DRUG DEVELOPMENT RESEARCH 34:91–109 (1995)

Research Overview

Some Practical Considerations and Applications of the National Cancer Institute In Vitro Anticancer Drug Discovery Screen

Michael R. Boyd and Kenneth D. Paull

Laboratory of Drug Discovery Research and Development, Developmental Therapeutics Program, Division of Cancer Treatment, National Cancer Institute-Frederick Cancer Research and Development Center, Frederick (M.R.B.), and Information Technology Branch, Developmental Therapeutics Program, Division of Cancer Treatment, National Cancer Institute, Bethesda (K.D.P.), Maryland

Strategy, Management and Health Policy				
Venture Capital Enabling Technology	Preclinical Research	Preclinical Development Toxicology, Formulation Drug Delivery, Pharmacokinetics	Clinical Development Phases I-III Regulatory, Quality, Manufacturing	Postmarketing Phase IV

ABSTRACT During 1985–1990 the U.S. National Cancer Institute (NCI) phased out its murine leukemia P388 anticancer drug screening program and developed as the replacement a new in vitro primary screen based upon a diverse panel of human tumor cell lines. For each substance tested, the screen generates a remarkably reproducible and characteristic profile of differential in vitro cellular sensitivity, or lack thereof, across the 60 different cell lines comprising the panel. Several investigational approaches to display, analysis, and interpretation of such profiles and databases, derived from the testing of tens of thousands of substances during the past 4–5 years since the NCI screen became fully operational, have been explored. A variety of useful, practical applications of the in vitro screen have become apparent. As these applications continue to evolve, they are proving to be complementary to diverse other anticancer screening and drug discovery strategies being developed or pursued elsewhere. Reviewed herein are some practical considerations and selected specific examples, particularly illustrating research applications of the NCI screen that may be more broadly applicable to the search for new anticancer drug development leads with novel profiles of antitumor activity and/or mechanisms of action.

Key Words: antineoplastics, cancer, drug development

INTRODUCTION

In simplest terms, the NCI in vitro primary screen consists of a panel of 60 different human tumor cell lines against which compounds are tested over a defined range of concentrations to determine the relative degree of growth inhibition or cytotoxicity against each cell line. The design and operation of the screen is such that for each compound tested, both the absolute and relative sensitivities of individual cell lines comprising the screen are sufficiently reproducible that a characteristic profile or "fingerprint" of cellular response is generated. Depending upon the extent of differential cellular response, the profile may contain much information which is useful for further research. The least interesting or useful (and expectedly most common) response to a random selection of chemical structures is none at all; that is, none of the cell lines show any evidence of growth inhibition or

Address reprint requests to Michael R. Boyd, Laboratory of Drug Discovery Research and Development, National Cancer Institute-Frederick Cancer Research and Development Center, Bldg. 1052, Rm. 121, Frederick, MD 21702-1201.

cytotoxicity. A similarly featureless profile may be obtained when one or more concentrations of the tested compound produce(s) growth inhibition and/or cytotoxicity of essentially the same magnitude across the entire panel of cell lines. Certainly, the NCI screen is capable of identifying highly potent, indiscriminant direct cell poisons; however, that is not a unique or particularly useful attribute of the screen.

In contrast, the cell lines comprising the NCI panel may show differential sensitivity to a given test substance. The degree of differential response between the most and least sensitive lines typically may be relatively small (e.g., 2- to 10-fold), or occasionally as great as 3–4 orders of magnitude. Furthermore, the cell lines may be widely heterogeneous in response to a given compound, or they may be comparatively homogeneous, with only a relatively few lines showing much greater or lesser sensitivity than average. Regardless of the magnitude of the differential or the degree of heterogeneity of response of the cell line panel, it is the reproducibility of the response fingerprint that is key to the useful information contained therein. This valuable information can be exploited productively in its own right, as well as in complement to other drug discovery research models and strategies.

Routine operation of the NCI in vitro screen began in 1990, after 5 years of extensive development and pilot evaluations during 1985–89. Reviews of the concept, rationale, and technical aspects of development of the screen are available elsewhere [e.g., see Boyd, 1986, 1989, 1993; Boyd et al., 1992]. From 1990 to the present, more than 30,000 compounds, submitted by cancer researchers worldwide, have been tested in the NCI screen. Screening databases derived therefrom have provided NCI staff and collaborators a unique opportunity to explore a considerable variety of data analysis strategies and methods. Reviews and other publications describing such studies are available [e.g., see Paull et al., 1989, 1995; Boyd et al., 1992; Hodes et al., 1992; Weinstein et al., 1992, 1994; van Osdol et al., 1994].

In many if not most of the important respects thus far examined, the results and conclusions from diverse analytical approaches have been convergent. Increasingly sophisticated mathematical and computational techniques are being developed and applied further, and undoubtedly these will add important new dimensions to the valuable information that can be derived from the in vitro screening panel. This seems particularly certain when the data from parallel ongoing efforts to further characterize the unique biology of the individual cell lines can be further integrated into the analyses.

The purpose of this brief review is to offer some practical considerations and to describe and illustrate some relatively simple and straightforward research applications that may be of immediate and considerable utility to many current and future users of the NCI service screen. In so doing, some selected examples are drawn from the authors' particular research experiences using the screen. The review is by no means intended to be comprehensive; the scope is limited to some generally useful applications that can be pursued by "nonexperts" using relatively simple analytical techniques with data generated and supplied routinely for pure compounds submitted for testing in the NCI screen.

THE SCREEN

Detailed descriptions of the screening assay in use as of 1990 are available elsewhere [Boyd, 1989; Monks et al., 1991; Skehan et al., 1990]. Some changes in the screen subsequently have been made, particularly in late 1992. These are noted briefly below. Investigators evaluating recent data in comparison with older data from the NCI in vitro screen may wish to take these differences into account.

Cell Line Panel

The identities, sources, derivation, morphological and immunocytochemical characteristics, and methods of maintenance of the cell lines comprising the NCI 60 cell line panel as of 1990 have been described in detail elsewhere [Boyd, 1989; Monks et al., 1991; Stinson et al., 1992]. On December 1, 1992, ten of the original cell lines were deleted from the panel to make way for ten breast cancer and prostate cell lines. The lines removed from the panel comprised: [lung] HOP-18, LXF-529L, DMS114, DMS273; [brain] XF-498, SNB-78; [colon] KM20-L2, DLD-1; [renal] RXF-631L; [melanoma] M19MEL. The lines added (and references to the original sources and/or corresponding descriptive publications) are as follows: [breast] MCF-7 [Soule et al., 1973], MCF-7ADR [Cohen et al., 1986], HS578T [Hackett et al., 1977], MDA-MB-231 [Cailleau et al., 1974; Siciliano et al., 1979], MDA-MB-435 [Cailleau et al., 1978; Brinkley et al., 1980], MDA-N (Steeg, NCI, unpublished data), BT-549 (American Tissue Culture Collection, Rockville, MD, unpublished data), T47-D [Keydar et al., 1979]; [prostate] DU-145 [Stone et al., 1978; Mickey et al., 1977], PC-3 [Ohnuki et al., 1980; Kaighn et al., 1979, 1981].

Screening Assay

In routine screening, each agent is tested over a broad concentration range against every cell line in

the panel. All lines are inoculated onto a series of standard 96-well microtitre plates on day zero, followed by a 24 h incubation in the absence of the test compound. The inoculation densities employed depend upon the particular cell line and its growth characteristics. Inoculation densities used currently in the screen for many of the cell lines are the same as originally published [Monks et al., 1991]. Exceptions, introduced as of December 1, 1992, are as follows (current densities used [cells/well] are indicated in parentheses): HOP-62 (10,000), UO-31 (15,000), 786-0 (5,000), LOX IMVI (7,500), SR (20,000). Inoculation densities used for the breast and prostate lines beginning with their addition on December 1, 1992, are as follows: MCF-7 (5,000), MCF-7ADR (15,000), HS578T (20,000), MDA-MB-231 (20,000), MDA-MB-435 (15,000), MDA-N (15,000), BT-549 (20,000), T47-D (20,000), DU-145 (15,000), PC-3 (7,500). Further exceptions introduced as of July 25, 1994, are: NCI-H226 (15,000), RXF-393 (15,000), ACHN (10,000), PC-3 (7,500). Test compounds are routinely evaluated at five 10-fold dilutions starting from a high of 10^{-4}M, unless otherwise requested. Following a 48-h incubation with the test compound, the cells are assayed by the sulforhodamine B procedure [Skehan et al., 1990; Monks et al., 1991; Rubinstein et al., 1990]. Optical densities are measured on automated plate readers, followed by computerized data acquisition, processing, storage and availability for display, and analysis.

DATA DISPLAY AND ANALYSIS

A detailed description of the contents and format of the data report package routinely provided to submitters of compounds for NCI screening has been published [Boyd et al., 1992]. The "dose-response matrix" part of the package is no longer provided or routinely used. The "dose-response curves" and the "mean-graphs" components of the report are the main interest of most investigators. Therefore, following are some brief descriptions and comments concerning the dose-response curves, calculated response parameters, and mean-graphs which are most germane to the examples to be presented. Also offered are some comments and suggestions as to standards for investigator reporting of NCI screening data in the scientific literature. The consistency, detail, format, and placement of such information has increasingly been a concern of some journal editors [e.g., see Editors, 1994].

Dose-Response Curves

Each successful test of a compound in the full screen generates 60 dose-response curves, which are printed in the NCI screening data report as a series of composites comprising the tumor-type subpanels, plus a composite comprising the entire panel. Data for any individual cell line(s) failing quality control criteria, or otherwise deficient for any cell line(s) not tested successfully, are eliminated from further analysis and are deleted from the screening report. Figure 1 shows contrasting patterns in the dose-response curves obtained from two different compounds. The figure was prepared directly from the corresponding NCI supplier reports by deleting extraneous or otherwise distracting information, and adding minimal scaling and reference information for clarity and substantial photoreduction. The cell line subpanels are identified in the figure legend (Fig. 1).

The "percentage growth" (PG) term, and meaning of the +50,0 and −50 response reference lines, and the calculated response parameters, GI_{50}, TGI, and LC_{50} have been defined elsewhere [see Boyd et al., 1992; Monks et al., 1991]. Although the response parameters are already calculated by computer and provided to the investigator in the data report package, it is important to appreciate how these values are determined, and likewise how this may affect data interpretation.

The 50% growth inhibition parameter (GI_{50}) is the concentration of test drug where $100 \times (T-T_0)/(C-T_0) = 50 = PG$. The optical density of the test well after the 48-h drug exposure is T; the optical density at time zero is T_0; and, the control optical density is C. The PG is a T/C-like parameter that can have values from +100 to −100. Whereas the GI_{50} may be viewed as a growth-inhibitory level of effect, the TGI signifies a "total growth inhibition" or cytostatic level of effect. The TGI is the drug concentration where $100 \times (T-T_0)/(C-T) = 0 = PG$. The LC_{50} is the lethal concentration, "net cell killing" or cytotoxicity parameter. It is the concentration where $100 \times (T-T_0)/T_0 = -50 = PG$. The control optical density is not used in the calculation of LC_{50}.

The GI_{50}, TGI, and LC_{50} values are calculated by interpolation using the tested concentrations that give PG values above and below the respective reference values (e.g., 50 for GI_{50}). Therefore, a "real" value for any of the three response parameters is obtained only if at least one of the tested drug concentrations falls above, and likewise at least one falls below, the respective appropriate PG reference value (i.e., the dose-response curve for that particular cell line must cross the respective PG reference line). If, however, for a given cell line all of the tested concentrations produce PGs exceeding the respective reference level of effect (PG value of +50,0 or −50 as appropriate), then the lowest tested concentration

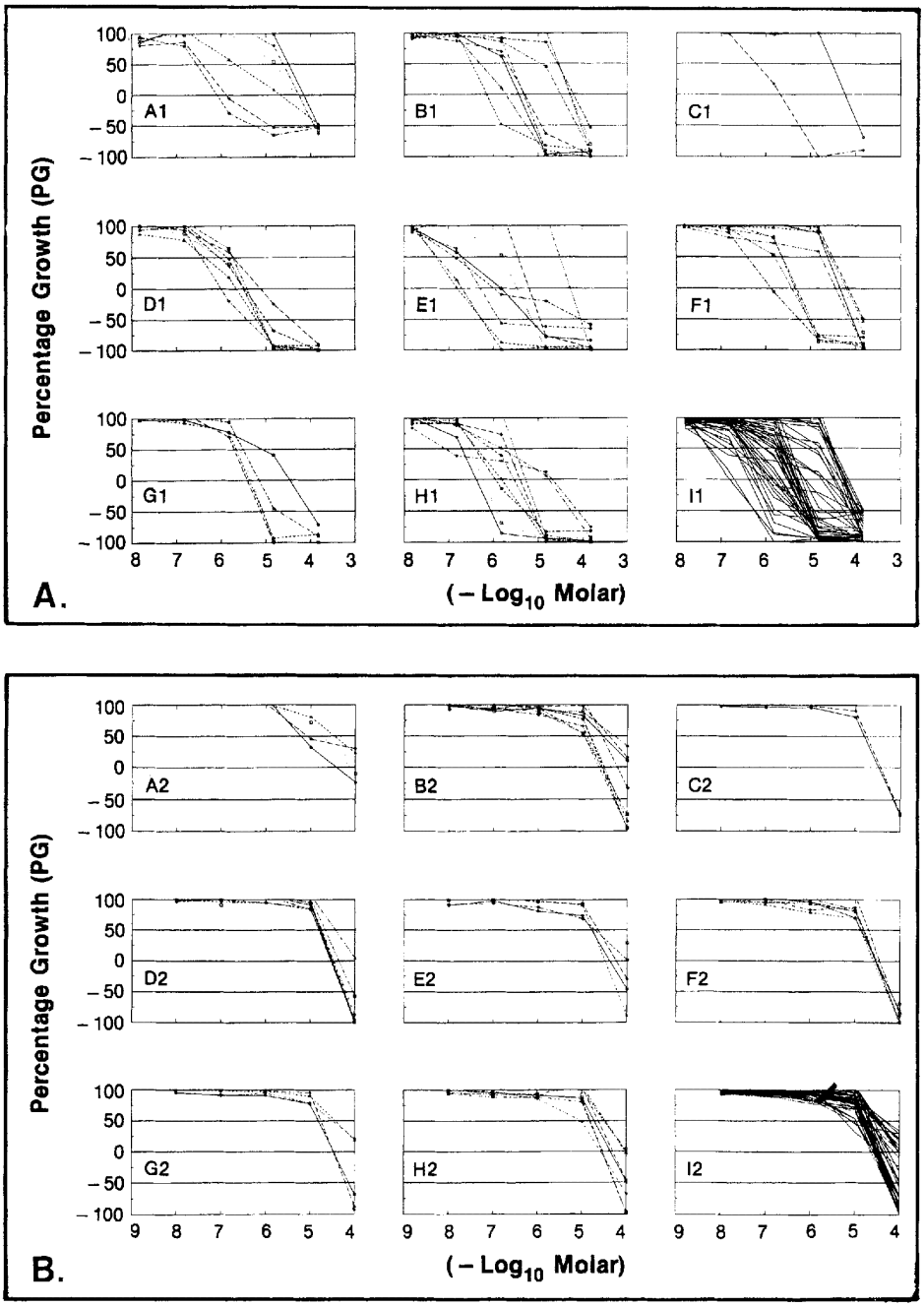

Figure 1. The top composite (A) of nine sets of dose-response curves is from the testing of halomon (structure 1 of Fig. 7) in the NCI in vitro screen. The bottom composite (B) of nine sets of dose-response curves is from the testing of a related natural product (structure 5 of Fig. 7) in the NCI in vitro screen. Individual cell line identifiers have been omitted for clarity. Graphs A1 and A2 are from the leukemia/lymphoma subpanel, graphs B1 and B2 are from the non-small-cell lung cancer subpanel, graphs C1 and C2 are from the small-cell lung cancer subpanel, graphs D1 and D2 are from the colon cancer subpanel, graphs E1 and E2 are from the brain tumor subpanel, graphs F1 and F2 are from the melanoma subpanel, graphs G1 and G2 are from the ovarian cancer subpanel, graphs H1 and H2 are from the renal cancer subpanel, and graphs I1 and I2 are composites of all the respective subpanels together. Reprinted with permission from the American Chemical Society from Fuller et al. [1992]. Copyright 1992 American Chemical Society.

(specified in negative $\log_{10}$ units) is assigned as the default value. In the screening data report, that default value is preceded by a "<" sign, signifying that the "real" value is something "less than" the lowest tested concentration. Likewise, if none of the tested concentrations produces the required PG reference level of effect or greater, then a ">" sign precedes the printed default value (which is the highest tested concentration or HICONC, specified in negative $\log_{10}$ units), signifying that the "real" value is something "greater than" the HICONC. in any case, either the "real" (interpolated) or the default (< or >) GI_{50}, TGI, and LC_{50} for every cell line in the panel are printed with the mean-graphs included in the screening data report. The investigator can, if desired, verify visually that for any printed response parameter concentration preceded by a "<" or ">" for a given cell line in the GI_{50}, TGI, or LC_{50} mean graphs, there must be a corresponding dose-response curve that either lies entirely *below* or entirely *above* the corresponding PG reference line, respectively.

For some applications, the occurrence of many default values for the response parameters in a given screening test can have a major impact on both the accuracy and the interpretation, and therefore the usefulness of the data. This problem may be particularly prominent, for example, in structure-activity studies where both quantitative (e.g., overall or panel-averaged potency) and qualitative (e.g., profile of differential cytotoxicity) comparisons of compounds are desired. For any given compound, the particular range of concentrations tested can be the major determinant of the extent of occurrence of "<" or ">" response parameter values. Therefore, it may be necessary to obtain further testing of a compound in concentration range(s) other than employed routinely in the screen, depending upon the intended use of the data. Indeed, in certain instances, data from the testing of a given compound in different concentration ranges may yield distinctly useful, non-overlapping information. Examples that follow may provide further clarification of these points. Before presentation of specific examples, however, some additional background and descriptive information concerning the "mean-graph" and the COMPARE analysis concepts are pertinent.

Mean-Graph

A mean-graph is a pattern created by plotting positive and negative values, termed "deltas," generated from a set of GI_{50}, TGI, or LC_{50} concentrations obtained for a given compound tested against each cell line in the NCI in vitro screen. Figure 2 shows the GI_{50}, TGI, and LC_{50} mean-graphs derived from the dose-response data of Figure 1. This figure was also prepared directly from the NCI supplier report, by manually cropping and editing the original mean graphs.

The deltas are generated from the GI_{50}, TGI, or LC_{50} data by a three-step calculation. For example, the GI_{50} value for each cell line successfully tested with a given compound is converted to its $\log_{10}$ GI_{50} value. The mean panel $\log_{10}$ GI_{50} value is obtained by averaging the individual $\log_{10}$ GI_{50} values. Both "real" and default values are used in the calculation. Each individual $\log_{10}$ GI_{50} value then is subtracted from the panel mean to create the corresponding delta.

To construct the mean-graph, the deltas are plotted horizontally in reference to a vertical line that represents the calculated mean panel GI_{50}. The mean panel GI_{50} may or may not represent, nor even approximate, a "true" mean, depending upon the extent to which defaults were among the values averaged (see Dose-Response Curves). In any case, the negative deltas are plotted to the right of the mean reference line, thereby proportionately representing cell lines more sensitive than the calculated average. Conversely, the positive deltas are plotted to the left of the reference line to represent the less sensitive cell lines to the given agent. Thus, for example, a bar projecting 3 units to the right of the vertical reference line in a GI_{50} mean-graph indicates that the GI_{50} concentration for that cell line is 1,000 times less than the panel-averaged GI_{50} concentration. The TGI and LC_{50} mean-graphs are prepared and interpreted similarly.

In the full standard NCI screening data report package, three additional numbers are printed at the base of each of the three respective mean-graphs provided. These numbers are the MG-MID, the Delta (not be confused with the "delta" for an individual cell line), and the Range. The MG-MID is the calculated mean panel GI_{50}, TGI, or LC_{50}. The Delta is the number of $\log_{10}$ units by which the delta of the most sensitive line(s) of the panel differs from the corresponding MG-MID. Similarly, the Range is the number of $\log_{10}$ units by which the delta of the most sensitive line(s) of the panel differs from the delta of the least sensitive line(s). The MG-MID, Delta, and Range are most meaningful when few if any default values are contained in the corresponding mean-graph; otherwise, they are not particularly meaningful or useful, and indeed can be misleading. Further clarification of this point follows in a discussion of data presented in Figures 3, 4, 5A and B.

COMPARE

COMPARE is a computerized, pattern-recognition algorithm which has considerable utility in the evaluation and exploitation of data generated by the

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.