Trials, Tribulations, and Trends in Tumor Modeling in Mice

JoAnn C. L. Schuh
Applied Veterinary Pathobiology

Follow this and additional works at: https://www.wellbeingintlstudiesrepository.org/toxramres

Part of the Animal Experimentation and Research Commons, Animal Studies Commons, and the Other Medical Sciences Commons

Recommended Citation

This material is brought to you for free and open access by WellBeing International. It has been accepted for inclusion by an authorized administrator of the WBI Studies Repository. For more information, please contact wbisr-info@wellbeingintl.org.
INTRODUCTION

Malignant neoplasms rank second as the leading cause of death in the United States and ranked first in those aged 45–74 (Anderson, 2002). As a result, anti-cancer therapies are a frequent focus for startup companies and represent major therapeutic classes for pharmaceutical, biopharmaceutical, medical device and drug delivery manufacturers. Despite the effort applied to cancer targets, the number of successful new therapies for treating human malignancies is discouragingly low. This is surprising in that many trials are now conducted using novel agents with specificity for molecular pathways and cellular components rather than broad targeting of chemotherapy and radiation to normal and neoplastic cells. Failures in clinical trials are multifactorial with lack of efficacy an important cause. Conversely, modulating and curing experimental cancer in mice is a relatively easy process. Many commonly used mouse models of neoplasia have proven to be biased towards false positive results and preclinical studies have not accurately predicted clinical responses. Therefore, unconditional acceptance of limited data from mouse models has to be avoided to prevent premature movement of development programs into clinical testing. Increasingly, the plethora of novel strategies undergoing testing requires greater attention to proper design and conduct of preclinical efficacy studies. Rationale design of preclinical efficacy studies requires understanding the biology of tumors and implantation techniques, selection of in vivo and ex vivo endpoints, and a willingness to integrate new and often costly testing strategies that more appropriately mimic the biology of human neoplasms.

BIOMEDICAL MODELS OF NEOPLASIA FOR PRECLINICAL TESTING

Spontaneous and Environmental Carcinogenesis Models

Historically, spontaneous, chemical, ultraviolet (UV), oncogene, and viral infection models helped to define many aspects of carcinogenesis (Harrison, 2002) and therapeutic intervention (Boone et al., 1992) and helped to promote development of inbred strains of mice (Corbett et al., 2002). Despite the significance of spontaneous and environmental models to biomedical research, the long latency of most of these models makes them impractical for most preclinical studies of tumor modulation. Spontaneous, chemical and UV and viral infected or transformed tumors are of greatest importance as the source of many cell lines used for in vitro studies and in vivo transplantation models.

Transplantation Models

There are many immortalized cell lines of human and murine origin available from commercial sources and privately held by research organizations that have been tested for tumorigenicity in mice (Giard et al., 1973; Gershwin et al., 1977; Fogh et al., 1977; Trainer et al., 1988). In addition to availability, these tumorigenic cell lines are generally easy to maintain, selectable for unique mutations in vitro and backed by numerous publications on in vivo behavior in immunodeficient (nude, beige, nude/beige, C.B-17 severe combined immunodeficient [SCID], nonobese diabetic [NOD]/SCID), immunosuppressed (thymectomized or corticosteroid treated), humanized (hu)-SCID or hu-NOD/SCID and immunocompetent strains of mice. However, quality control is an issue. Many cell lines have undocumented source and passage histories, poorly characterized receptor and oncogene expression and cellular secretions, and inconsistent designations in publications. Features that make these tumor lines suitable for transplantation may affect experimental design...
because of the need for specific mouse strains, sex specificity, and altered host immunity. Furthermore, spontaneous mutations in vitro that allow selection of tumor cell line subclones with unique behaviors can also result in point mutations that lead to changes in histomorphology, sensitivity, and behavior of these tumors in vivo. Dissociated solid tumors and fragments of tumors (explants) that retain histomorphology relationships of the tumor and associated stroma are also suitable for implantation (Gershwin et al., 1977). Tumor cells from cell lines and solid masses of both human and mouse origin may require handling as biohazardous material due to their histogenesis by viral transformation or inadvertent contamination (Hay, 1991; Nicklas et al., 1993). The literature on tumors that have been used in mice to model important therapeutic targets in humans is vast. Due to the rapidly changing nature of biomedical research, a thorough and current literature search is warranted prior to selecting the most appropriate models for any therapeutic development program. A selection of representative mouse models (transplantation and genomically) for malignant neoplasms causing death in humans is provided in Table 1.

Criteria other than historical usage, and availability of the tumor line and a suitable mouse host need to be considered when testing efficacy of tumor modulation with therapeutic (Skipper, 1968; Huseby, 1969; Klausner, 1999). Even with genetically engineered models, multistep progression and clonal derivation of tumors in humans are difficult to model in mice. Therefore, foibles of these models need to be understood to prevent overinterpretation of positive or negative results.

**CONSIDERATIONS IN SELECTING TUMOR MODELS IN MICE**

It is well recognized that our understanding of modeling of tumor biology and therapeutic intervention is constrained by many factors (Siemann, 1987; Rew, 2000a; Rew, 2000b). However, validation studies of these model systems for their ability to adequately predict therapeutic responses in patients have been rare (Hann and Balmain, 2001). Consequently, design and interpretation of preclinical studies for tumor modeling must be undertaken carefully.

**Study Design**

Candidate anti-tumor agents can be identified and selected using a broad panel of in vivo tumors (Atassi et al., 1988) but false positive and negative results may occur due to incompatible host-tumor-therapeutic interactions or technical incompetence. In efficacy studies, greater depth needs to be achieved through evaluation of at least several subtypes of the tumor representative of the clinical target, and through use of different models systems in mice (transplantation, genetically engineered, and orthotopic models) as well as use of models in other appropriate species. Conduct of animal studies that mimic expected exposure, scheduling, and duration of therapeutic and posttreatment periods of clinical studies are valuable designs prior to initiating clinical trials. Optimization of therapeutic dose and schedule through pharmacokinetic studies are also important early procedures that should be evaluated in one or more model systems. Efficacy studies should also incorporate histopathology endpoints to confirm expected therapeutic target and to help refine dose scheduling relative to growth and cell loss fractions.

Investigators need to be aware that many historical and commonly used model systems in mice were originally established and optimized for use in mechanistic studies (Burger, 2000). Conduct of efficacy testing in such systems has not been optimized and inadvertent selection of systems with excess curability, spontaneous regressions or failure to establish adequate tumor burden (Corbett et al., 2002) can lead to overly optimistic projections of clinical success. Bias in the selection of the model(s) may also arise from experience, knowledge base and objectives of the investigator. Whereas pharmacologists, immunobiologists or cell biologists may consider direct intratumoral injection of therapeutics acceptable, needle tracks and pressure-induced necrosis (compartment syndrome) may interfere with adequate evaluation of such models by a histopathologist. Conversely, intratumoral administration of light activated substances is a common and appropriate route for photodynamic therapy of tumors (Casas et al., 1999). Design of efficacy studies for combination protocols is often difficult. However, such studies can provide important efficacy data for the transition between novel therapies and standard treatment practices that may

**Table 1.**—Rank order of deaths due to malignant neoplasms in the United States in 2001 (Arias and Smith, 2003) and selected published reviews on applicable mouse models of these neoplasms.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Malignant Neoplasms</th>
<th>Cases</th>
<th>Selected Mouse Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Lung—trachea</td>
<td>156,005</td>
<td>Tuveson and Jacks, 1999; Malkinson, 2001; Liu and Johnston, 2002</td>
</tr>
<tr>
<td>2</td>
<td>Colon—rectum</td>
<td>56,799</td>
<td>Heye et al., 1999; Koback-Larsen et al., 2000; Horig et al., 2001; Boivin et al., 2003</td>
</tr>
<tr>
<td>3</td>
<td>Lymphoid—hematopoietic</td>
<td>56,350</td>
<td>Dykes and Waud, 2002; Uckun and Sensel, 2002; Vanderkerken et al., 2003</td>
</tr>
<tr>
<td>4</td>
<td>Breast</td>
<td>41,844</td>
<td>Hutchinson and Muller, 2000; Cardif, 2001; Rosner et al., 2002; Clarke, 2002</td>
</tr>
<tr>
<td>5</td>
<td>Prostate</td>
<td>30,714</td>
<td>Royal et al., 1996; Navone et al., 1999; Abate-Shen and Shen, 2002; Nyska et al., 2002</td>
</tr>
<tr>
<td>6</td>
<td>Pancreas</td>
<td>29,723</td>
<td>Fu et al., 1992; Hotz et al., 2000; Standop et al., 2001; Bardeesy et al., 2001</td>
</tr>
<tr>
<td>7</td>
<td>Ovary</td>
<td>14,361</td>
<td>Rahman et al., 1998; Rahman and Hultaniemi, 2001; Orsulic et al., 2002</td>
</tr>
<tr>
<td>8</td>
<td>Liver</td>
<td>13,263</td>
<td>Fausto, 1999; Feitelson and Larkin, 2001</td>
</tr>
<tr>
<td>9</td>
<td>Brain—meninges</td>
<td>12,567</td>
<td>Holland, 2001; Reilly and Jacks, 2001; Beigmann et al., 2002; Gutmann et al., 2003</td>
</tr>
<tr>
<td>10</td>
<td>Esophagus</td>
<td>12,509</td>
<td>Opitz et al., 2002</td>
</tr>
<tr>
<td>11</td>
<td>Stomach</td>
<td>12,340</td>
<td>Furukawa et al., 1993b</td>
</tr>
<tr>
<td>12</td>
<td>Bladder</td>
<td>12,115</td>
<td>Eto et al., 2000; Bonfill et al., 2002</td>
</tr>
<tr>
<td>13</td>
<td>Kidney</td>
<td>12,084</td>
<td>Naito et al., 1987b; An et al., 1999; Hillman, 2002</td>
</tr>
<tr>
<td>14</td>
<td>Oral</td>
<td>7,638</td>
<td>Waters et al., 1998; Myers et al., 2002</td>
</tr>
<tr>
<td>15</td>
<td>Skin</td>
<td>7,543</td>
<td>Alvarez, 2002; Eccles, 2002; Carson III and Walker, 2002</td>
</tr>
<tr>
<td>16</td>
<td>Uterus</td>
<td>6,835</td>
<td>Couse et al., 1997; Keshavarzi et al., 2002</td>
</tr>
<tr>
<td>17</td>
<td>Cervix</td>
<td>4,064</td>
<td>Herbst et al., 1996</td>
</tr>
<tr>
<td>18</td>
<td>Larynx</td>
<td>3,826</td>
<td>Chen et al., 2001; Kennel et al., 2002</td>
</tr>
</tbody>
</table>
include a combination of surgery, chemotherapy or radiation, and other supportive therapies (Bogden et al., 1974; Corbett et al., 1979). In designing tumor model systems, investigators must guard against template designs that ignore inherent differences in model systems, and in using insufficient animals per group. This is particularly important in investigational new drug (IND) enabling studies where more animals per group and few groups may be required to adequately control for variation in tumor burden and therapeutic responses. Consultation with a biostatistician to assist in determination of sample size appropriate to each model system is highly recommended.

**Context of the Study and Available Resources**

Therapeutic intervention conducted in tumor-bearing animals may occur as proof of principal studies for discovery, target validation, and in vivo pharmacology, or as investigations for therapeutic efficacy, safety, and interactions. These types of studies often need to be designed and conducted differently than studies focusing on mechanistic studies. Additionally, an available supply of suitable strains of mice and investigators capable of providing necessary manipulations (injections, surgery, polytherapy, evaluations, and pathology services) may materially affect both quantity and quality of data that can be generated from selected models.

**Lack of Quality Control**

Investigators using cell lines should develop and maintain in-house databases that allow ready access to known features for tumor cell lines in their collections. A minimum list might include all applicable information that defines origin (human or animal), sex and strain of original host, treatment history (particularly human lines), name and subclone of the tumor line, source (commercial or private), passage history of all aliquots, species, strain and sex and routes susceptible to implantation, in vitro and in vivo growth characteristics, growth rate to maximum humanely accepted and lethal sizes, histomorphological characteristics, metastatic potential and method, immunogenicity, receptor and oncogene expression, cellular products, microbiological screening history, biohazard potential (generally viral) and unique characteristics. This information is invaluable for preventing and investigating problems such as implantation failures, excess curability, microbial contamination, mislabeled or cross-contaminated stocks, and alterations in histomorphology and behavior that may arise from haphazard use of cell lines (Corbett et al., 2002). With rare exception (Rosner et al., 2002), thorough characterization of comparative histology of tumor models using authenticated stocks has not been adequate. Additionally, microbial and cross-species contamination of cell lines has lead to erroneous conclusions in tumor biology (Moseley et al., 2003; Drexler et al., 2003). These problems are compounded by lack of source and passage history information reported by most laboratories conducting either mechanistic or efficacy studies. This is not a trivial problem, as the histomorphology and behavior of tumors grown from stocks in different laboratories may eventually differ from historical photomicrographs and behavior descriptions of the original tumor and its cell line (Schuh, unpublished data). Although the original passage history may not be known, investigators should strive to use contaminant free tumor lines of a consistent number of passages from an authenticated stock for preclinical testing. In addition to providing a known in-house passage history, testing from a stock provides a baseline and helps to reduce the potential for point mutations that may cause inter-study variation or erroneous test results.

**Paraneoplastic Syndromes**

Paraneoplastic syndromes typically result in clinical manifestations of altered physiological responses to autotchthonous (spontaneous) neoplasms. These syndromes in mouse models of neoplasia have been seldom described (Liebelt et al., 1974) even though they exist. This oversight is brought on by the focused and short-term nature of most studies. Unlike safety studies where a complete set of tissues are collected and a clinical pathology examination conducted, efficacy studies are often completed without benefit of histopathology and clinical pathology examinations. This type of study design neglects important interactions of the host with tumor receptor expression and secretions produced by tumor cells or by stimulation of host cells by the tumor. Paraneoplastic syndromes include extramedullary hematopoiesis, bone marrow hyperplasia, peripheral granulocytosis and leukocytosis (leukemoid reactions), thrombocytosis, anemia, altered lipid metabolism, hypercalcemia of malignancy, hypoglycemia, cachexia and organomegaly in nontumor-bearing tissues (Liebelt et al., 1974; Castillo et al., 1982; Yoneda et al., 1991; Tanaka et al., 1996; Diament et al., 1998, Schuh, unpublished data). Paraneoplastic syndromes represent potential models for similar syndromes in humans but causation are poorly characterized and effects of these syndromes on pharmacology, safety, and efficacy studies using experimental tumors are unknown.

**Transplantation Protocols: Sites of Implantation Can Affect Study Outcome**

Autogeneic or autotchthonous tumors are seldom practical for tumor modeling for therapeutic intervention. Most transplantable tumors are placed heterotopically (ectopically) in syngeneic (same species, genetically identical), allogeneic (same species, genetically different) or xenogeneic (different species and genetics) host systems. Tumor lines in use have been specifically selected for mutations that allow heterotopic growth in mice. Although these tumors will grow and respond to therapeutics, heterotopic sites are not ideal and selection of the transplantation site may modulate tumor growth (Naito et al., 1987a; Corbett et al., 2002) and success of therapeutic intervention (Averbook et al., 2002).

Subcutaneous (SQ) and less frequently intradermal areas are used for primary tumors for reasons of accessibility, lack of distress and interference with mobility in mice, and visibility for monitoring. Generally, SQ refers to placement by injection or surgical implantation in the flank, a region referring to the posterior lateral abdominal quadrant. Some investigators erroneously include the hindlimb, back and axillary region in their description of the flank. Placement is generally done in fat and mammary gland tissues near popliteal, inguinal or accessory axillary lymph nodes. Despite the common use of SQ sites, it is important to note that even large tumors rarely, if ever, metastasize after implantation in this site (Gershwin et al., 1977; Eccles, 2002). Implantation in the hindlimb (including popliteal lymph node
fat) is sometimes considered superior to placement near body cavities for improved visualization and avoidance of accidental intra-peritoneal or thoracic implantation. Rarely, tumors may not be viable after SQ implantation and other sites may be required. Implantation of tumors into footpads is generally not acceptable for humane considerations (UKCCCR, 1998; Wallace, 2000). Intramuscular implantation is not common as cell lines and tumor explants may not grow well in muscle compared to SQ, leg mobility is restricted by large masses, expansion space is limited, measurements with calipers are more difficult and this site appears to be more prone to self-mutilation and a target for cage mate aggression. Fat pads other than classic SQ sites, including retroperitoneal, epididymal, intrascapular and mediastinal/thymic sites are also valuable locations for injection and surgical implantation to provide a highly vascular milieu that appears to assist in establishment of xenogeneic tumors. A comparison of tumor viability in brown fat in interscapular areas, compared to the predominate white fat in other fat pads does not appear to have been made.

Primary and sometimes metastatic models are also established by direct surgical or percutaneous intra-organ injections into spleen, liver, lymph nodes, mammary gland, prostate, base of the tongue (head and neck carcinoma), intraluminal (bronchial, urinary bladder, thorax), cecal wall implantation (colorectal metastases), and kidney capsule. Injection into brain and eye are often not considered acceptable by institutional animal care and use committees (IACUC) (UKCCCR, 1998; Wallace, 2000).

Metastatic and sometimes primary tumor models are created by intravenous, intracardiac (ventricle), intraosseous injections or intravascular injections proximal to organs (e.g., portal vein for liver neoplasms). Despite careful preparation to reduce cell clumps and to inject slowly, such tumors are embolic and may grow in unintended sites. Most tumors show selective tissue tropisms, but tumors with broad tropism (e.g., lymphoma/leukemia) may grow within all tissues and result in large unmonitorable tumor burdens (Schuh, unpublished data). Investigators rarely evaluate tissues for metastases outside of their specific area of interest (most often lung and bone) so that characterization of additional tumor burdens and thromboembolism of tumor cells on validity of these tumor models is lacking. Intraperitoneal injections of tumors, once a common route for leukemias and metastatic models, have humane considerations that need to be considered (UKCCCR, 1998; Wallace, 2000).

Implantation techniques must be practiced, as accidental injection into muscle masses or visceral cavities, and postsurgical inflammation can cause intra-study variability and lack of interstudy reproducibility. Migration through SQ tissues or leakage of cell suspensions is generally not an issue with proper technique and minimal volumes injected with a small gauge needle. Interstudy reproducibility of implanted tumor burdens may be affected by extent of cell separation, particularly for disruption of in vitro cultures or solid tumors (An et al., 1999). In some efficacy studies, surgical implantation of hollow fibers (Sadar et al., 2002), matrigel and discs (Eccles, 2002), polymers (Righi et al., 2003), liposomess (Kunstfeld et al., 2003), and transparent windows (Dellian et al., 1996; Li et al., 2000; Jain et al., 2002) can provide a tightly contained tumor environment to optimize reproducibility for certain endpoints such as angiogenesis. Inflammation and interference with tumor biology must be considered when utilizing surgically implanted devices for containment of tumor masses.

Other Host-Tumor-Therapeutic Interactions

Differences in tumor burden potential, angiogenic and therapeutic response have been shown to be due to differences in strains of mice tested (Naito et al., 1987b), primary implantation site and interval between tumor implantation and therapeutic manipulations (Wilmanns et al., 1992; Chakrabarty et al., 1994; Averbook et al., 2002; Monsky et al., 2002), and metastatic microenvironments (Averbook et al., 2002; Seki et al., 2003). For xenogeneic and syngeneic tumors that require immunodeficient mice, differences in immunological defects between these strains should be understood. Similarly, immunogenicity of tumors in immunocompetent hosts may be an important modifying factor. Other factors that can affect tumor growth in vivo include concurrent manipulations such as surgery, radiation or concurrent treatments that produce inflammation and inhibit establishment of tumor cells. Conversely, stress and seasonal effects may increase tumor burden and distribution (Giraldi et al., 2000). Finally, for long standing tumors, tumor-induced cachexia may suppress tumor growth and enhance efficacy (Laster et al., 1961; Chakrabarty et al., 1994; Mukherjee et al., 2002), similar to delays in tumor development found with intentional caloric restriction in rodents (Sutie al., 2003). Therapeutic efficacy is also modulated through tissue specific and immunologic modifiers, and differences in drug metabolism and disposition (Gershwin et al., 1977; Naito et al., 1987a; Wilmanns et al., 1992; Averbook et al., 2002; Seki et al., 2003).

Specific strains of mice used for tumor models are usually matched to origin of the transplanted tumor. Where multiple tumor cell lines are available to model certain tumors, unique characteristics of these tumors such as host origin, oncogenes, receptors and secretions, and tumor stage may be an important consideration. However, these parameters are often difficult to match to comparable stages in humans (Harrison, 2002). With these limitations in available models, selection of or reporting only on tumors that are dramatically and positively affected by test therapeutics should be avoided. Testing therapeutic interventions on similar tumor types derived from multiple cell lines provides a test situation that partially addresses heterogeneity of tumors in humans. That some of these tumors may not be modulated by the test therapeutic should be expected, and these nonresponders should not be treated as preclinical failures or dismissed. Rather, both negative and positive tumor responses can provide investigators with a more realistic expectation of therapeutic potential in humans, and the totality of the response across several models provides superior insight into activity of therapeutics.

ENDPOINTS AND EVALUATION CRITERIA FOR TUMOR MODELS IN VIVO

Critical host-tumor-therapeutic interactions, dose-response, treatment protocol design, and selection of endpoints interact to produce outcomes in tumor modeling (Skipper, 1990; Kerbel, 1999). The goal in clinical oncology is regulation to improve survival and quality of life, and
to prevent recurrent disease rather than to cure/kill cancer (Schipper et al., 1995). More stringent criteria than curability of experimental tumors in mice need to be assessed to determine therapeutic efficacy. Counts of tumor-bearing animals and measurement of tumor burden are the easiest and most frequently used outcomes of efficacy in preclinical studies. In reality, endpoints need to be matched to type of tumor (solid, leukemia or metastatic), context of the study, accessibility of the implantation site, type of implantation, and therapeutic class. Simplicistic criteria used in mouse models do not match criteria of partial and complete responses used in clinical oncology, and they contribute to the conflicting opinions about the relevance and predictability of mouse models. Therefore, use of other metrics, and evaluations of angiogenesis, immunomodulation, metastases, and detailed histopathology need to be incorporated into most study designs for preclinical testing. Despite the technical difficulty, labor-intensive nature, and expense commonly cited as limitations for detailed examinations, the utility of mouse models is improved by multiple and appropriate endpoints (Table 2).

**In Vivo Endpoints**

Tumor growth inhibition studies where treatment is prophylactically administered before or on the day of tumor induction are not realistic for preclinical evaluation of clinical responses. Typically, preclinical efficacy studies utilize tumor growth delay in which tumors are induced by injection or surgical implantation and allowed to establish for a number of days prior to initiation of treatment. Solid tumors in accessible sites are amenable to a variety of metrics that are not applicable to primary and metastatic tumors of internal organs and hematologic neoplasms. A count of tumor-bearing animals assumes that non-tumor-bearing animals represent tumor regression or cures. Spontaneous regressions, failure of tumors to become established or displaced tumor mass into adjacent body cavities may account for a small percentage of false cures. For this reason, tumor onset and progression should be monitored daily to ensure adequate and similar tumor masses prior to treatment and monitor onset of regression in each group. In models with log-phase tumor growth, animals can be randomized into treatment groups after a predictable period of development, usually 3–14 days. Conversely, less well-developed tumor models may require enrollment of individual animals into the study when the tumor burden reaches a minimum size. Such enrollment studies are difficult to evaluate. Allowing extra days for tumor growth may be as misleading as starting treatment on small tumor masses that have not established and show enhanced regression after onset of treatment. In vivo progression of tumor burden should be evaluated on a daily basis, excluding tumor-free animals as they appear. Burden is commonly measured with calipers and volume estimated from measurement of two (length and width) dimensions (Corbett et al., 2002; Teicher, 2002). Estimates of tumor weight (length[width²]/2) using the typically inaccurate measurements derived from tumors of varying shapes and boundaries is not recommended. Tumor growth delay measures the difference in days for the mean or median tumors in test and control animals to reach a specific volume, usually

| Endpoint | Comment | Formula

| In vivo  | Tumor onset | Day of palpable tumor mass of preselected size | Length × width
| Tumor progression | Daily plot of tumor burden | Length × (width × 2)
| Number of tumor-bearing animals | Tumor free assumes cure | Length × (width)² × 0.5
| Tumor burden—in vivo | Measured (mm²) | T − C in days
| Volume estimated (mm³) — 2-dimensional measurement | T − C in days)/3.32
| Tumor growth delay | Delay to reach specific volume | 1) (T − C in days)/3.32 × Td
| Tumor cell kill | 1) Log₁₀ total tumor cell kill | 2) (T − C) (duration of treatment in days)/3.32 × Td
| Survival – life span or long-term | T/C (%) | T/C
| Ex Vivo | Survival—number alive | Increase in mean or median lifespan or long-term survivors
| 1) Treatment termination | Length × width × depth
| 2) Posttreatment | Tumor weight/body weight
| Tumor burden—gross pathology | 1) Volume estimated (mm³) — 2- or 3-dimensional measurements | Length × width × depth
| 2) Absolute weight (mass in mg as water displacement or wet weight) | Tumor weight/body weight
| 3) Tumor weight relative to body weight | 4) Uceration
| 5) Invasion or tissue distribution and gross lesions (e.g., infarction) | 5) Invasion or tissue distribution and gross lesions (e.g., infarction)
| Hematology | 1) Complete blood count | a. Confirm histogenesis and differentiation
| 2) Differential blood count | b. Identify invasion and metastases
| 3) Bone marrow differential | c. Confirm expected therapeutic activity
| Histopathology | 1) Hematoxylin and eosin | d. Necrosis—characteristics and estimate percentage
| a. Confirm histogenesis and differentiation | e. Evaluate angiogenesis, hemorrhage, edema, and immune and stromal responses
| b. Identify invasion and metastases | 2) Morphometrics
| c. Confirm expected therapeutic activity | 3) Immunohistochemistry and in situ hybridization
| d. Necrosis—characteristics and estimate percentage | 4) Molecular pathology

aT = test; C = control; Td = doubling time; 3.32 = doublings to increase 1 log₁₀ unit.
bFormulas used for volume estimates vary.
0.5–1.0 cm³. This value has been suggested to mimic clinical endpoints and disease progression, but this measure is often incorrectly applied. A 50% reduction in tumor mass as a measure of cytoreduction used to gauge clinical responsiveness is not equivalent to 50% inhibition of tumor growth commonly used as a measure of preclinical efficacy (Corbett et al., 2002; Teicher, 2002). Tumor cell kill (net and total) for leukemic or solid tumors requires tumor titration using 10-fold dilutions to determine doubling time, and comparison to tumor growth delay and duration of treatment. Despite the value and more exact nature of these determinations, extra time and additional test animals required generally result in infrequent use of these endpoints (Corbett et al., 1999; Harrison, 2002; Teicher, 2002). Survival data should not be mistaken for mortality, an unacceptable endpoint (UKCCCR, 1998; Wallace, 2000). An increase in percentage mean or median survival may be measured for the lifespan or in long-term survivors. Clinically, posttreatment survival is a valuable endpoint that evaluates treatment efficacy during and after treatment. Survival at study termination is inadequate due to the short duration of most assays and lack of posttreatment data. For many tumors, a log-linear growth cannot be assumed, and more importantly, regrowth of tumors post-treatment may escalate due to increased doubling times (Gunduz et al., 1979; Teicher, 2002). Retention of a subset of animals for several weeks after treatment ends may demonstrate a rapid regrowth of tumor at the efficacious dose that quickly parallels morbidity of controls and lower dose groups. In spite of a delay in tumor growth over the treatment period, a rebound effect post-treatment would be an indication of lower overall efficacy.

**Ex Vivo Endpoints**

At study termination, volume measurements similar to those performed in vivo can be made at gross pathology, but mass (water displacement) or absolute wet weight ex vivo is more exacting. Volume measurements at termination may show significant disparity from final in vivo measurements. Irregular shapes, small or multilobular masses and necrosis, edema, and hemorrhage all contribute to variation in estimates of tumor burden measurements. Relative tumor to body weight ratios are often useful to gauge therapeutic efficacy in slow growing tumors and when cachexia is present. The relative ratio can reduce or expand differences between controls and treated animals and is a valuable and easily obtained supplement to absolute tumor weight. Visualization and quantification of metastases is particularly difficult to monitor as multifocal and occult tumor burdens are frequent. Flooding of airways with India ink to highlight lung masses for counting, “bread loafing” of the target organ for metastatic counts and organ weight differences compared to controls are inexact measurements prone to intra- and inter-study variability. Such models are better served by tagged tumors and advanced imaging technologies. Gross pathology examination should evaluate presence or absence of ulceration, extent of intentional or inadvertent invasion, tissue distribution beyond the primary implant site, and identify other gross lesions. Accidental implantation of tumors into body cavities can skew results by partial growth of tumor at the intended site that responds to treatment, along with a large internal tumor burden that is unresponsive to treatment. Hematology (complete blood count and differential) and bone marrow differentials (smears or fluorescent activated cell sorting analysis) can provide data about paraneoplastic syndromes and toxicity. However, hematology and histopathology endpoints are often considered elective. Histologic screening of tumors is valuable to confirm histogenesis and state of differentiation of tumors for quality control and to identify and separate local invasion from metastasis. Characteristics of the tumor such as necrosis, angiogenesis, hemorrhage and immune cell and stromal responses should be monitored. More importantly, histology should confirm the expected activity and target of the therapeutic class. Novel therapeutics currently under development are driving the need for additional ex vivo studies such as morphometry to evaluate angiogenesis and apoptosis, immunohistochemistry, in situ hybridization and molecular pathology to evaluate changes in macromolecules, immunophenotypes and nucleic acids.

**Pharmacology and Safety Endpoints in Tumor-Bearing Animals**

Toxicokinetic (TK), and absorption, distribution, metabolism, and excretion (ADME) studies are generally performed early in drug discovery (Lin et al., 2003) and are often incorporated into anti-tumor protocols to determine dose and scheduling. Conversely, safety testing is almost exclusively performed in nontumor-bearing animals. Considering the altered homeostasis of most tumor-bearing animals and humans, safety in a tumor-bearing host should be considered to fully evaluate the therapeutic response and potential. Adverse events can be monitored using a standard safety protocol in tumor and a non-tumor-bearing mouse model, but tolerability of treatment is more difficult to monitor. Limited resources and economic restraints would not support conduct of a full safety program in tumor-bearing animals, but examination of a full panel of tissues from at least one model may be useful to identify adverse effects, modulation of paraneoplastic syndromes, and tolerability issues. Although safety information is generally collected in immunocompetent animals, use of immunodeficient strains may be a better predictor of TK/ADME and safety as a mimic of immune dysfunction after chemotherapy, radiation therapy, surgery (Wichmann et al., 2003), infections and paraneoplastic syndromes (Tanaka et al., 1996). A modified safety study can be incorporated into an efficacy study by collecting a complete or selected list of tissues from tumor-bearing animals and by including treated but non-tumor-bearing animals as additional controls. Although safety information in tumor-bearing animals may be incomplete, this data supplements efficacy, TK and ADME data and can provide needed redirection in dosage and scheduling for subsequent studies.

**Humane Considerations**

Therapeutic intervention trials may use both short-term and long-term tumor growth models depending on growth rate and aggressiveness of the tumor. Mortality should not be used as an endpoint. Commonly used endpoints such as tumor burden should be limited according to absolute values and relative to body weight. Therefore, careful monitoring and humane endpoints need to be developed in conjunction with the IACUC. Ulceration, tissue and body cavity distension,
inanition, cachexia, anemia, increased intracranial pressure, self-mutilation, cannibalism, and metastases also contribute to premature death of study animals and may require a limited protocol or modified study design (UKCCCR, 1998; Wallace, 2000).

**Trends in Tumor Modeling**

*Bioinformatics and Mouse Models of Cancer*


**Technical Trends in Tumor Modeling**

Technical advances in tumor modeling have included use of skin fold and transparent windows (Dellian et al., 1996; Li et al., 2000; Jain et al., 2002), matrigel (Eccles, 2002), polymers (Righi et al., 2003), liposomes (Kunstfeld et al., 2003), chambers (Dvorak et al., 1987) and hollow fibers (Sadar et al., 2002) to contain tumor growth and monitor angiogenesis. Ex vivo perfusion (Kristjansen, 2002) and intravital microscopy are providing insights into drug metabolism, angiogenesis and in situ tumor responses (Jain et al., 2002).

Improved imaging of tumors has revolutionized noninvasive monitoring of distribution, growth, metastasis and morphometrics of tumor models in mice. Available systems include micro positron electromagnetic imaging (Ray et al., 2003; Yang et al., 2003), magnetic resonance imaging (Berr et al., 2003; Nelson et al., 2003), in situ visualization of primary and metastatic tumors and occult metastases (Menon and Teicher, 2002), using improved fluorochromes (Rosenberg et al., 2003) and quantum dots (Watson et al., 2003), green fluorescent protein (GFP) (Hoffman, 2002), and luciferase (Edinger et al., 1999; Burgos et al., 2003). Improved fluorochromes, GFP, and LacZ (Kruger et al., 1999; Culp et al., 2001) are also retained within tissues and allow microscopic evaluation by fluorescent microscopy and histochemical or immunohistochemical staining of tumor tissues. Laser capture microdissection is a proven technology for identification of genetic heterogeneity of tumors that continues to expand our knowledge of tumor biology (Culp et al., 2001; Hoon et al., 2002). Specialized microscopy using confocal (Paddock, 1999) and deconvolution (Maierhofer et al., 2003) microscopes are also positioned to provide optical sectioning for identification of 3-dimensional distribution of tumor elements and drug distribution within tumors (Mavivasager et al., 2002). While many of these improvements are expensive and impractical for incorporation into preclinical studies, selective use of advanced technologies in drug discovery, pharmacology and preclinical studies can provide minimally invasive and detailed information about host-tumor-therapeutic interactions in vivo at multiple time-points.

**Biological Trends: Genetically Engineered Mice**

Neoplastic transformation and progression requires a series of genetic alterations that disrupt the balance of cellular mechanisms involving cellular growth and deletion. Mice have played an important role in defining the genetic mechanisms of carcinogenesis. Thus, it is no surprise that genetically engineered mice (GEM) are beginning to take their rightful place as models that show accelerated tumor development and recapitulate the genetics and behavior of human cancer states and cancer resistance (Mickisch et al., 1991; Piot, 2001; Klatt and Serrano, 2003). Selected target genetic events allow creation of gene-driven gain of function transgenic, loss of function deletion (Reilly and Jacks, 2001; Meuwissen et al., 2001; Resor et al., 2001; Balmain, 2002; Jackson-Grousby, 2002; Herzig and Christofori, 2002; Tuveson and Jacks, 2002; van Dyke and Jacks, 2002), conditional function (Jonkers and Berns, 2002), clones (Rideout III et al., 2000) and phenotypically-driven N-ethyl-N-nitrosourea mutants (Justice et al., 1999; Balmain, 2002).

Although genetically engineered mice have been advocated for preclinical testing (Feitelson and Larkin, 2001; Horig et al., 2001), problems associated with these models include discordance in etiology and histogenesis between human and mouse tumors, multifocal tumors due to multi-tissue deletions or promoter promiscuity, failure to metastasize, variable penetrance of transgenes, long latency (Rosenberg and Bortner, 1999; Moore and Nagle, 2000), limited availability, lack of extensive historical pathology databases for founder strains (FVB, 129 strains, BALB/c), and costs and effort in creating and maintaining specialized GEM animal colonies. Regardless, the ability to use GEM to genetically, anatomically, pathophysiologically and histologically mimic tumors found in humans is becoming a reality. As more logistical problems are overcome, these genomic-derived models will likely become more consistently incorporated into preclinical testing. Compared to transplantation models, GEM will be particularly useful to create tumor states that were previously difficult to model including nervous system (Gutmann et al., 2003), pancreas (Hotz et al., 2000; Bardeesy et al., 2001), lung (Tuveson and Jacks, 1999; Liu and Johnston, 2002), breast (Hutchinson and Muller, 2000), ovary (Rahman et al., 1998; Rahman and Hultaniemi, 2001; Orsulic et al., 2002), prostate (Sharma and Schreiber-Agus, 1999; Navone et al., 1999), oral (Opitz et al., 2002), liver (Fausto, 1999; Feitelson and Larkin, 2001; Koike, 2002), hematologic malignancies (Bernardi et al., 2002; Herzig and Christofori, 2002), pediatric tumors (Beltinger and Debatin, 2001; Houghton et al., 2002), gene-environment interactions (Hursting, 1997), and metastases (McClatchey, 1999; Herzig and Christofori, 2002), and to test interventional feasibility for novel therapeutics against bone metastasis (Fausto, 1999), cell adhesion dysfunction (Herzig and Christofori, 2002), telomere dysfunction (Goytisolo and Blasco, 2002; Artandi, 2002; Granger et al., 2002), ribosomal RNA modification (Ruggiero et al., 2003), and DNA hypomethylation (Goodman and Watson, 2002; Gaudet et al., 2003). Despite these ongoing advances in model systems, disease states associated with tumors such as minimal residual disease (Teicher, 1997; Wetterwald et al., 2002), concurrent opportunistic infections, and treatments such as surgical debulking and combination
therapy will continue to resist modeling in mice for practical and humane concerns.

**Biological Trends: Modified Immunodeficient Mice**

Nude (athymic) and C.B-17 SCID (T and B cell deficient mice) fail to reject engraftment with a variety of human and mouse tumor cells (Pettaway et al., 1996; Lapidot et al., 1997; Bankert et al., 2002; Giovanella, 2002; Uckun and Sensel, 2002) by virtue of their T and B cell defects. The presence of innate immunity, particularly natural killer (NK) cell activity is probably an important factor in limiting tumorigenesis and metastases in these models. The nude mutation results in a mouse that is T cell deficient and has B cell maturational defects, but with intact innate immunity including tumoricidal macrophages and an increase in NK cells. This mutation has been introduced into many mouse strains. Introduction of the beige mutation blocks NK and myeloid-derived cell activity and additional B cell defects result by crossing with the X-linked immunodeficiency mouse. The Chédiak–Higashi syndrome (hypopigmentation, bleeding diathesis and recurrent bacterial infections) in mice or hybrids with the beige anomaly limits their usefulness in surgical, including orthotopic models (Clarke, 2002). The T and B cell deficient SCID mouse has been particularly amenable to further manipulation and modifications involving expanded immunodeficiencies and “humanization.” These mice are useful for mechanistic studies and for evaluating anti-tumor therapies on human xenographs (Bankert et al., 2001). However, NK cell activity and the tendency of some SCID mice to become “leaky” and develop active T and B cells, and early onset lymphoproliferative diseases can be problematic. Introduction of the nonobese diabetic (NOD) mutation and recombination activating gene (Rag) deficiency are useful modifiers for the SCID mutation (Clarke, 2002; Eccles, 2002). The NOD mouse is a complex model of immune defects that includes autoimmune mediated type-1 diabetes and sialitis, and Rag-1 and Rag-2 null state produces a severe combined immunodeficiency. Combination of these mutations with SCID defects results in mice that are no longer leaky and NK cell activity is reduced while retaining the ability to support xenographs (Bankert et al., 2001).

SCID or NOD-SCID mice made chimeric with implanted human lymphoid tissues, peripheral blood cells or bone marrow cells are the most commonly used “humanized” mouse models. These models have been generated to provide surrogate human microenvironments and to test immunotherapies (Bankert et al., 2001). Anti-tumor responses in humanized mouse models have not been extensively validated and their use supplements rather than supplants other mouse models for antitumor testing.

**Biological Trends: Orthotopic Models**

Less than ideal growth and behavior of heterotopic implants has lead to development and characterization of surgical and cellular orthotopic implants (Manzotti et al., 1993). Both normal and neoplastic cells (individual or clusters) and histologically intact tissues, of fetal and adult origin can be used in syngeneic and xenogeneic orthotopic mouse systems. Normal tissues can be implanted and mice injected with tumor cells to produce an improved metastatic model, and using tumor cells with a reporter gene allows enhanced imaging capabilities. Implantation of tumor tissue into anatomically correct tissue for the histogenesis of the tumor (i.e., kidney tumor in kidney) rather than heterotopic implantation promotes improved tumor growth and metastases (Naito et al., 1987a, 1987b; Menon and Teicher, 2002). Orthotopic models also allow correlation of experimental responses with the original tumor in the host (Steel, 1987).

Although orthotopic models have been utilized for many years, the need for surgical manipulation in many of the models and need for a fresh source of human tissues (normal or neoplastic) adds to the difficulty and expense of this type of implantation. Regardless, orthotopic models provide a useful model for establishing tumors previously difficult to model (Fu et al., 1992; Furukawa et al., 1993a, 1993b; An et al., 1999; Myers et al., 2002) and can be used alone or in combination with humanized mice (Klausner, 1999; Kunstfeld et al., 2003). Orthotopic models allow efficacy testing for tumor inhibition and metastases (Menon and Teicher, 2002; Zhang et al., 2002; Boyd et al., 2003) and appear to be better predictors of clinical success than heterotopic models (Kuo et al., 1993; Manzotti et al., 1993; Kilian et al., 1999; Hoffman, 1999; Bloomston et al., 2002).

**Other Animal Models for Preclinical Efficacy Studies**

Their size, fecundity, ease of handling, relatively economical production and care (excluding genetically altered mice), strain selection, and short gestation backed by massive databases of susceptibilities, genetics, immunology, physiology, pathology, and microbiology make mice an ideal candidate for tumor biology and preclinical efficacy studies. Rats and hamsters are also useful for biological and therapeutic modulation because of their propensity to develop a variety of spontaneous tumors, their amenability to tumor implantation, availability of cell lines (Schwartz and Gu, 2002; Thompson and Sporn, 2002), and recent development of genetically altered rats lacking suppressor genes linked to breast and ovarian cancer (Zan et al., 2003). Beyond rodents, a limited number of spontaneous neoplasms in companion and domestic animals (Vail and MacEwen, 1997; Knapp and Waters, 1997; Dewhirst et al., 2002) and other species including genetically altered fish (Vanchieri, 2001; Spitsbergen and Kent, 2003) also have some potential value in preclinical efficacy testing. Dogs, particularly older males provide a very useful model of prostatic disease (Waters and Bostwick, 1997; Waters et al., 1998; Strandberg, 2000) and dogs larger than purpose-bred Beagles are often the only relevant model to test medical devices or device/drug combinations for anti-tumor applications.

**Towards an Ideal Preclinical Efficacy Testing Program**

In order to transition into clinical trials, regulatory agencies must be provided with evidence that the new therapeutic has an improved safety and/or efficacy profile compared to current therapies. Therefore, selection of tumor models in mice is best approached by using specific criteria that match biologic (multistage, clonal, progression, histomorphology, and metastasis), genetic (multiple mutations, altered chromosomes and cell signaling, genetic expression profile and susceptibility), inductive etiology (chemicals, UV light, diet, hormones and viral), immunogenicity and therapeutic
FIGURE 1.—A guide for preclinical efficacy testing of anti-tumor therapeutics. In addition to toxicokinetics (TK) and absorption, distribution, metabolism and excretion (ADME) studies, development programs should progress through an increasingly complex series of tumors in one or more mouse models, genetically engineered mice (GEM) and suitable models in other species. Once a probable level of efficacy is established by use of stringent endpoints in multiple models, an investigational new drug (IND) application may be feasible. Safety studies in tumor-bearing mice should be considered to evaluate toxicity during altered host homeostasis. If insufficient efficacy is established, additional studies utilizing orthotopic models, GEM and combination therapy models may be necessary to provide sufficient evidence of probable clinical efficacy. Well-characterized orthotopic and GEM models may be biologically superior to testing in other models and may provide a shortcut (dotted line) to demonstrating efficacy.

potential (target and predictivity for clinical success) between mice and humans (Siemann, 1987; Hann and Balmain, 2001; Balmain, 2002). Although a generic program is not advocated, in vivo preclinical efficacy programs for anti-tumor therapeutics should follow a guide for testing similar to that outlined in Figure 1. After selecting the appropriate dose and scheduling, an efficacy program most economically progresses from simple to more complex tumor model systems. Use of one or more simple models alone to determine possible clinical efficacy is contrary to the complexity and issues associated with tumor models in mice and the variation in accepted therapeutic protocols in clinical oncology. Utilizing multiple surrogate models (cells and explants, different stains of mice and/or additional species, different tumor stages), transplants using variants of the target tumor, and tumor testing in genetically altered animals begins to provide sufficient efficacy data that can support an IND application. This efficacy data should be accompanied by results from PK and ADME studies that evaluate kinetics in one or more tumor model systems. The IND application can also be strengthened by conducting safety studies in tumor-bearing animals as a model to mimic altered homeostasis in tumor-bearing humans.

Ideally, the IND should only be filed after completion of additional preclinical studies that match the genetic heterogeneity of human tumors (multiple animal and tumor models), at least partially match tumor prone genetic profiles in a tumor-bearing host (genetically altered animals), match tissue tropism (orthotopic models), and match the most accepted therapeutic protocols (combination therapy models). In some cases, the most appropriate models will be GEM or orthotopic models and a limited number of studies in standard mouse models are justified. Incorporation of multiple levels of efficacy testing along with careful selection of endpoints will lead to a development program that most closely matches molecular targets, therapeutic protocols and outcomes encountered in the clinic. Testing through all tiers may not be applicable or warranted and completion of all steps is not a guarantee of correlation with clinical efficacy. However, a thorough preclinical efficacy program provides greater assurance of potential therapeutic efficacy and safety than tumor modeling restricted to one or two implanted tumors in a single strain of mouse. The expense of comprehensive modeling and additional time spent in preclinical testing is modest compared to failure in phase II and phase III clinical trials.

SUMMARY

Transitions from in vitro to preclinical and then to clinical testing for tumor modulation remain difficult with a low rate of clinical entry for most therapeutic classes. Increased understanding of mechanisms of neoplasia through macro-molecular biology, genomics and bioinformatics is helping to address treatment bottlenecks such as lack of specificity, low efficacy, toxicity and drug resistance, and helping to identify critical targets for clinical exploitation. In addition to cytotoxic, hormonal, adjunct, and medical device therapies, numerous novel strategies for enhanced and targeted drug deliver, anti-angiogenesis (inhibitors and enhanced
permeation), immunotherapy (vaccines, monoclonal antibodies, toxin conjugates, prodrug activators, cytokine antagonists), small molecules (inhibitors of growth, matrix and adhesion), apoptosis (enhancers, inducers, proteasome inhibitors, reverse DNA methylation), anti-sense and gene therapy (tumor suppressor genes) and cell cycle alterations (inhibitors) are being developed for the anti-tumor market. However, our ability to model and accurately predict clinical efficacy is limited. Despite historical significance and ongoing utility, tumor models in mice used for preclinical therapeutic intervention often error towards false positive results and curing cancer in mice. The inadequacy of classic transplantation models for anti-tumor therapy is helping to drive development and use of new models based on genetic and technical modifications. However, underlying limitations of tumor models reinforce the need for careful attention to design (applying correct models to the question), conduct (using multiple models) and interpretation (recognizing limitations and applying stringent criteria to outcomes) of efficacy studies for tumor modulation. Animal models can provide quick answers but application of these results to predicting clinical outcomes is often undertaken prematurely. New strategies and techniques, and continued improvements in stringency and consistency of criteria used for evaluating outcomes will be necessary to ensure that tumor models in mice remain a useful tool for development of anticancer agents and devices.

REFERENCES


Tumor Models in Cancer Research


