicities, mild forms of the adverse event (grade 1 and 2) would likely occur before serious or life-threatening grade 3 or 4 toxicities, so that examination of the types and grades of side effects seen in early trials is at least as important as the number of patients treated. Once toxicity prevalence data are analyzed, statistical probabilities can be used to estimate the likelihood of encountering a serious toxicity given the number of patients previously treated and the incidence of toxicity (Table 2). For example, if 50 patients have been treated with the agent at the phase II dose, there is a 95% probability of observing at least one
event if the true underlying rate of a given toxicity is 6%. If 82 patients were observed, there is a 99% probability of observing at least one event if the true underlying rate was 5%. Identification of more rare toxicities (≤1% in frequency) would require a much larger number of patients.

### Table 2. Minimal number of patients required in phase I study to evaluate toxicity at specific prevalence and confidence level

<table>
<thead>
<tr>
<th>True prevalence of event</th>
<th>Number of patients studied</th>
<th>Probability of observing event*</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>45</td>
<td>0.90</td>
</tr>
<tr>
<td>0.05</td>
<td>57</td>
<td>0.95</td>
</tr>
<tr>
<td>0.06</td>
<td>82</td>
<td>0.99</td>
</tr>
<tr>
<td>0.06</td>
<td>37</td>
<td>0.90</td>
</tr>
<tr>
<td>0.06</td>
<td>50</td>
<td>0.95</td>
</tr>
<tr>
<td>0.06</td>
<td>68</td>
<td>0.99</td>
</tr>
<tr>
<td>0.07</td>
<td>32</td>
<td>0.90</td>
</tr>
<tr>
<td>0.07</td>
<td>40</td>
<td>0.95</td>
</tr>
<tr>
<td>0.08</td>
<td>57</td>
<td>0.99</td>
</tr>
<tr>
<td>0.08</td>
<td>28</td>
<td>0.90</td>
</tr>
<tr>
<td>0.08</td>
<td>35</td>
<td>0.95</td>
</tr>
<tr>
<td>0.08</td>
<td>51</td>
<td>0.99</td>
</tr>
<tr>
<td>0.09</td>
<td>24</td>
<td>0.90</td>
</tr>
<tr>
<td>0.09</td>
<td>31</td>
<td>0.95</td>
</tr>
<tr>
<td>0.09</td>
<td>45</td>
<td>0.99</td>
</tr>
<tr>
<td>0.10</td>
<td>22</td>
<td>0.90</td>
</tr>
<tr>
<td>0.10</td>
<td>28</td>
<td>0.95</td>
</tr>
<tr>
<td>0.10</td>
<td>40</td>
<td>0.99</td>
</tr>
<tr>
<td>0.15</td>
<td>14</td>
<td>0.90</td>
</tr>
<tr>
<td>0.15</td>
<td>18</td>
<td>0.95</td>
</tr>
<tr>
<td>0.15</td>
<td>26</td>
<td>0.99</td>
</tr>
</tbody>
</table>

*The probability of observing one or more events assuming a binomial distribution with an event rate given by the true prevalence and the number of binomial observations given by the number of patients studied.

### Anticipating the Potential for Harmful Interactions between Investigational and Standard Agents

Another important consideration in the testing of novel agents in combination with standard chemotherapy is the potential for antagonism between therapies that would render standard therapy less effective than if it was given alone. Although several examples of this phenomenon can be cited in phase III studies, arguably these negative interactions could have been predicted from preclinical studies showing antagonism. For example, a combination of EGF receptor (EGFR) tyrosine kinase inhibitors (gefitinib or erlotinib) with chemotherapy was inferior to chemotherapy alone in several large phase III randomized trials in lung cancer (10–14); yet, preclinical models clearly showed that in most EGFR wild-type or sensitizing mutant non–small cell lung carcinoma cells, the concomitant gefitinib/cisplatin combination showed antagonism, likely because gefitinib interfered with cisplatin entry into the cell (15). Similar results were predicted for the antagonistic effects of combined chemo/endocrine therapy in breast cancer (16, 17).

Thus, preclinical and phase Ib studies play a critical role in screening the interactive effects of coadministration of both the investigational agent and the standard treatment to which it will be added at a dose and schedule similar to that of the planned phase II neoadjuvant trial. It is not essential that these data be obtained specifically in the disease of interest; data from other diseases are acceptable for this assessment. Fewer patients may be needed in the phase Ib study than in the preceding single-agent phase I study, given that the single-agent toxicity is already known. Table 2 shows the minimal sample size needed for a phase Ib study to have a reasonably high probability of observing at least one event given varying underlying true prevalence of the event. For example, to have a 95% probability of observing an event if the true prevalence rate is 1%, we would need to study 57 patients. To have a 95% probability of observing an event if the true rate is 5%, we would need to study 57 patients. To have a 95% probability of observing an event if the true rate is 10%, we would need to study 28 patients. To have a 90% probability of observing an event if the true rate is 15%, we would need to study 14 patients. In such a study a serious adverse event is observed, additional patients could then be enrolled in a phase I expansion cohort or a subsequent phase II study in advanced disease before moving to the curative setting.

### Minimizing Risk by Using Study Designs That Minimize Exposure of Study Subjects to Ineffective Therapies

Given the need to limit patient exposure to potentially harmful agents in neoadjuvant trials, adaptive study designs that minimize exposure without compromising statistical power are best suited to these studies. In an adaptive randomization design, the randomization ratio changes during a period of time on the basis of the current (Bayesian) probability that an arm is the better treatment (18, 19). Thus, the sample size needed to complete the study is reduced while patients preferentially receive assignment to more efficacious arms as the study progresses, thereby further optimizing the ratio of benefits to risks.

### Conducting Intensive Safety Monitoring during the Course of Trial Therapy

Even with hundreds of patients treated in a phase I study, a serious toxicity may not come to light, simply by chance or sufficiently low frequency. Most clinical trials monitor only serious adverse events in real time; data collection for standard, nonserious adverse events is not fully assembled until the end of the study. Given the risks associated with novel agents in the neoadjuvant setting and limited prior patient exposure, it is important that both serious adverse events and adverse events be collected and reviewed in real time during the trial to identify any safety signals that arise. Intensifying this approach with frequent Data and Safety Monitoring Board (DSMB) review can be used as another important safety step in the conduct of an investigational

---

Table 2. Minimal number of patients required in phase I study to evaluate toxicity at specific prevalence and confidence level.
phase II trial in curable patients. Although DSMBs are standard in phase III investigations, they are not typically used in phase II trials and commonly review safety data at a few predetermined time points in the course of a multiyear trial. Using DSMBs as a safety measure at frequent time points is feasible, but requires “real-time” data entry from study sites and rapid verification and analysis by the Data Coordinating Center, challenges not typically faced with a standard phase II trial.

Implementing These Strategies in the I-SPY2 Trial

The I-SPY2 Trial is a multicenter, randomized phase II neoadjuvant breast cancer trial designed to assess the incremental benefit of targeted investigational agents added to standard chemotherapy (Fig. 1; refs. 20, 21). In the planning and design process for the I-SPY2 trial, trial investigators engaged a panel of outside experts in drug development and biostatistics from the National Cancer Institute and the patient advocacy community to develop a set of guidelines for drug selection in the trial that are based on the points above. The trial design of I-SPY2 assigns patients to therapy based upon validated biomarkers of risk and response to therapy, i.e., ER, progesterone receptor, Her2, and MammaPrint. The decision to identify high-risk patients is based upon the prospective evaluation of these markers, which is based upon data from numerous neoadjuvant trials of standard therapy, including the I-SPY 1 Trial (6), in which the expected 3-year recurrence-free survival for patients with tumors more than 3 cm in size receiving standard anthracycline/taxane-based neoadjuvant therapy was 80% overall, but significantly worse for patients with Her2+ and with triple-negative tumors, as well as those with ER+/MammaPrint “high” disease who did not achieve a complete pathologic response. Patients with ER+/MammaPrint low tumors are excluded from I-SPY2 because the risk of exposure to an investigational agent outweighs the potential benefit. Patients receive investigational agents in combination with paclitaxel; phase Ib studies of a proposed agent in combination with paclitaxel are required at a dose and schedule approximating the I-SPY2 treatment plan. The threshold set for likelihood of a grade 4 toxicity is less than 15%, requiring at least 15 to 20 patients in the expansion cohort to rule this out with 91% to 96% certainty (Table 2). For more serious anticipated toxicities (such as liver or cardiac toxicity), the threshold to rule out grade 4 toxicity must be 5% or less, necessitating study of at least 50 patients. Drugs “graduate” from the trial when they have shown sufficient activity to have an 85% likelihood of being efficacious in a phase III trial, or failing that, when they have been given to 120 patients, providing a limit to the number of patients who would be exposed to an inactive drug. Safety data are collected electronically through a web-based application (TRANSCEND) developed specifically for the trial in collaboration with the National Cancer Institute Center for Bioinformatics. The I-SPY2

![Diagram of the I-SPY2 trial](attachment:image.png)

Figure 1. Schema of the I-SPY2 trial. In the screening phase of I-SPY2, consented patients have core breast tumor biopsies that undergo profiling on a 44K microarray that includes MammaPrint under an IDE. Patients who are ER+ /MammaPrint high, ER−, or Her2+ are eligible for the study, provided they meet other eligibility criteria (including ability to obtain tumor volume by MRI and adequate organ function). Patients who are screen-eligible are then randomized and sign treatment consent to either standard neoadjuvant chemotherapy with weekly paclitaxel (with trastuzumab for Her2+) or paclitaxel combined with one of several investigational agents. This is followed in all patients by 4 cycles of doxorubicin/cyclophosphamide. Patients on treatment undergo serial tissue collection and imaging. An adaptive design uses pathologic response at surgery and imaging response to modify randomization probabilities, increasing efficiency, and minimizing exposure to less efficacious agents. AC, Adriamycin/cyclophosphamide; R, randomized.
DSMBC meets monthly to review safety and outcome data, providing the opportunity to conduct frequent monitoring for unexpected safety signals and constantly reevaluate the risk/benefit ratio for a given drug or combination in the trial. To date, more than 200 patients have been adaptively randomized to one of the 4 different investigational treatment arms with intense safety review monthly by the DSMB and rapid response to toxicities encountered in early stages of the trial.

In summary, the neoadjuvant setting provides a unique opportunity to speed drug development by enabling in vivo examination of tumor response and proximate outcomes, including pathologic response and relapse in high-risk patients. However, to accelerate drug development in this setting, new approaches are needed to assure the safety and well-being of potentially curable patients. The I-SPY 2 Trial is an example of a phase II neoadjuvant study in which such approaches have been developed. Given the recent FDA guidance outlining a path to regulatory approval for agents screened and active in the neoadjuvant setting, these issues will require close attention as more patients enter such trials.

Disclosure of Potential Conflicts of Interest

D.A. Berry is employed as a co-owner and statistical scientist; has ownership interest (including patents); and is a consultant/advisory board member of Berry Consultants, LLC. K.S. Albain is a consultant/advisory board member of Genentech, Roche, Pfizer, Novartis, Genomic Health, and Amgen. J.W. Park has received honoraria from the speakers’ bureau from Genentech, Agenda, Novartis, and Bristol-Mysers Squibb; has ownership interest (including patents) in Merrimack Pharmaceuticals; and is a consultant/advisory board member of Merrimack Pharmaceuticals and Genentech. B.A. Parker has a commercial research grant from GlaxoSmithKline and Sanofi-Aventis and is a consultant/advisory board member of Roche. No potential conflicts of interest were disclosed by the other authors.

References


Developing Safety Criteria for Introducing New Agents into Neoadjuvant Trials

Angela DeMichele, Donald A. Berry, JoAnne Zujewski, et al.