Issues in experimental design and endpoint analysis in the study of experimental cytotoxic agents *in vivo* in breast cancer and other models

Robert Clarke

Vincent T. Lombardi Cancer Center, Georgetown University Medical School, Washington DC, USA

**Key words**: xenografts, breast cancer, cell lines, resistance, cytotoxic drugs, synergy

**Summary**

Considerable effort has been placed into the identification of new antineoplastic agents to treat breast cancer and other malignant diseases. The basic approaches, in terms of model selection, endpoints, and data analysis, have changed in the previous few decades. This article deals with many of the issues associated with designing in *vivo* studies to investigate the activity of experimental and established compounds and their potential interactions. Endpoints for both in *situ* and excision assays are described, including approaches for determining cell kill, tumor growth delay, survival, and other estimates of activity. Suggestions for approaches that may limit the number of animals also are included, as are possible alternatives for death as an experimental endpoint. Other concerns, such routes for drug administration, drug dosage, and preliminary assessments of toxicity also are addressed. Statistical considerations are only briefly discussed, since these are addressed in detail in the accompanying article by Hanfelt (Hanfelt JJ, Breast Cancer Res Treat 46:279-302, 1997). The approaches suggested within this article are presented to draw attention to many of the key issues in experimental design and are not intended to exclude other approaches.

**Introduction**

The disseminated nature of breast cancer and the development of crossresistant tumors are the primary causes of failure of current therapies. By the time many tumors are detected, there is a high probability that metastatic lesions will be present, many of which may already contain resistant subpopulations [1]. Not surprisingly, there is substantial interest in the identification of novel cytotoxic and endocrine agents for the treatment of breast cancer.

A systemic approach is required to eradicate the majority of metastatic breast disease, and this remains primarily in the form of cytotoxic chemotherapy or endocrine manipulation. The latter began with the initial studies on oophorectomy by Beatson [2] and was followed by the administration of high dose estrogens. These were largely replaced by the development of antiestrogens,

Address for correspondence and affinities: Robert Clarke PhD, Vincent T. Lombardi Cancer Center and Department of Physiology & Biophysics, W405A The Research Building, Georgetown University Medical School, 3870 Reservoir Road NW, Washington DC 20007, USA; Telephone: (202) 687-3755; Fax: (202) 687-7505; email: clarke@gnet.georgetown.edu
<table>
<thead>
<tr>
<th>Cell Line</th>
<th>ER status</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>BT-549</td>
<td>negative</td>
<td>**</td>
</tr>
<tr>
<td>MCF-7</td>
<td>positive</td>
<td>(107)</td>
</tr>
<tr>
<td>MCF7/ADR</td>
<td>negative</td>
<td>(39,40,106)</td>
</tr>
<tr>
<td>MDA-MB-231</td>
<td>negative</td>
<td>(108,109)</td>
</tr>
<tr>
<td>MDA-MB-435</td>
<td>negative</td>
<td>(108-110)</td>
</tr>
<tr>
<td>T47D</td>
<td>positive</td>
<td>(111)</td>
</tr>
</tbody>
</table>

* Kindly provided by Dr. Joseph Mayo.

** There is no primary citation for this cell line listed by the provider (ATCC, Rockville, MD), who indicate in their "Catalogue of Cell Lines and Hybridomas" that the cells were derived from a papillary invasive ductal tumor in a 72-year old patient. Metastatic disease was found in three of seven regional lymph nodes. The origins were Drs. W.G. Coutinho and E.Y. Lasfargues. A discussion of other human breast cancer cell lines can be found elsewhere [112].

Initially the triphenylene Tamoxifen [3] and, more recently, introduction of the steroidal anti-estrogen ICI 182,780 [4]. Other endocrine strategies include the use of inhibitors of estrogen biosynthesis, primarily aromatase inhibitors, and the use of LHRH agonists and antagonists.

The use of chemotherapeutic agents in the management of neoplastic disease began with Rhoad's description of the use of nitrogen mustard for the treatment of Hodgkin's lymphoma [5]. The number of cytotoxic agents available has increased substantially over the intervening years, with breast tumors generally exhibiting good overall response rates to several cytotoxic drug combinations. While chemotherapy can produce gains in overall survival, most patients with metastatic breast cancer will eventually recur. Many reasons may account for this failure, including poor dose scheduling, inappropriate combinations of drugs and the emergence of cell populations resistant to the antineoplastic agents.

Animal models provide one approach for the optimization of drug scheduling and for the identification of novel compounds that exhibit promise in such in vitro prescreens as those currently utilized by the National Cancer Institute [6]. For many years a primary in vivo screen, utilizing the L1210 and P388 murine leukemias, was an integral part of the NCI's preclinical drug development program. However, the lack of a significant representation of solid human tumors drew criticism, and partly explained the weakness of the screen to identify new agents active against the more common solid tumors, including breast cancer. Perhaps not surprisingly, the screen appeared to have been more successful in identifying agents active against hematologic malignancies. Some drugs with well established clinical efficacy (busulfan, hexamethylmelamine) fail to demonstrate substantial activity in the L1210/P388 screen [7,8].

The P388 and L1210 in vivo screen has largely been replaced by panels of disease-specific human tumor cell lines (prescreen) and xenografts (primary screen). While there are relatively few ideal in vivo models for breast cancer, several of the available human cell lines fulfill some of the requirements for screening, e.g., stable phenotype, high tumor take rate, predictable and reproducible kinetic properties. Several of those currently used by NCI are shown in Table 1.

Some considerations for choice of model, scheduling, dosage, and endpoint will be discussed below. The discussions and issues raised are to draw attention to different approaches to experimental design. As such, these should be considered in the light of the other articles on experimental design and data analysis in this issue and elsewhere. Some issues are generic, while others are more directly applicable to screening cytotoxic activities, and may be of lesser relevance to testing chemopreventive and endocrine agents. It is hoped that the topics discussed will assist investigators to consider key issues in experimental design. However, these are provided only as suggestions. There are many ways to identify potentially active compounds and drug combinations, and these are constantly being
modified by improvements in both cellular models and approaches to the use of animals in biomedical research.

Log cell kill and tumor kinetics in cytotoxic cancer chemotherapy

Early attempts to modify cytotoxic therapies were largely empirical. The initial criteria for designing cytotoxic therapies were based on the observations of Skipper, who demonstrated that cytotoxic drug-induced cell kill follows a similar kinetic pattern to that established by Arhenius at the turn of this century for the killing of bacteria. The log cell kill hypothesis states that cytotoxic drugs kill cells by first order kinetics [9]. Thus, a constant proportion of the cells will be killed regardless of the size of the cell population. For many cytotoxic drugs a 4-log cell kill is achievable and would eradicate a tumor population of $10^5$ cells and have a one in ten chance of eliminating a population of $10^6$ cells [10]. However, clinically detectable primary tumors have a cell population frequently in excess of $10^9$ cells. Many metastases also could contain cell populations greater than $10^9$ cells. While theoretically sound, log cell kill can be affected by several biological parameters, including the presence of de novo resistant cells, the degree of tumor vascularity, and other factors that may affect drug perfusion and metabolism.

The principles elucidated by Skipper were further modified by Norton & Simon, who applied Gompertzian growth kinetics to tumor cell populations [11]. This clarified the inverse relationship between growth fraction and tumor size. The ability of a cytotoxic drug to inhibit tumor cell growth was determined to be directly related to the tumor’s growth rate, which also is related to tumor volume [12]. The tumor mass killed by a cytotoxic treatment is proportional to the growth fraction multiplied by the total tumor volume [12].

A Gompertzian kinetic growth pattern produces a growth fraction that ultimately decreases exponentially with time. This inverse relationship between tumor size and growth fraction implies that micrometastases should be more kinetically sensitive to cytotoxic chemotherapy than larger tumor masses [13,14]. Thus, early intervention when the tumor mass is small should provide the greatest opportunity for induction of remission. Metastases tend to exhibit a more rapid tumor doubling time ($T_D$) than primary tumors, particularly for the common solid tumors like those of the breast [15].

There are a number of factors that contribute to the apparent $T_D$ of any tumor. These include the rate of cell production, the size of the growth fraction, cell recruitment from $G_0$, and the rate of cell loss. Cell loss includes shedding of cells into other compartments, e.g., metastasis, differentiation to a non-proliferating cell type, or entry into prolonged $G_0$ and apoptotic cell death. The high proportion of non-proliferating or normal cells present in many solid tumors, and the frequently high rate of cell loss, generally produce tumors with a long $T_D$. The rate of cell loss may be as high as 80% of the rate of cell production [16].

Growth kinetics in human tumors and animal models

The kinetic parameters of tumor growth represent one of the major differences between animal models and the human disease. While Gompertzian kinetics apply to experimental tumors and those in patients, the $T_D$ and growth fractions are frequently quite different. For example, many human breast tumors exhibit long $T_D$ and often small relative growth fractions. In marked contrast, human breast tumor xenografts generally have short $T_D$ and high growth fractions. Estrogen-treated MCF-7 tumors (ER-positive) have $T_D$ of approximately 10-12 days [17-19] compared with over 100 days for many tumors in patients [20]. We have generated several estrogen-independent MCF-7 variants. These have $T_D$ as long as 100 days when grown in the absence of
estrogen, but grow as rapidly as parental MCF-7 tumors in estrogen-supplemented mice [17-19]. MDA435/LCCOCell is ER-negative, ascites variant of MDA-MB-435 growing as solid tumors have much shorter T/2 of 3-5 days [21]. Most of the breast cancer xenografts used in the current NCI screen have mean T/2 of 2-10 days.

The differences in kinetic properties between xenografts and tumors in patients would tend to make the xenografts more sensitive to agents with a strong cell cycle-related phase specificity. For the purposes of a primary in vivo screen for novel agents, a limited overestimate of activity may not be a major concern. For studies to optimize scheduling or combinations of established drugs, the relative sensitivity of the in vivo screen is a concern only if the model is either too sensitive or too resistant to a combination, when it will become difficult to assess interactions. Most of these concerns are readily addressed by a careful choice of in vivo model(s). Various schedules have already been identified for the established drugs; examples are provided in Table 2.

The dose response relationship and dose intensity

The relationship between treatment and response is described by:

\[ k = C \times t \]

where C = concentration; t = time. Thus, response should be approximately equivalent where C x t values (area under the concentration time curve) are equivalent. This can enable the design of clinically relevant in vitro analyses of established drug combinations based on pharmacokinetically measured previously obtained in patients or animals. Clinical studies can utilize the reasonable across-species dosage relationship of mg/m² to estimate dose from data obtained in preclinical animal screening. The doses may require some further modification, since the serum half-life of some drugs can be longer in man than in rodents [22]. One approach is to use 1/10th the maximum tolerated dose (MTD) in rodents as the approximate starting dose for a Phase I trial in humans [23]. Where possible, it may be better to use a dose that produces comparable pharmacokinetics, since this can increase the predictability of the mouse xenograft-to-human tumor model [24].

The pharmacokinetics for cytotoxic drugs are frequently similar in mice and men [25]. For many drugs, the mouse LD₅₀ also approximates the MTD in humans when expressed as mg/m² [26]. However, there are exceptions. The C x t values at the LD₅₀ are higher for mitomycin C, vincristine, and cyclophosphamide, and lower for methotrexate and 5-fluorouracil, in mice when compared with humans [reviewed in (26)].

In clinical practice, a narrow therapeutic index is frequently responsible for the reduction of dosage due to side effects. However, the steep dose response curve for most cytotoxic drugs implies that even a small perturbation in dosage may produce a significant change in response. The effects of alterations in dosage on clinical response has been widely reviewed [22,27]. It has been suggested that a major contributing factor to the failure of many treatments is the ad hoc reduction of drug dosage [22].

It has been widely acknowledged that the most effective treatments involve a high dose intensity chemotherapeutic regimen. This is partly based on the steep dose response relationship for most cytotoxic drugs and various other clinical observations. Dodwell et al. [27] have reviewed the published data, and reexamined the role of dose intensity in response, for a number of the more common malignancies. While they conclude that high intensity regimens can produce significant advantages in disease free survival, clear demonstrations of increased overall survival are obtained much less frequently. The relationship between dose intensity and response often varies with both drug, tumor model, and disease. Animal models provide a safe and logical means to explore this and related issues, rather than attempting to identify appropriate or potentially dangerous schedules directly in patients.
<table>
<thead>
<tr>
<th>Drug</th>
<th>Dose (mg/kg)</th>
<th>Route</th>
<th>Schedule</th>
<th>Toxicity</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adriamycin</td>
<td>6</td>
<td>i.v.</td>
<td>4,5,6</td>
<td>ND&lt;sup&gt;1&lt;/sup&gt;</td>
<td>[67,113]</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>i.v.</td>
<td>1,5,9</td>
<td>ND</td>
<td>[114]</td>
</tr>
<tr>
<td></td>
<td>6.8</td>
<td>i.p.</td>
<td>single</td>
<td>dose&lt;sup&gt;4&lt;/sup&gt;</td>
<td>none</td>
</tr>
<tr>
<td></td>
<td>8.5</td>
<td>i.p.</td>
<td>single</td>
<td>dose</td>
<td>LD&lt;sup&gt;3&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>i.v.</td>
<td>single</td>
<td>dose</td>
<td>MTD</td>
</tr>
<tr>
<td>Ara-C&lt;sup&gt;4&lt;/sup&gt;</td>
<td>40</td>
<td>i.m.</td>
<td>24 h infusion</td>
<td>none</td>
<td>[115]</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>i.p.</td>
<td>1,7</td>
<td>alopecia</td>
<td>[116]</td>
</tr>
<tr>
<td>BCNU</td>
<td>18</td>
<td>i.p.</td>
<td>single</td>
<td>dose</td>
<td>LD</td>
</tr>
<tr>
<td>Methy CCNU</td>
<td>18</td>
<td>i.p.</td>
<td>single</td>
<td>dose</td>
<td>MTD&lt;sup&gt;7&lt;/sup&gt;</td>
</tr>
<tr>
<td>Cyclophosphamide</td>
<td>35</td>
<td>i.p.</td>
<td>single</td>
<td>dose</td>
<td>alopecia</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>i.p.</td>
<td>0.4, 7-11</td>
<td>ND</td>
<td>[117]</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>i.v.</td>
<td>1,5,9</td>
<td>ND</td>
<td>[114]</td>
</tr>
<tr>
<td></td>
<td>143</td>
<td>i.p.</td>
<td>single</td>
<td>dose</td>
<td>NOD&lt;sup&gt;11&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>i.p.</td>
<td>1,15</td>
<td>LD</td>
<td>[119]</td>
</tr>
<tr>
<td></td>
<td>286</td>
<td>i.p.</td>
<td>single</td>
<td>dose</td>
<td>LD</td>
</tr>
<tr>
<td></td>
<td>290</td>
<td>i.p.</td>
<td>single</td>
<td>dose</td>
<td>none/LD&lt;sup&gt;10&lt;/sup&gt;</td>
</tr>
<tr>
<td>5-Fluorouracil</td>
<td>32</td>
<td>i.p.</td>
<td>5,6,13</td>
<td>ND</td>
<td>[72]</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>i.p.</td>
<td>1-4, 15-48</td>
<td>ND</td>
<td>[119]</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>i.p.</td>
<td>0.4</td>
<td>ND</td>
<td>[117]</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>i.v.</td>
<td>1,5,9</td>
<td>ND</td>
<td>[114]</td>
</tr>
<tr>
<td></td>
<td>180</td>
<td>i.p.</td>
<td>single</td>
<td>dose</td>
<td>MTD&lt;sup&gt;25&lt;/sup&gt;</td>
</tr>
<tr>
<td>Ifosfamide</td>
<td>150</td>
<td>i.v.</td>
<td>single</td>
<td>dose</td>
<td>none</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>i.v.</td>
<td>single</td>
<td>dose</td>
<td>none</td>
</tr>
<tr>
<td>Melphalan</td>
<td>12</td>
<td>i.p.</td>
<td>single</td>
<td>ND</td>
<td>[120]</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>i.p.</td>
<td>0.4</td>
<td>ND</td>
<td>[120]</td>
</tr>
<tr>
<td>Mitomycin C</td>
<td>1</td>
<td>i.p.</td>
<td>1,4,8,11,15,18</td>
<td>ND</td>
<td>[119]</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>i.p.</td>
<td>1-5</td>
<td>ND</td>
<td>[119]</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>i.p.</td>
<td>0-3</td>
<td>ND</td>
<td>[117]</td>
</tr>
<tr>
<td>Mitomycin C</td>
<td>2</td>
<td>i.v.</td>
<td>1,15</td>
<td>LD</td>
<td>[130]</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>i.p.</td>
<td>single</td>
<td>dose</td>
<td>LD</td>
</tr>
<tr>
<td></td>
<td>4.5</td>
<td>i.p.</td>
<td>single</td>
<td>dose</td>
<td>none</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>i.p.</td>
<td>single</td>
<td>dose</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>i.p.</td>
<td>single</td>
<td>dose</td>
<td>MTD&lt;sup&gt;25&lt;/sup&gt;</td>
</tr>
<tr>
<td>cisplatin</td>
<td>4</td>
<td>i.p.</td>
<td>1,15,10</td>
<td>ND</td>
<td>[123]</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>i.v.</td>
<td>1,15,10</td>
<td>ND</td>
<td>[123]</td>
</tr>
<tr>
<td></td>
<td>7.5</td>
<td>i.p.</td>
<td>single</td>
<td>dose</td>
<td>none</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>i.p.</td>
<td>single</td>
<td>dose</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>27</td>
<td>i.p.</td>
<td>single</td>
<td>dose</td>
<td>MTD&lt;sup&gt;25&lt;/sup&gt;</td>
</tr>
<tr>
<td>Taxol</td>
<td>20</td>
<td>i.p.</td>
<td>single</td>
<td>dose</td>
<td>none</td>
</tr>
<tr>
<td>Temozolomide</td>
<td>15</td>
<td>i.v.</td>
<td>4,4,8</td>
<td>MTD</td>
<td>[67]</td>
</tr>
<tr>
<td>Vinblastine</td>
<td>3</td>
<td>i.v.</td>
<td>1,15</td>
<td>ND</td>
<td>[119]</td>
</tr>
</tbody>
</table>

<sup>1</sup> = Dose/Injection; <sup>2</sup> = Schedule is given as days unless otherwise indicated; <sup>3</sup> = Not defined or set dose<sup>4</sup> = None reported or no deaths; <sup>4</sup> = Lethal dose (one or more deaths attributed to drug-induced toxicity); <sup>5</sup> = 1-B-D-xylofuranosylcytosine; <sup>7</sup> = Approximate MTD as defined by the investigators.
General considerations in experimental design

Choice of host for xenografts

There are several considerations relating to the choice of immunocompromised host, including the degree of immune-competence and the ability of the strain to support the growth of the tumor. For syngeneic tumors the choice of host should be obvious. However, the choice of host for xenografts is more complex. The host not only should facilitate the growth of the tumor but also should enable the tumor to maintain a biologically relevant phenotype. It also is useful if the selected model enables some estimation of the selectivity of drug action and the determination of potentially lethal toxicities. For most of the breast cancer cell lines/xenografts available, the nu/nu mouse is sufficient, irrespective of the background strain [28].

The potential for immunological modulation to contribute to tumor response to combination chemotherapy is an important consideration. This is particularly relevant when the combination includes biological response modifiers that can influence the activation of cells associated with cell mediated immunity, e.g., interferons and interleukins. In some cases, it may be desirable to attempt to either eliminate or exclude effects on the immune competence of the host. Thus, the choice of an appropriate immune-deficient strain may become paramount. Since the different immunophenotypes of available rodent hosts have been recently reviewed [28], they will not be further discussed.

Choice of appropriate tumor model

The model used for the screening of a drug suspected of activity against a particular tumor should reflect the biological properties of that tumor as closely as can reasonably be achieved. For example, a drug active against a leukemia with a short T$_D$, high growth fraction, and relatively short cell cycle time would be less likely to demonstrate activity in a screen against a breast tumor with a longer T$_D$, small growth fraction, and long cell cycle time. Thus, choice of an appropriate model to screen combinations of agents with known pharmacological and kinetic requirements should (where possible) closely reflect the major biological properties of the human disease. To some extent the model also should reflect the requirements of the drug. For example, there would be limited value assessing a drug expected to show no crossresistance to P-glycoprotein without including a model that expresses P-glycoprotein.

Selecting a breast tumor model for screening cytotoxic compounds can be problematic. Relative to the murine ascites models, most solid breast tumors exhibit a relatively slow T$_D$. For example, MCF-7 tumors have a T$_D$ of approximately 10-12 days when growing in appropriately estrogen supplemented animals [19]. While this may be acceptable for screening antiestrogenic or chemopreventive agents, estrogen supplementation can alter the activity of some cytotoxic drugs [29,30]. In general, the most rapidly proliferating human breast tumor xenografts do not express estrogen receptors. MDA-A435/LCC6 tumors have a T$_D$ of 2-3 days. Since the parental cell line (MDA-MB-435) was obtained from a patient who had not received chemotherapy, these models may be useful in screening cytotoxic agents for activity against breast cancer [21]. The MDA-MB-231 cell line also has a comparably rapid T$_D$ in vivo (R. Clarke, unpublished observations). The more rapid T$_D$ of ER-negative xenografts broadly reflect the characteristics of such tumors in patients, where ER-negative tumors tend to have shorter T$_D$ [31].

From a purely practical viewpoint, relatively slowly proliferating breast tumors can produce significant problems in logistics and experimental design. The slower growing tumors frequently exhibit significant intertumor variability and can require substantial numbers of animals to enable meaningful statistical analysis of data. Determining the period at which a tumor is considered
"cured" also can become problematic. Even the more rapidly proliferating solid tumors with $T_{50} = 48$ hr may require up to four months of post-treatment observation to establish "cure" [32].

It is unlikely that any one tumor model will adequately represent the major biological characteristics of a particular malignancy. Thus, the use of a series of tumors (where appropriate/available) may be required to determine the sensitivity of a particular neoplastic disease to a either a single or a combination chemotherapy regimen. However, this must be considered in the context of reducing animal usage, cost, and the value of the additional data obtained.

For breast cancer, there are several potential models available for screening (Table 1). Most of the ER-positive models require estrogenic supplementation for tumorigenicity or maximal growth. We have generated ER-positive models that will grow without supplementation, but the respective $T_{50}$s are relatively long [19]. While this may be more representative of breast tumors in general, this characteristic is inappropriate for screening cell cycle or cell phase specific cytotoxic compounds. We also have developed an ascites model based on a variant of the MDA-MB-435 cell line (MDA435/LCC6). The pattern of response to a variety of cytotoxic drugs appears to reflect closely that seen in breast cancer patients [21]. For example, breast cancers in general respond poorly to nitrosoureas [33], as do MDA435/LCC6 ascites to BCNU. Etoposide also does not produce long term survivors, and this drug generally has been ineffective as a single agent in breast cancer [34,35]. Adriamycin [34], mitomycin C [36], and taxol [37] are among the most effective single agents in previously untreated breast cancer, and all of these drugs produced long term survivors in mice bearing the MDA435/LCC6 ascites. The characteristics of several breast cancer xenografts have been reviewed elsewhere [28].

Another example of the choice of tumor model is in studies to evaluate P-glycoprotein reversing agents, compounds which may have significant potential in some breast cancer patients [38]. For these types of analyses the choice of tumor model is critical. Cells to be used as xenografts should be transfectants rather than selected for resistance, since selection can produce multiple unrelated resistance mechanisms. For example, MCF-7ADR (selected for resistance to adriamycin [39]), but not MDR1-transduced MCF-7 cells (CL 10.3), are cross resistant to Tumor Necrosis Factor [40]. Since both adriamycin and Tumor Necrosis Factor can inhibit cells by the generation of free radicals [41,42], this cross resistance in MCF7ADR cells suggests the presence of functional adriamycin resistance mechanisms in addition to P-glycoprotein, including altered expression of manganous superoxide dismutase [40]. These cells also exhibit increased glutathione transferase and topoisomerase II activities [43,44]. The use of transfected cells allows for a clearer interpretation of the data. Cells concurrently expressing multiple resistance mechanisms may more closely reflect the drug resistance that occurs in patients [38], but interpreting responses in a mechanistic light may be difficult. This does not invalidate their use where the purpose is simply to screen compounds for potential antineoplastic activity. Indeed, the choice of a series of models that are too sensitive will likely identify compounds with limited activity in patients, whereas active compounds identified in otherwise resistant models may have a higher probability of being active in patients [45]. This is likely to be true in principle, but the extent to which it applies will depend upon whether the resistance mechanisms operating in the tumor model contribute significantly to the resistance phenotype in patients.

It is apparent that there are two potential types of screening approaches, each with different objectives that will result in different choices of models. Where a broad based, non-mechanism oriented screen is required, a disease-specific panel of xenografts with widely differing but biologically relevant phenotypes is likely to be optimal. Knowledge of the pattern of response to a series of established drugs, for each component of the panel, is required. Such a panel might be
expected to contain both sensitive and resistant models (see Table 1 for examples included in one possible panel for breast cancer). For a mechanism or structure/function-based screen, the choice of components is likely to depend upon assumptions inherent in the mechanism. For example, where a specific target is identified, the panel may contain several models with different levels of expression of the target, e.g., P-glycoprotein, multidrug resistance related protein [21,38,46,47]. In many cases, this requirement may be most effectively met by a series of transfected cell lines and their respective control populations.

**Phenotypic stability**

The stability of the phenotype is a critical determinant for tumor model selection. Some tumor xenografts may require periodic cycles of in vivo/in vitro growth in order to maintain the ease of reestablishment in vitro for some excision cytotoxicity assays. Prolonged in vivo growth of some established cell lines can result in significant phenotypic alterations. We have described the isolation of hormone-independent sublines of the estrogen-dependent MCF-7 human breast cancer cell line following prolonged selection in vivo [17,18]. While responses to antiestrogens remain unaltered [18,48], there are significant changes in their responsiveness to estrogens [17,18,48]. For many cell lines, this problem can be overcome by using cells within a limited number of passages (≤10) from a single frozen stock of cells. The frozen stock should be from a single passage of cells with a well-characterized phenotype.

The intertumor stability of the growth patterns and cell cycle profiles also may be important considerations. A high degree of variability in intertumor growth fractions could significantly influence the reliability or ease of data interpretation. The stability of the metastatic potential must be well defined. Cells with an unpredictable metastatic capacity may alter tumor burden and affect survival and/or the host's sensitivity to the toxicity of cytotoxic treatments.

**Therapeutic index and dose scheduling**

The difference in the dose response curves of normal and neoplastic tissues, often referred to as the therapeutic window or therapeutic index, is widely applied in the clinical pharmacology of cytotoxic drugs [22,49-51]. This difference in drug sensitivity enables the administration of sufficient drug to produce cytotoxic effects in the tumor, but not to induce significant and irreversible toxicity to normal cells. Unfortunately, many cytotoxic drugs exhibit a steep dose response curve with a small therapeutic index. In many cases, the development of unacceptable toxicity is the dose limiting factor for cytotoxic chemotherapy.

The sensitivity of normal cell populations reflects the rapid growth rate and high growth fraction of some cells. For example, cells in normal bone marrow and the intestinal crypts have respective thymidine labeling indices of 30%-70% and 12%-18% [15]. These values are generally higher than observed in most solid tumor cell populations. This accounts for the high incidence of hematopoietic and gastrointestinal side effects associated with many chemotherapeutic regimens. Granulocytes are highly susceptible to cytotoxic agents because of their high growth fraction and short lifespan, but a pool of non-cycling stem cells can regenerate [15]. Generally, the purpose of dose scheduling is to administer the second and subsequent treatments at times that will allow some normal stem cells to evade the drug and enable repopulation to occur without permanent damage to bone marrow. Examples of some drug schedules for rodents are provided in Table 2.

**Drug administration in vivo**

The route of administration can influence drug pharmacokinetics, toxicity, and antitumor activity. For example, injection i.v. introduces the drug directly into the blood, with a relatively rapid exposure of hematopoietic stem cells. Adminis-
tration s.c. or i.p. would be expected to produce a slower drug equilibration with this compartment, and produce a delayed or reduced myelo-suppression. However, local tissue damage could be increased relative to the i.v. route. The potential toxic/pharmacokinetic differences associated with routes of administration can be controlled by careful planning. For example, while an i.v. bolus of an agent may produce unacceptable toxicity, the same dose in mg/kg body weight can occasionally be administered either by infusions, s.c., p.o., or multiple lower doses given s.c., with significant alterations in host toxicity.

For the majority of solid tumors, an i.v. administration most closely reflects the clinical administration of the drug. However, for experimental drugs, the route of administration will vary primarily with the physico-chemical properties of the drug. The i.v. route is generally limited to water soluble compounds with a pH >4.3 and <8.5. Water-insoluble drugs can be administered either s.c., provided they do not produce unacceptable local damage, or p.o. if they have sufficient chemical stability. For steroid hormones and antihormones that can require sustained delivery, i.p. or s.c. depot in peanut oil, s.c. Alzet mini pumps [Alza Scientific, Palo Alto CA], cholesterol-based slow release pellets [Innovative Research of America, Sarasota, FL] or silastic pellets all can produce appropriate plasma levels of drug for sustained periods of time.

The timing of administration also is an important consideration. Initiation of treatment within a few days of tumor cell inoculation may produce evidence of activity, when administration to established tumors indicates inactivity. For cytotoxic agents, administration within a few days of cell inoculation is often inappropriate. Treatment of established tumors, where these are clearly palpable and fall within a predelineated size range, is generally more appropriate and allows for assessment of the most widely used endpoints. This use should not be so large as to influence response to the drug. Early administration of drug is usually appropriate when the tumor is directly xenografted from another animal and has a rapid T". Of only a few days, or perhaps when sensitivity is required in a primary screen. Another exception is for the testing of chemopreventive agents, e.g. stilbestrols or retinoids, which would be given to high risk women without evidence of clinically detectable disease. With these compounds, chemopreventive/chemosuppressive activity against low tumor burdens, e.g. recently inoculated tumor cell suspensions, may more closely reflect the proposed clinical use. Widespread initiation of treatment around the time of tumor cell inoculation is justified, i.e. before the appearance of palpable tumors, the more common endpoints include time to tumor appearance and tumor incidence.

Pharmacokinetics and toxicity in drug combination studies

There are a number of pharmacokinetic considerations that can influence activity of a drug in vivo when drugs are combined. The biologic activity of one agent may alter the metabolism, absorption, distribution, or toxicity of another. Insulin can alter the metabolism and sub sequent cytotoxicity of methotrexate [52]. Procarbazine reduces host toxicity and enhances the antitumor effects of nitrogen mustard, mephalan, and chlorambucil [53]. In some experimental tumor systems progestosterone can reduce the systemic toxicity of chlorambucil [52]. Maris et al. [54] have reported that adriamycin, lomustine, carmustine, semustine, and vincristine can increase mephalan uptake in L1210 cells. In patients with advanced breast cancer, the toxicity of a combination of cyclophosphamide, methotrexate, and 5-fluorouracil is reduced by fluoxymesterone [48]. Bleomycin cytotoxicity is increased by several membrane-acting drugs [55]. Illiger and Herdrich [56] have exhaustively reviewed many of the drug interactions encountered in cytotoxic chemotherapy.

The purpose of some experimental designs is to specifically test the interactions of drugs.
example, there is considerable interest in the generation of drugs that may reverse the efflux activities of P-glycoprotein, the glycoprotein product of the MDR1 gene [57]. There is clear evidence that some reversing agents, if not all, alter the pharmacokinetics of cytotoxic drugs [58]. This may reflect modulation of normal P-glyco-

protein function in the liver and other tissues, leading to an effective increase in drug exposure. Thus, increased activity of a cytotoxic agent requires careful evaluation.

One simple way to control for such effects is to compare equitoxic doses, e.g. the cytotoxic drug alone and in the presence of the reversing agent using schedules each of which produce an MTD. While this is a reasonable approach for pilot studies, directly assessing the pharmacokinetics of the cytotoxic drug with and without the reversing agent is definitive. Ultimately, such pharmacokinetic data may be required to clearly demonstrate that effects are not simply due to perturbations in the cytotoxic drug: mean serum concentration/plasma residence time.

Estimation of drug dosage and choice of starting dose/dose range

Where established drugs are to be used, well-doc-
umented protocols are available from the literature (Table 2). This is, at best, a general guide, since some rodent strains may be more or less sensitive to the toxic side-effects of specific agents.

Identifying a dose regimen for unknown or experimental agents is frequently empirical. Often the initial studies are performed to obtain estimates of the MTD/LDL. A simple dose escalation study with a limited number of animals per group is a common strategy. The choice of starting dose to obtain an MTD estimate could be based on one of several criteria, such as the known in vitro toxicity, or toxicity of a closely related compound. Where agents under investi-
gation are natural products or analogues thereof, e.g. phytochemicals, the levels of exposure in human populations may be available and enable estimation of the starting dose. One simple approach where resources are limited, is to do a pilot (dose range finding) study, using a broad range of doses but a small number of animals (one or two animals per dose). This has the advantage of limiting the number of animals that may be exposed to particularly toxic/lethal doses when the intent is to rapidly establish a toxic dose. A second, more definitive, follow-up study can then be done to obtain the MTD, using a limited number of mostly sublethal doses with larger numbers of animals per group. Animal usage can be further limited by restricting these follow-up studies only to those drugs that show antineoplastic activity and are identified for further evaluation in subsequent secondary screen-
ings.

For preliminary toxicological analyses, e.g., estimates of MTD, all animals should be moni-
tored twice daily at a minimum. Recording animal body weights and food consumption twice weekly can identify the onset of some toxicities that are not immediately apparent. Loss of 10% of the starting body weight, or failure to gain body weight at the same rate as controls, provide useful endpoints for estimating an MTD. Many cytotoxic agents will produce more immediate effects, which are often apparent from the altered behavior of the animals, e.g. crossing, somno-

lence, or reduced activity. These also can be used as endpoints to establish an MTD. Such effects may be transient, with the animals recovering within several hours if no further drug is admin-
istered, or they can persist and lead to morbidity and death if prolonged over several days or weeks.

Where drug related deaths occur within several hours of administration, these data should not be used to establish the MTD, since toxicity could merely reflect sensitivity to the peak plasma concentration resulting from the bolus. Rather, the dosage and/or schedule should be modified, e.g. using lower doses and perhaps more frequent administration. This pattern of toxicity, i.e. schedule-independent with peak plasma level toxicity, has been referred to as category III [59].
Other categories reflect whether toxicity is schedule dependent (category i) or independent (category ii, where total dose determines toxicity) [60,61].

Hematologic toxicity is often assessed as a toxicological endpoint, but the transient suppression of some cell populations, and the timing of their likely recovery, needs to be considered. It can be useful to measure white blood cell counts and hematocrits twice weekly, or as appropriate, on two or more individuals in each group. In many cases this can be done by retroorbital bleeding. However, the same animal should not be bled repetitively, since this can be sufficient to influence the parameters under investigation, independent of the treatment. It should also be noted that hematopoietic toxicity is not always evident from peripheral blood assays, and may require spleen colony formation analyses [49]. Additional evidence of toxicity may be apparent on determining other parameters, e.g. SPGT, total bilirubin, BUN, creatinine. Such detailed blood analyses are usually restricted to more intensive toxicological studies, where identifying the dose limiting or lethal toxicity is required. However, if the aim is to obtain an approximate MTD in a pilot study, it can be informative to sacrifice moribund or clearly affected animals by terminal bleeding under appropriate anesthesia. The blood can then be used for a wider panel of preliminary tests. It is usually routine to sacrifice all animals at the end of the study and perform necropsies with subsequent examination of the major tissues and organs for evidence of gross and/or microanatomic toxicity.

Site of inoculation, tumor-host interactions, and drug delivery

The site of tumor cell inoculation can significantly alter tumor growth and metastatic potential [62-65] and has previously been discussed in detail [28]. Cells re-inoculated into the site of the original tumor (orthotopic implantation) produce metastatic lesions more frequently than those inoculated elsewhere [64]. We routinely use the mammary fat pad area as the preferred inoculation site for mammary tumors [17]. Where orthotopic implantation is either inappropriate or not required, s.c. inoculation into the flank facilitates good responses for cytotoxic agents and s.c. inoculation into the back facilitates good responses for irradiation regimens [66].

The "tumor bed effect", which relates to modifications in the tissue at the site of implantation, has been most widely studied in relation to radiosensitivity but can also influence response to cytotoxic drugs [54]. Some tumor models utilize inoculation into preirradiated subcutaneous sites. Prior irradiation of normal tissues can significantly impair their ability to produce new vascular tissues in response to the tumor. Thus, drug delivery to the tumor can be reduced when tumors are grown in preirradiated sites. The "tumor bed effect" has been reviewed by Milas [54].

Techniques for the determination of in vivo antitumor activity

Several techniques have been used to determine the activity of antineoplastic agents against experimental tumors growing in vivo. These tend to fall into one of two main categories, in situ and excision assays. In situ assays are performed entirely in vivo. Tumors are inoculated into the appropriate host, treated in vivo, and the effects of drug treatment estimated either on various parameters of tumor growth, or on the duration of survival. Excision assays also are based on the treatment of tumors growing in vivo. However, the estimations of cytotoxicity require removal of the tumor for further evaluation. This can include estimating colony forming ability in vitro, or TD50 estimations of the number of treated cells required to form tumors on reimplantation into a second host. Each of these techniques has unique advantages and disadvantages. Irrespective of the tumor endpoint, many investigators include additional groups of tumor-bearing animals treated with a drug known to be active against the tumor model [67].
in situ assays — tumor growth delay

Perhaps the most widely used in situ technique determines the effect of drug treatment on the kinetics of tumor regrowth. Tumors are inoculated into the appropriate host, treatment being initiated when a specific tumor size is reached, and measured at regular intervals until both treated and untreated tumors have reached a predetermined size. Growth delay is assessed as the time difference between treated and control tumors to reach this pre-determined size. This approach measures growth of tumors of the same size, an important consideration when competitive/kinetics are involved.

There are a number of considerations specific to the design of tumor growth delay assessments. The metastasis should be performed on proliferating tumors, so that it is advisable to obtain some estimate of pretreatment tumor growth characteristics to enable tumor selection. For example, where the size at treatment is to be 5 mm, the range of tumor sizes for inclusion in combination chemotherapy-study might be ±1-2 mm in diameter. The issue of appropriate randomization of animals into each group is discussed elsewhere [68]. The optimal size at treatment will vary depending on the individual growth characteristics of the tumor cell line but should be sufficient to produce tumors that are easily measured yet not so large that they include significant areas of hypoxia or necrosis. The optimal endpoint size is close to the size at treatment, e.g. twice the treatment volume. This minimizes any effects of treatment on growth rate [69].

An alternative endpoint to growth delay for in situ techniques is overall survival. This method has proved reliable for the mouse ascites tumors, since the growth properties and lethal tumor burden are well established. Survival can easily be compared with an untreated tumor-bearing population. Survival is a less reliable endpoint for many solid tumors. The majority of solid human breast tumor models exhibit a poor or unpredictable metastatic potential. Thus, tumor burden is almost exclusively provided by the primary tumor. In these cases, the lethal tumor burden (primary tumor), as a percent of total body weight, can be far in excess of that observed in humans. Some poorly vascularized tumors may increase in volume and yet contain a relatively stable volume of viable issue.

Tumor measurement: area, volume, and weight

Tumor area and/or volume is usually recorded every 1-4 days, depending on the growth characteristics of the tumor. The length of the longest axis and the width perpendicular to the longest axis are sufficient to obtain tumor area. For easily accessible tumors, a third perpendicular measurement can be obtained to determine tumor volume. This is useful when it is apparent that palpable tumors are particularly irregular in shape and unlikely to meet the requirements for volume assessments from measurements of tumor area. For a more detailed description of tumor measurements see the accompanying article by Rygaard and Sparre-Thomsen [70].

Tumor weight, as determined at necropsy, is usually used as an endpoint in some drug studies. However, it is generally not advisable as the sole endpoint. While it provides a convenient measurement of the final tumor, it does not allow for assessments of the activity of the treatment on growth kinetics, being only a "snap shot" of the tumor's growth. For example, a cytotoxic and cytostatic treatment could produce the same reduction in final tumor weight. The cytotoxic treatment could have induced an initial complete remission (disappearance and regrowth of the tumor), implying a possible induction of cell death and the potential for an alternative dosing schedule to produce cure. The cytostatic agent could have decreased the rate of cell proliferation without inducing any significant cell kill. Where there is heterogeneity in initial cell volume, the ability to statistically demonstrate activity also may be compromised. This is less problematic when consecutive measurements are
obtain on each tumor, e.g. to obtain individual $T_2$. It also is possible to estimate tumor weight/volume from caliper measurements of tumor area. Tumor volume can be obtained from area measurements by [71]:

$$\text{tumor weight (mg)} = \left(\text{length}\ \times \text{width}\right)/2$$

where: measurements $= \text{mm}$ for all equations. Where the volume is estimated at necropsy and the area of necrosis can be measured, this is modified to [72]:

$$\text{tumor weight (mg)} = \left\{\frac{\left(\text{length} \times \text{width}\right)^3}{2}\right\} - \left\{\frac{\left(\text{necrosis length} \times \text{necrosis width}\right)^3}{2}\right\}$$

For nonspherical tumors, the volume can be estimated by [73]:

$$\text{volume} = \frac{4}{3}\pi\left(\text{largest diameter}\right)^2 \times \text{smallest diameter}/2^{1/2}$$

Other approaches include measuring three perpendicular dimensions and simply multiplying the three estimates [74]. While obtaining these measurements is often feasible with s.c. tumors in the flanks or back of nude mice, there can be difficulty in obtaining reproducible measurements of height for some tumors. The underlying assumption that the tumors are “box shaped” appears reasonable, since assuming an ellipsoid shape does not appear to give a better estimate, at least for human lung tumors [75].

Where necessary, the ellipsoid volume can be estimated by [76]:

$$\text{volume} = \frac{4}{3}\pi\left(\text{length}/2 \times \text{width}/2 \times \text{height}/2\right)$$

When such tumor volume/weight estimates are used, it is advisable to confirm the validity of the relationship by at least comparing the predicted tumor weights/volumes with the actual wet weights/volumes at necropsy at the end of the experiment. Tumor volume measurements can be used to estimate $T_2$ and growth delay. Data also can be transformed and plotted as changes in relative tumor volume. Relative tumor volume can be obtained from [77]:

relative volume $= V_i/V_0$

where $V_0$ = volume at the start of treatment, and $V_i$ = volume at day $i$.

Excision assays

The main purpose of excision assays is to directly estimate the fraction of cells in a tumor that have retained their clonogenicity (fractional cell survival). The greatest methodological variation in these assays relates to the technique chosen to determine clonogenicity. Probably the more widely utilized clonogenicity assays determine the number of tumor cells, in an ex vivo suspension, able to form anchorage-independent colonies in vitro. Clonogenicity is usually assessed in a semi-solid medium (agar or methyl cellulose). This approach has significant advantages over other excision assays in time, cost, and intra-experimental variability [55], and in reducing the numbers of animals.

The $T_{D_{50}}$ assay determines tumorigenicity in vivo and, therefore, the tumorigenic potential of the treated cell populations. The excised tumor is disaggregated and a dilution cloning technique used to determine the minimum number of cells required to form tumors upon reinoculation. While immunologic properties of the tumor can be problematic, reinoculation into an immune-compromised host can largely eliminate these problems [55]. Aspects, such as the MDA455/LCC6 [21], can be inoculated i.p. and treated i.v., s.c., or p.o. Surviving cells can easily be removed and reinoculated either into the mammary fat pads of recipient mice to assess tumorigenicity as solid tumors, or i.p. to assess tumorigenicity by ascites formation. This approach, while providing a rigorous determination of effects on in vivo clonogenicity, can require relatively large numbers of animals depending on the study design.

The need to appropriately reduce animal usage is likely to effectively eliminate many of these in vivo or in vitro experimental approaches. However, a modification of this approach is to
assess clonogenicity of in vivo treated tumors in vitro. Thus, cells are treated as solid tumors or ascites in vivo, with cells removed and fractional cell survival determined by an anchorage-independent colony formation assay in vitro (above). This has the disadvantage that in vitro clonogenicity is used as a surrogate for tumorigenicity.

Some cells also may not immediately readapt to in vitro growth in a manner that allows for an adequate assessment of cell survival. However, for many cellular models there is no significant effect on estimates of cell survival, but there are advantages in substantially reduced costs and fewer animals required.

All excision assays suffer from the disadvantage that, at some point, cells are removed from the animal for manipulation. Thus, the critical tumor/host relationship is lost. The disaggregation of solid tumors, for either in vivo or in vitro analysis, can result in a cell population no longer representative of the tumor. The process of enzymatic digestion can reduce clonogenic potential and thereby appear to increase the activity of the treatment. The time of excision can also influence the results, since some cells may eventually repair potentially lethal damage [78]. One advantage of an ascites model for these types of studies is that the cells may not require damaging enzymatic and/or physical disaggregation.

Endpoint comparisons

The most important technical disadvantage of the excision compared with the in situ techniques is the necessity for removal and manipulation of the tumor prior to assessing cytotoxicity. The loss of the tumor microenvironment also is problematic, since the ability of normal cells to stimulate tumor regrowth may be a crucial factor in the apparent failure of some cytotoxic regimens. However, excision assays eliminate the problem of remaining dead (or reproductively dead) cells contributing to tumor volume, a concern that can arise with in situ assays. The tumor growth delay assays have the disadvantage that the kinetics of tumor regrowth may be unpredictable for novel drug combinations. Tumors that regrow may have altered proportions of infiltrating normal cells that could cause an overestimation or underestimation of cell kill. The variability of inter-tumor regrowth patterns also can be problematic. Nevertheless, the in situ assays generally appear to be favored in the literature.

Death/survival as an endpoint

Many institutional animal care and use committees are restricting or eliminating death as an endpoint in drug screening and other in vivo studies. Nevertheless, some investigators may have experimental designs where death/morbidity/survival is a justified requirement of the study. Several related issues require consideration. Survival can be measured at a fixed time, e.g., the proportion of animals remaining alive at a predetermined time point beyond when the last untreated animal dies, or the median duration of survival when all animals in all groups die within the observation time. When "cure" or the proportion of "long term survivors" are estimated, it is necessary to define the time point at which "cure" or "survival" is attributed. A major concern is the T_50 since sufficient time must be allowed for any significant number of remaining cells to proliferate to the point where a palpable tumor/ascites would be expected. A solid tumor with T_50 = 48 hr may require up to four months of post treatment observation to establish "cure" [32]. Defining long term survivors (where death/morbidity is the primary endpoint) often is based on the duration of survival of mice bearing untreated tumors, e.g., three times their mean or median survival. By this criterion, with a mean survival of 30 days in the untreated group, treated mice that survive for 90 days could be considered long term survivors [21]. When there are survivors, the data analysis procedures need to take this into account [68].

There is little ethical justification for using
death as an endpoint for solid tumors that are easily accessible for estimations of tumor growth delay or excision assays. In general, survival as an endpoint should probably be restricted to ascites models, and perhaps also to solid tumors known to achieve a lethal tumor burden within an appropriate and predictable time. Even for these tumors, it may be possible to substitute morbidity for death. We have found this to be a viable and more humane alternative in several ascites studies. While it requires knowledge of the time from the onset of morbidity to death, and evidence that this period is sufficiently consistent, this information can be obtained, in advance, on a relatively few animals.

Where possible, morbidity should be considered as a potential surrogate for death as an endpoint in survival analyses. To assist other investigators, the criteria used to define morbidity, and the verification of its applicability as a surrogate for death, should be reported.

Approaches to data analysis

Estimation of tumor doubling times

Estimates of $T_D$ can be obtained from measurements of either tumor area or tumor volume. A detailed description of this approach can be found in the accompanying article by Rygaard & Spang-Thomsen [70]. The $T_D$ for each individual tumor can be obtained, with those within an experimental group combined for further analysis, e.g., use of an appropriate ANOVA to explore differences among different treatment groups. While there is evidence of investigators using simple linear regression models to estimate $T_D$ from growth curves, this is potentially confounded by the Gompertzian nature of tumor growth kinetics [11,12,70]. However, in short term studies, particularly where the growth data approximate exponential growth, the Gompertzian model may reduce to an exponential growth model, which may provide a more efficient model for the data [68]. There also can be statistical concerns where early deaths limit the data available for the adequate determination of a $T_D$ [68].

Cell kill estimates from in situ studies

There are a number of methods for estimating cell kill from in situ analyses. For ascites models, estimations of percent increased life span (%ILS) provide an indication of therapeutic activity. %ILS can be estimated from:

$%\text{ILS} = 100 \times \frac{C}{C_{\text{ref}}}$

where $C = \text{median survival time of treated populations}$ and $C_{\text{ref}} = \text{median survival time of untreated control population of tumor bearing mice}$ [67]. It should be noted that median survival times focus on the 50% survival estimates, and do not efficiently use all of the data in the survival curve. This can become problematic when there is substantial heterogeneity in the duration of survival in some treatment groups. One approach is to use the hazard ratio (HR) of death for the control group versus the treated group (J Hanfelt, personal communication). Thus, where there is considerable heterogeneity in survival data, %ILS may be better estimated from:

$%\text{ILS} = 100 \times \frac{HR-1}{HR}$

where $HR = \text{hazard ratio of death for the control group versus the treated group}$. The HR can be obtained by Cox proportional-hazards regression analysis [79,80].

Estimates of cell kill can be derived from solid tumor growth curves, where repeated measurements are obtained over the period of time required for tumors to reach a predetermined size. Provided the $T_D$ for untreated and regrowing (treated) tumors are equivalent, the cell kill can be estimated by a number of related formulae [81]. Cell kill can be determined from growth delay measurements using an estimate of the $T_D$ [9] and the number of cell divisions required for each log increase in growth [3.32] [82].

Total cell kill can be estimated from [81]:

$\log_{10} \text{cell kill} = (T_C-3.32) / T_D$
Cell kill/dose can be estimated from [81]:

$$\log_{10} \text{cell kill/dose} = (T-C) \cdot (3.32) \cdot (T_D) / n$$

where: $T =$ median time to predetermined size in treated populations; $C =$ median time to predetermined size in untreated control populations; $n =$ number of treatments.

Net cell kill can be estimated from [81]:

$$\log_{10} \text{net cell kill} = (T-C) - \text{duration of treatment} / (3.32) (T_D)$$

The specific growth delay for solid tumors can be estimated from the times taken for both treated $(T)$ and control $(C)$ tumors to reach a predetermined size [66]. If the specific growth delay value exceeds 2, the tumor model is often considered responsive. This is generally considered the most important estimate of antitumor activity, and can be reported along with the total cell kill estimated from the same primary data (above).

Specific growth delay can be estimated from [77,83]:

$$\text{Specific growth delay} = (T-C)/C$$

Where the $T_D$ of tumor regrowth is significantly different from the $T_D$ of untreated tumors, these estimates are invalid. For example, the $T_D$ of treated populations may be slower than equivalent untreated tumors following irradiation [69]. Other limitations of these analyses are indicated when treatment is begun shortly after tumor cell inoculation, where a significant immune component is suspected, or where the treatments are close to being curative [72]. The degree of response in tumors that exhibit alterations in the $T_D$ following treatment can be more appropriately expressed as delay/doubling time [69].

**Determination of synergy**

Studies are often performed to determine the nature of the interaction between two or more treatments. Many investigators claim synergistic interactions without having performed the necessary analyses. Several authors have more clearly defined these terms and described the conditions required to determine whether an interaction is synergistic, additive, or antagonistic [84-86].

Synergism is a mathematically defined interaction, most widely determined from isobologram analyses. Construction of classical isobolograms (isoeffect curves) for the determination of additivity, synergy, and antagonism generally requires an initial estimation of the dose response curve of each agent alone and dose responses of one agent in the presence of each of several concentrations of the second agent [72]. While this may be feasible for routine in vitro studies [84], classical isobologram analyses generally require too much information to be readily applicable to in vivo data. A minimum requirement for in vivo studies has been suggested, using five doses of each drug alone and three or four combination treatments using intermediate doses [87]. An interaction index $(I_I)$ can then be estimated from the isobologram equation [88,89]:

$$I_I = (D/D) + (T/T)$$

where $ED_2 =$ some fixed level of activity, $D$ and $T =$ dose of each drug alone required to produce $ED_2$, and $D$ and $T =$ dose for each drug in combination that produce $ED_2$. $I_I$ values of $c1$ indicate synergy, a value of 1 reflects additivity, while values $>1$ indicate antagonism. Ideally, each $ED_2$ value should be estimated by performing appropriate probit or logit analyses of the dose response data [90,91].

An alternative experimental design is to fix the dose of the first drug and vary the concentration of the second [92]. This approach has not been widely applied to in vivo studies, and its ability to adequately define the nature of drug interactions does not appear to have been extensively confirmed. Nevertheless, it could be applied to in vivo studies and may be worthy of consideration.

Application of the median effect analysis is one of several alternative approaches to the classical isobologram approach for assessing synergy [85]. While this method may be amenable for use in in vivo studies, this has yet to be firmly
established. Other related approaches to assessing synergy also are reported [93] but their application to in vivo studies also remains to be confirmed. The concern is not the validity of the approaches, which is beyond the scope of the current article, but how appropriate in vivo experiments could be designed, within the constraints of time, cost, and appropriate animal usage, to produce sufficient data to generate statistically reliable estimates of the nature of the drug interactions.

Where one drug has no activity, synergy is more readily determined [93] and may simply require statistical evidence that the combination is significantly different from the inhibitory drug alone. This simplistic approach is potentially confounded when the "inactive" compound is active at higher concentrations. If the dose used was at a theoretical ID_{50}, i.e., activity was present but could not be detected because of the sensitivity of the assay, assumptions of synergy could well be invalid. Thus, if this approach is to be used, it should be at least restricted to doses well below the minimally active dose, or better yet, restricted to studies where one compound is known to have no antineoplastic activity.

Where resources are available, isobologram-based approaches are preferred. However, some interactive definitions may be approximated from limited data as follows [94,95]:

**Antagonistic**
\[(AB)/(C) > (A/C) x (B/C)\]

**Additive**
\[(AB)/(C) = (A/C) x (B/C)\]

**Synergistic**
\[(AB)/(C) < (A/C) x (B/C)\]

where A = response to treatment 1; B = response to treatment 2; C = response to no treatment/vehicle; AB = combination of treatments A and B. Some investigators define a fourth interaction category (subadditive) as \[(AB)/(C) > (A/C) x (B/C)\] < [B/C] (where B/C > A/C). However, a sub-additive interaction is essentially antagonistic, making the utility of this term somewhat unclear from a pharmacologic perspective.

The application of this approach may vary depending on the endpoints and experimental design chosen for a particular study. For excision assays using in vitro colony formation as an endpoint, the response to treatment would be represented by the colony forming ability or surviving fraction. Growth delay assays could utilize the estimates of cell kill/survival described above or T_{50} parameters. A description of approaches to the statistical analysis of data obtained from these studies can be found in the accompanying article by HanSich [96].

These terms are approximations, at best, when compared with more classical isologram approaches, and should be used with some caution. The outcome provides only a general approximation of the nature of the interaction, and should not be considered definitive. While reasonable approximations may be obtained where tumors approximate logarithmic growth, the analysis does not take into account the shape of the respective dose response curves, a central component of isologram approaches for determining the nature of drug interactions. This could lead to both underestimation and overestimation of potential synergy or antagonism.

Therapeutic synergism has been defined as occurring when the response of a combination is in excess of the maximum response of either drug alone (e.g., A = 100%, B = 150%, A+B = 200%) at equitoxic doses [67,96,97]. The term "clinical synergism" also has been applied when a combination chemotherapy regimen is curative [96]. These are not particularly informative definitions from an interpretive or mechanistic perspective, and could be applicable when the "true" biochemical interaction is only additive or even antagonistic.

In practice, a drug combination can be curative and yet produce a nonsynergistic interaction. If two drugs with different mechanisms of action each kill 10^{5} cells, a tumor of 10^{6} cells could still be cured by an additive or even antagonistic interaction [96]. Thus, the determination of synergy or antagonism is not required where the purpose is simply to show that a combination of two or more drugs is curative relative to either agent administered alone.
The toxicity of a particular combination should be considered when determining the nature of an interaction between two or more drugs. In order to make appropriate comparisons, the treatments should be approximately equitoxic [97]. For example, the improved antitumor activity of a combination regimen which is accompanied by significantly increased toxicity might be no more effective than a higher dose of either agent alone that produced equivalent toxicity.

Some suggestions for statistical approaches to data analysis and the assessment of activity

The statistical analytical procedures used to explore data from animal studies is dependent on the study design. Statistical approaches will be dealt with only cursorily, since a detailed discussion of methods and approaches is provided elsewhere in this issue [68]. Readers are encouraged to consult this and other articles which deal with endpoints and considerations in depth beyond the scope of this section. While the use of some specific tests is suggested below, these should be used with some caution, since the primary data may violate assumptions implicit within the analyses. For example, some of these tests described below require that the data approximate a normal or other distributional form. When this is not apparent, or cannot be achieved by transformation, e.g. converting to log_{10} nonparametric analyses may be required. For investigators unfamiliar with statistical analyses, there is no substitute for consultation with a biostatistician.

One of the key first steps in the design of animal experiments should be the determination of appropriate group sizes. The number of animals per group will depend on the endpoint and the magnitude of the expected response. In this regard, it is necessary to perform appropriate power estimates to ensure that the design will have sufficient statistical power to adequately identify significant or nonsignificant differences. This is discussed in some detail in the accompanying article by Hanfelt [68]. It is important to ensure that group size is sufficient to maintain power when some animals may either die from other causes, e.g. drug toxicity or secondary infection, or bear tumors that do not exhibit appropriate pretreatment kinetics. For pilot exploratory studies, where more definitive follow-up studies are planned, this is of lesser concern. The data from such pilot studies can often provide useful information on the expected response rate and toxicity, thereby facilitating the generation of potentially more relevant power estimates.

Tumor growth delay is generally assessed in terms of the time necessary for tumor growth in each group to increase from the initial size at the time of treatment to a larger, predetermined size [66,81,98]. The most common approaches are to compare the T/C or T-C values. Both parameters are usually based on median values, with zeros included for the T/C estimates and "tumor-free cures/tumor-free survivors" being excluded from the estimates and analyzed separately. A T/C value 542% is generally required to demonstrate activity [67]. The necessity to exclude cures may be problematic if it affects the power of the analysis. As discussed by Hanfelt [68], the use of median times also may be inefficient, with longitudinal analyses providing a more effective use of the data.

There are several other ways to explore data from tumor growth delay studies. For example, survival analyses also can be used to assess "time-to-event" endpoints, e.g. time to reach a predetermined size. A further approach is to use a repeated measures ANOVA to compare tumor size at each time point across the analysis. This may be most applicable if a significant number of tumors in a treatment group do not reach the predetermined size in the control group. Alternatively, tumor doubling times can be estimated for each individual tumor in each group by applying Gompertzian kinetic analyses, such as those performed by the GROWTH software [73]. Tumor doubling times can be compared among experimental groups by either ANOVA or multivariate ANOVA [99]. This approach also is useful for
cytostatic treatments, e.g., hormones or anti-hormones, and to ensure that the kinetics of regrowth are appropriate for cell kill estimates.

For tumor excision assays, in vitro colony forming ability among groups, or in vivo tumor incidence if reseeded into recipient mice, can be compared by either ANOVA or multivariate ANOVA.

Percent increased life span (%ILS) is the most widely used activity measure for ascites tumors. Since some animals may survive, survival curves may be estimated using the Kaplan-Meier approach [100], and differences in survival between treatment groups estimated by the Log-Rank test [101]. Where early deaths occur from drug toxicity and later deaths from tumor burden, a model more sensitive to early events may be necessary [68]. In general, a %ILS >27% in the P388 ascites model is considered the minimum for activity when both drug and tumor are administered i.p. When drugs are administered i.v. to an i.p. tumor, a %ILS >40% is considered sufficient to demonstrate activity [67].

Tumor incidence, proportion of survivors, or long-term survivors, e.g. 2x2 analysis using single treatment and control, can be compared among groups by \( \chi^2 \). If more than one site is used per mouse, General Estimating Equation methods are required to account for any lack of independence of tumors within animals [68]. If the number of observations is small, a Fisher's exact test or Pearson's \( \chi^2 \) test may be used.

Body to tumor weights, where the data are continuous and randomly distributed, can be compared by ANOVA followed by a multiple comparison post hoc test such as Dunnett's multiple range test [102]. Where it is required to merely compare several individual groups with the same control, e.g. dose-response analysis, and where the group sizes are equal, Dunnett's t-test can be used [103]. Where group sizes are different, Scheffe's multiple comparison test can be applied [104]. However, it should be noted that some endpoints also are associated with several variables. For example, organ weights increase with increasing body weight. For such endpoints, analysis of covariance approaches are required [105].

**Conclusions**

As indicated in the introduction, many of the issues are raised to assist investigators in approaching some of the major concerns that arise in designing studies to test the activity of cytotoxic agents. There are alternative approaches to several of these issues, but a detailed discussion is beyond the scope of the current article. Some are described in detail elsewhere in this journal issue, and others can be found in many of the publications cited herein. While the relative importance of many of the topics addressed will vary with the hypothesis and endpoint chosen, some issues will apply irrespective of these, e.g. the need to design experiments with sufficient statistical power.

**Acknowledgments**

The author wishes to thank Dr. John Hanfelt, Fabio Lenoness, and Robert Clazer for critical reading of the manuscript. This work was supported by research grant DAMD17-96-1-5231 from the U.S. Department of Defense, U.S. Army Medical Research and Materiel Command.

**References**

4. Howell A, DeFriend DI, Robertson JFR, Blamey
57. Ueda K, Clark DP, Chen CJ, Rominon IB, Gotten-
man MM, Passan J. The human molding resistance
(MXR) gene. C-Dean and transcriptional regu-

58. Norton IK, Therasse KN, Houghton IA, Horowitz
ME, Houghton PJ. Modulation by verapamil of vin-


60. Bollag W, Wearnau A, Polin L, Capps D, Leopold
WR, Werbel LM, Biart L, Dan ME, Corbett TH. Antitumor efficacy of PD 159343 (NSC 566146) against solid tumors of mice. Cancer Res 50:4993-
4995, 1990


63. Volpe JG, Bullen L. Influence of tumor trans-
plantation method on tumor growth rate and meta-


66. Yamamoto S, Fruh K, Kanan-T, Nozue S, Namba M. In vivo studies on potential of cyto-


70. Rygaard K, Sjost-Rotmer M. Quantification and


75. Cleland CA. Tumour volume. Br J Cancer 73:464-
465, 1997


81. Corbett TH, Valenstein PA. Rodent models in experi-


---


