Abstract. The rapid emergence of new anticancer agents is a tremendous challenge for basic, pre-clinical and clinical research to evaluate and eventually integrate these new agents into clinical routine. Standardized, well-established in vitro and in vivo methods are available for the experimental evaluation of new anticancer agents. A step-wise procedure from in vitro to in vivo experiments using non-functional, functional non-clonogenic and, if applicable, clonogenic assays allows reduction of the number of promising agents for further clinical testing.

New anticancer agents are either designed for tumor-specific targets based on a biological rationale or are generated by large-scale drug screening programs. Because of their higher specificity, these new therapeutics promise higher efficacy combined with a lower toxicity than classical cytotoxic agents such as chemotherapy and ionizing irradiation. The fascinating, rapid emergence of thousands of new drugs offers great hope for patients. On the other hand, it is a tremendous challenge for basic, pre-clinical and clinical research to evaluate and eventually integrate these new agents into the clinical routine.

The question of whether a new drug improves cancer therapy in patients can ultimately only be answered in a clinical trial. However, because of ethical, medical and economic limitations and constraints on the number of patients eligible for clinical trials, most of the research has to be done in experimental systems. Over many decades, researchers in experimental tumor therapy have developed well-proven, reliable in vitro and in vivo methods to evaluate treatment response. The use of these standardized experimental methods is time-consuming and costly, and there is, consequently, a gap between the quantity of new agents and the resources available for their evaluation. Even in an ideal world where cancer research is appropriately funded, well organized in multi-institutional networks and focused on a few, particularly promising drugs, it will take a long time from drug discovery to approval for clinical use. It is clear that the continuous improvement of standardized experimental methods to expedite evaluation of new drugs is an important part of cancer research, while neglecting these methods is potentially harmful and certainly a waste of resources.

Translational research chain in evaluation of anticancer agents

Evaluation of anticancer agents depends critically on the interaction of basic, pre-clinical and clinical research in a structured network (Figure 1). The so-called translational research chain is usually envisaged as a more or less step-wise, hierarchical system of in vitro studies and animal models converging towards clinical trials and eventually to standard of patient care (1). However, in practice the translational process is not uni-directional, but rather may reverse direction at each step.

Functional versus non-functional assays

Experimental evaluation of new anticancer agents is realized by means of in vitro and in vivo methods to describe whether or not a new drug is effective against cancer cells. The so-called functional assays basically measure survival of
tumor cells with and without therapy, e.g. as a total number of cells, a number of colonies, tumor volume or tumor cure rate. Non-functional assays are often also referred to as mechanistic investigations, e.g. assessment of drug effects on apoptotic pathways or intracellular signaling, and are important to improve our understanding of the underlying mechanisms of action. Both functional and non-functional assays are essential for the evaluation of anticancer agents.

**Clonogenic versus non-clonogenic tumor cells**

Tumor cells with the capacity to produce an expanding family of descendents are clonogenic (2, 3). Experimental data show that only a small percentage of cells in the tumor are clonogenic cells. Most tumor cells are non-clonogenic and die without any therapy after some cell divisions. As an example, 90-99% of tumor cells in FaDu tumors, a human squamous cell carcinoma growing in nude mice, are non-clonogenic (4). To cure a tumor, which is the ultimate goal of cancer treatment, it is necessary to inactivate all clonogenic cells either by cell kill or by inducing a permanent state of dormancy, i.e. the loss of clonogenic capacity. The evaluation of whether a drug has a curative potential, i.e. effectively inactivates clonogenic tumor cells, requires the use of experimental endpoints that represent the response of clonogenic cells. Clonogenic endpoints are of particular importance when new anticancer agents are integrated into curative therapeutic settings, e.g. in combination with radiotherapy or chemotherapy.

**Intertumoral heterogeneity in response to new anticancer agents**

Not all tumor cell lines show the same magnitude of response to anticancer agents. For most anticancer agents the underlying reasons for intertumoral heterogeneity are poorly understood. Experimental data suggest that expression levels of the molecular target and specific genetic alterations are important determinants for response. For example, the response to inhibitors of epidermal growth factor receptor (EGFR) shows a considerable heterogeneity between different cell lines in vitro and in vivo (5-8). Expression patterns of EGFR and HER2/neu are distinct between different tumor cell lines and seem to correlate with response to the corresponding inhibitor (7-10). It has been suggested that the specific mutational pattern of down-stream pathways determines whether cell proliferation or cell survival is preferentially affected by EGFR inhibitors (11). In line with experimental studies, clinical data show that specific mutations correlate with the individual response of tumors to EGFR inhibition (12, 13). Experience with EGFR inhibitors clearly shows the importance of evaluating new anticancer drugs in a range of different tumor cell lines. The use of “out-liner” or “best-responding” cell lines may help in studying the mechanism of action of a particular drug, but may also lead to an overestimation of its therapeutic potential. Systematic exploration of heterogeneity by molecular profiling will help to tailor new approaches and to identify patients who might benefit from new anticancer agents.
**In vitro methods**

For most anticancer agents the initial step of evaluation is cell culture. Compared to animal tumor models, *in vitro* methods are less expensive and less time-consuming, thereby allowing evaluation of large quantities of new anticancer agents. Molecular methods to prove and quantify the potential of several drugs to affect the molecular target, *e.g.* to decrease the activity of a specific kinase, facilitate the selection of promising candidate drugs. Sophisticated *in vitro* experiments provide data on mechanisms of action, which, when combined with detailed characterization of tumor cell lines, help to identify tumor entities which may respond to the drug. Based on these data, further selection of promising drugs for *in vivo* testing requires data obtained from functional assays.

In general, for functional assays cells are exposed to different drug concentrations and the response is monitored. Cell counting or dye-based assays such as MTT are quick and robust methods to estimate the total number of surviving cells. Assessment of the fraction of surviving clonogenic cells can be done either by the colony forming assay or by the dilution assay (14). Clonogenic assays are laborious and require experience. As a consequence, non-clonogenic and non-functional tests are preferred. However, data obtained from non-clonogenic and clonogenic assays are not necessarily consistent (15-17). Thus, for evaluation of the effects of a new drug on clonogenic tumor cells, non-clonogenic assays cannot replace clonogenic assays.

From the response data *in vitro* characteristic parameters can be calculated. For example, the IC$_{50}$ value describes the drug concentration necessary to reduce the number/fraction of cells to 50% compared with the controls. The IC$_{50}$ value allows comparison with results obtained with other drugs and other cell lines. Moreover, from IC$_{50}$ and pharmacokinetic data, it can be estimated whether effective drug concentrations are achievable *in vivo*. Thus, quantitative *in vitro* evaluation of anticancer drugs is fundamental for further testing in animal models. Recent developments, such as co-culture models (18) and the use of genetically manipulated cell lines, have improved our methods of studying the mechanisms underlying the drug effect on cancer cells (19). Despite their importance for drug testing, *in vitro* methods are beset by pitfalls and inherent limitations.

The results from clonogenic assays critically depend on experimental design, especially on drug exposure times. Misleading results may be obtained from the colony-forming assay (CFA) if cells are continuously exposed to a drug. For example, drugs like EGFR inhibitors may inhibit proliferation without pronounced cell kill, *i.e.* clonogens proliferate more slowly but are not inactivated. The experimental endpoint of the CFA is the surviving fraction of clonogenic cells. For this, after an incubation period of several days, the medium is removed, the cells are stained and the number of colonies is counted under the microscope. A certain number, usually 50 or more, cells descending presumably from a single surviving clonogenic cell, defines a colony. For the example discussed here it is conceivable that the number of colonies is smaller after drug treatment than in the controls. This result would indicate that the drug has a curative potential because it reduces the number of colonies. However, if also smaller colonies are counted or colony counting is performed at a later time point, it may become clear that the clonogenic cells are not inactivated and the drug has solely an anti-proliferative but no curative potential. Evaluation of more than one time point for colony counting or plating after drug exposure may help to reduce the problem.

Results from clonogenic assays also depend on culture conditions. Different cell lines may require different conditions, such as media composition. Also, the drug effects on clonogenic survival are affected by culture conditions (19). Enhancing the drug effects by additives to the medium may be useful for studying mechanisms, but artificial culture conditions may well not represent the situation *in vivo*. Optimizing culture conditions for clonogenic assays with different cell lines and drugs is very laborious. In practice this restricts the large-scale use of clonogenic assays for drug screening on different cell lines. However, if the drug is aimed for a potential curative treatment as monotherapy or in combination with other modalities, the use of clonogenic assays *in vitro* before proceeding to *in vivo* is indispensable.

It is obvious that cells in culture represent an artificial and simplified system. Unlike the situation *in vitro*, a tumor is a 3-dimensional complex consisting of interacting malignant and non-malignant cells. Vascularisation, perfusion and, thereby, drug access to the tumor cells are not evenly distributed and this fact consists an important source of heterogeneity in tumor response to drugs that does not exist *in vitro*. Therefore, prediction of drug effects in cancer patients based solely on *in vitro* data is not reliable and further evaluation in animal tumor systems is essential.

Given that it is practically impossible to test large quantities of new anticancer agents *in vivo*, the most important function of *in vitro* experiments is to select promising candidates for further testing and to gain insights into cellular mechanisms of action. To reduce the number of potential candidates for *in vivo* testing, it seems reasonable to proceed step-wise from non-functional to functional *in vitro* tests and, if applicable, from non-clonogenic to clonogenic assays. Although this procedure might exclude drugs that have no activity at a certain level, but which would have an anticancer effect at the next level, there is no obvious alternative to this strategy.
In vivo methods

An enormous variety of different tumor systems for in vivo evaluation of new anticancer agents is available. Mostly murine host systems are used for experimental tumor therapy because of the availability of in-bred lines at relatively low costs, the ease of obtaining tumor models and established, widely accepted experimental endpoints (20). Spontaneous or transplanted murine tumors can be studied in immunocompetent mice whereas investigation of human tumors requires an immunodeficient host, e.g. nude mice, to avoid tissue rejection. Spontaneous tumor models offer some advantages over transplanted tumor cell lines, e.g. genetic diversity, growth in the original environment, angiogenesis more likely to resemble the situation in patients, but are rather difficult to obtain and maintain (21). Genetically engineered mice may help to improve this situation (22). Experimental data show that tumor characteristics such as growth rate and potential to metastasize depend on implantation site. Tumors injected orthotopically, i.e. into the organ of origin, apparently behave more similarly to the clinical situation (23). Also, the response to anticancer drugs may depend on the implantation site (24). For practical reasons, scientists mainly use ectopically-implanted, subcutaneously-growing tumor models. Most pre-clinical data on new anticancer drugs were obtained using transplanted tumors in mice, frequently as xenografts of human origin. Considerable scepticism about the value of fast growing, ectopic tumors arose when some new drugs in clinical trails were not as effective as in the pre-clinical setting (25, 26). However, detailed comparison of pre-clinical results and clinical data reveals that ectopically-implanted tumor models can be remarkably predictive when experiments are performed under clinically relevant conditions (26). Thus, these models can provide proof of principle, but the magnitude of effect does not necessarily correspond to the clinical situation.

Animal tumor systems have to meet several requirements to be suitable for experimental tumor therapy (20, 27). It is very important that the tumor precisely reflects treatment response, and that the natural history of the host allows the study of the experimental endpoint, e.g. a sufficient life-span for follow-up to assess local tumor control. Stable biological characteristics of the tumor system such as expression of the molecular target, growth rate, differentiation and immune response are also required to assure the high quality of experiments. To avoid undetected changes of characteristics of tumor systems, it is indispensable that each researcher follows strict quality assurance protocols (28). The results obtained from quality assurance measures should always be reported. In our laboratory, we determine, for each experiment, the tumor identity by microsatellite assay, the human origin of the xenograft by LDH isoenzyme pattern, the growth rate of control tumors, histology and DNA index by flowcytometry.

Tumor growth delay assay. This functional assay is robust, standardized, widely accepted and used in most experiments to study anticancer agents in vivo (29). A delay in tumor growth would provide a benefit to cancer patients and is, therefore, an experimental endpoint of clinical relevance. A
large body of experimental data clearly shows that the growth delay assay is a valuable tool to evaluate new anticancer agents in animals. The tumor growth delay assay provides evidence for further drug evaluation in clinical trials. Moreover, results from pre-clinical investigations in animal tumor models may help to design clinical trials, while detailed in vivo experiments may help in understanding the results from clinical trials. The results from pre-clinical and clinical studies on inhibitors of EGFR and angiogenesis demonstrate that animal models can prove the principle of a new therapeutic approach.

Knowledge and experience is required to correctly perform and interpret the growth delay assay. Usually tumors are allocated to two experimental groups. Animals of the first group are treated with the anticancer agent. Animals receiving the so-called vehicle, e.g. the compounds and solutions that were used to prepare the drug solution, are controls. There is no consensus about the minimal group size to perform the tumor growth delay assay. Of course, the group size to detect a difference in tumor growth times depends on the magnitude of effect and intertumoral heterogeneity. Unfortunately, both factors usually are unknown when the experiment is designed. To account for intertumoral heterogeneity, it is useful to randomize the animals over the experimental matrix and to treat both experimental groups in parallel. As many tumor characteristics such as growth rate, cell loss, hypoxia, angiogenesis and response to anticancer agents may change with increasing tumor volume, it is necessary that the tumors in both experimental groups be enrolled into the treatment protocol at a similar tumor volume. Apparently most anticancer agents are more effective in smaller than in larger tumors. This is an important caveat because in clinical trials often patients with advanced stages and tumor masses are treated.

To determine tumor growth delay, the tumor volume is repeatedly measured, and for each individual tumor the time to reach a multiple of the starting volume, e.g. two, five or ten times the starting volume, is recorded. The calculated growth delay (tumor growth time of treated tumors minus tumor growth time of control tumors) is a direct measure of the drug effect on tumor growth. To generalize the data for comparison with other tumor models and drugs, the so-called specific tumor growth delay (ratio of growth delay to growth time of control tumors) is calculated. It is important to note that the endpoint of the tumor growth delay assay is a time to reach a volume but not a volume at a given time point. For many drugs the tumor growth delay increases with increasing endpoint sizes, because, in experiments with multiple drug administrations, the tumor growth delay increases with time because of the accumulation of drug effect. There is no consensus about the optimal endpoint size to report data from growth delay assays. If tumor cell kill is the major mechanism of action of an anticancer agent, dead and doomed cells and their clearance will contribute more and more to the tumor volume. Especially in slow shrinking tumors, this may mask the rapid regrowth of surviving tumor cells. Therefore, it appears that the smaller the endpoint size the more closely this will reflect the actual anticancer effect of the drug (29, 30). Multiple administrations of antiproliferative agents probably result in an increasing tumor growth delay with time. In fast growing tumors, the effect on tumor growth rate is detectable only after some drug administrations and, thereby, at later time points. Thus, for antiproliferative agents, larger endpoint sizes seem preferable. As the mechanisms of action of new drugs are usually unknown before the experiment, it is reasonable to analyze and report tumor growth delay with multiple endpoints.

Anticancer drugs may prolong tumor growth by several mechanisms. Agents may affect tumor cells directly or indirectly, e.g. via targeting stromal cells by inhibiting angiogenesis. Both directly and indirectly acting anticancer agents can reduce the tumor growth rate by inhibition of tumor cell production, increased tumor cell death, or improved clearance of dead and doomed cells. Determination of the mechanism underlying the anticancer effect of an anticancer agent by a simple tumor growth delay assay is impossible and requires more detailed in vitro and in vivo experiments. Whether a new drug affects proliferation or survival is of particular significance for designing more complex in vivo experiments and clinical trials.

Tumor control assay. In contrast to tumor growth delay, the results from the tumor control assay solely depend on the therapeutic effect on clonogenic cells. Permanent tumor control is the most relevant experimental endpoint for testing of potentially curative settings (31). In practical terms, after therapy tumors are followed-up and regrowth of the recurrent tumor is recorded. This requires sufficient follow-up times to detect virtually all recurrences. An alternative to this time-consuming procedure is the tumor-excision assay. For this, tumors are excised after treatment, a single cell suspension is prepared and cells are seeded into flasks or multi-well plates. After incubation, the fraction of surviving clonogens can be determined and compared with control tumors without treatment. Although this assay has the limitation that the survival of clonogenic cells is not determined in their original environment, the tumor-excision assay is less expensive than the tumor control assay because no follow-up is necessary and the number of animals required is smaller.

Many of the new anticancer drugs reduce tumor growth but do not eradicate the tumor. Combination of new anticancer agents with potentially curative therapy modalities, such as radiotherapy, can improve the results compared with radiotherapy alone. For example, inhibitors of the EGFR or VEGF-dependent angiogenesis are not curative as a monotherapy. However, the combination of these inhibitors
with irradiation in animal models consistently resulted in longer tumor growth delay than either treatment alone (6, 32, 33). Administration of the VEGFR2 mAB DC101 to tumor-bearing animals exposed concomitantly to fractionated irradiation improved the results of the tumor control assay (34). However, results from tumor growth delay and tumor control assays are not necessarily consistent. BIBX1382BS is a potent inhibitor of the receptor tyrosine kinase of EGFR, resulting in clear-cut effects on tumor cell proliferation \textit{in vitro} and \textit{in vivo} using the human squamous cell carcinoma FaDu, which shows membranous expression of the molecular target, \textit{i.e.} the EGFR (16). In combination with fractionated irradiation, tumors treated with BIBX1382BS showed a longer tumor growth delay than irradiated tumors or tumors treated only with BIBX1382BS (Figure 2). This clearly shows that the drug is also effective on the growth of irradiated tumors. However, BIBX1382BS did not improve the tumor control probability in the same tumor model (16) (Figure 2). The underlying reason for the discordance of the growth delay assay and the tumor control assay is unclear. From this example, it is quite obvious that an extrapolation of results from non-clonogenic assays to predict response of clonogenic cells can be misleading and may cause incorrect conclusions with far-reaching consequences for clinical trials. In our opinion, tumor control is the most relevant endpoint for pre-clinical testing of anticancer agents. Alternatively, large growth delay studies using different dose levels may yield results similar to those obtained from tumor control assays.

Monoclonal antibodies against EGFR have been shown to improve tumor control after radiotherapy in patients with head and neck cancer (35). Interestingly, xenografted FaDu tumors also showed a higher local tumor control rate after anti-EGFR antibody therapy with C225 and irradiation (36). Comparison of pre-clinical and clinical data of EGFR inhibition and radiotherapy corroborates the importance of detailed \textit{in vivo} studies with suitable, well-characterized tumor models in a clinically relevant setting. Neglect of clonogenic endpoints might result in misleading strategies for further clinical testing. Although failure of new approaches in the clinic cannot be prevented by \textit{in vivo} animal studies, consideration of data from carefully performed \textit{in vivo} studies on efficiency, curative potential and optimal regimen are valuable for the design of clinical trials and the investigation of mechanisms of action.

**Conclusion**

Standardized, well-established \textit{in vitro} and \textit{in vivo} methods are available for experimental evaluation of new anticancer agents. A step-wise procedure from \textit{in vitro} to \textit{in vivo} seems reasonable to reduce the large quantity of potential drugs to a few promising agents for further clinical testing. The clinical application for which the drug is aimed, \textit{e.g.} palliative, curative, tumor entity, or combination with other modalities, needs to be considered in the experimental evaluation. For evaluation of new anticancer agents, we advocate \textit{in vitro} and \textit{in vivo} experiments with at least two or three different tumor cell lines, applying functional non-clonogenic and, if applicable, clonogenic assays.

**Acknowledgements**

Supported by the Deutsche Forschungsgemeinschaft (Grant Ba 1433-4 to D.Z. and M.B.) and research grants from Schering AG (to D.Z. and M.B.) and Boehringer Ingelheim Austria (to D.Z. and M.B.), and the Buchanan-Seeger Research Professorship (HDT).

**References**


