A Statistical Model for Predicting Response of Breast Cancer Patients to Cytotoxic Chemotherapy\(^1,2\)

Michael L. Feldstein,\(^3\) Edwin D. Savlov, and Russell Hilf

Division of Biostatistics [M. L. F.], Department of Surgery [E. D. S.], Department of Biochemistry [R. H.], and University of Rochester Cancer Center [M. L. F., E. D. S., R. H.], University of Rochester School of Medicine and Dentistry, Rochester, New York 14642

ABSTRACT

A binary logistic model is used for predicting response to cytotoxic chemotherapy for a breast cancer patient on the basis of her tumor enzyme activity profile. The enzymes used in the model are lactate dehydrogenase, nicotinamide adenine dinucleotide phosphate-isocitrate dehydrogenase, and phosphoglucomutase, all of which were measured on primary tumor specimens from each patient. The statistical model provides an estimate of the probability that an individual will respond to treatment. Chemotherapeutic treatment consisting of combination cytotoxic drugs and subsequent evaluation of patient response followed cooperative group protocol guidelines, including outside review to confirm the patient evaluation.

The model based on this study, which represents 5 years of patient follow-up, correctly predicts clinical outcome in 32 of the 37 cases available.

INTRODUCTION

The possibility of accurately predicting a breast cancer patient's response to a particular mode of therapy has been the subject of much investigation. Bulbrook et al. (1, 3, 8) attempted to predict for response to hypophysectomy on the basis of urinary steroids, using a statistical technique known as discriminant analysis. While this approach is appropriate for data of this type, it presupposes the existence of 2 populations and gives rise to a statistical decision rule that assigns each individual into 1 of the 2 groups (responder or nonresponder). Brennan has pointed out in Ref. 9 a shortcoming of this technique insofar as individuals will be inclined to respond to varying degrees. Furthermore, the assumptions required for inferential purposes (multivariate normality of the discriminatory variables) have never been examined. Dao et al. (6, 7) were concerned with predicting response to adrenalectomy, on the basis of the capacity of the neoplasm to sulfurylate dehydroepiandrosterone relative to estradiol. Statistically, their approach consisted of arbitrarily partitioning the sulfokinase activity in the tumor into mutually exclusive intervals and comparing the respective proportions of patients responding to adrenalectomy. More recently, attention has focused on the presence in breast cancers of estrogen receptors as indicative of clinical response. Several groups of investigators (14), extending the work of Jensen et al. (12), have reported a correlation between the lack of estrogen receptors and failure to respond to hormonal treatment, either ablative or additive. Conversely, the presence of estrogen receptors was correlated with response to hormonal therapy in 50 to 60% of the cases studied.

The purpose of this paper is to apply the statistical technique known as logistic regression to the problem of predicting response to cytotoxic combination chemotherapy. Unlike the discriminant function approach, this method provides the clinician with an estimate of each patient's chance to respond, utilizing as prognostic variables a selected enzyme activity profile, obtained by assay on a portion of the primary tumor. Part of the data, which form the basis of this paper, can be found in the report of Hilf et al. (10), and additional data have since become available. They concluded that a high glycolytic enzyme profile was associated with patients who responded to cytotoxic chemotherapy, while those with lower profiles failed to respond.

MATERIALS AND METHODS

A full account of the laboratory procedures used to measure enzyme activity (expressed as \(\mu\)mol of pyridine nucleotide cofactor, reduced or oxidized per min per mg DNA), as well as a description of the patients under study can be found in the paper of Hilf et al. (10). Briefly, tissue was obtained from patients who underwent mastectomy and subsequently developed disseminated breast cancer, with gross measurable tumor burden. Lesions from all patients were classified as infiltrating ductal carcinoma, and all received a cyclic combination chemotherapy regimen including cyclophosphamide, methotrexate, and 5-fluorouracil. Such treatment and subsequent evaluation of clinical response followed the Eastern Cooperative Oncology Group protocol outlined in Ref. 16, and all of the cases were evaluated by 2 outside reviewers who confirmed the nature of the response. For purposes of this paper, each patient was classified as either a "responder" (R) or a "nonresponder" (NR), the category of "partial responder" being incorporated into "responder." For completeness, the definitions of "response" are given in "Appendix 1."

Not all of the 11 enzymes considered by Hilf et al. (10) were measured for each of the 37 individuals reported in this paper. Missing information is directly related to the amount of tissue available for assay, and a priority system


\(^{2}\) This paper is dedicated to the memory of Joan Bulbrook.

\(^{3}\) To whom requests for reprints should be addressed, at Division of Biostatistics, Box 930, University of Rochester School of Medicine and Dentistry, Rochester, N. Y. 14642.

Received November 10, 1977; accepted May 15, 1978.
based on the report (10) evolved. Thus, GPI, LDH, ICD, and PK, which were the enzymes reported as showing the most significant differences between responders and nonresponders, were almost always done, followed next by G6PD, \( \alpha \)-glycerolphosphate dehydrogenase, glutamate dehydrogenase, PGM, 6-phosphogluconate dehydrogenase, aspartate aminotransferase, and hexokinase, in that order.

Only 3 of the 11 enzymes considered by Hilf are utilized here in an attempt to predict response outcome. The enzymes selected were LDH, ICD, and PGM. The rationale for selecting these enzymes and a discussion of the statistical model utilizing them are given under "Results." As more patients become eligible for study, it is possible that the selected enzymes and the manner in which they appear in the model will have to be modified. The statistical methodology, which is well documented (5, 17), remains unchanged and can easily accommodate such modifications. The actual enzyme data as discussed so far are given in Table 1.

RESULTS

The problem confronting the clinician is that of assessing a patient’s chance to respond to cytotoxic chemotherapy on the assumption that the patient’s enzyme activity profile plays a major role in determining what the response will be. Statistically, the problem is to devise a quantitative relationship between the probability that a patient will respond and the knowledge of the patient’s enzyme activity profile.

Denoting by \( p_i \) the probability that the \( i \)th patient will respond conditional upon knowing the patient’s enzyme activity profile, we may write

\[
 p_i = e^{h_i} / (1 + e^{h_i})
\]

in which \( e \) is the base of natural logarithms, and \( h_i \) is an unknown function of the enzyme profile of the \( i \)th patient. In view of the one-to-one relationship between \( p_i \) and \( h_i \), Equation A makes no assumptions whatever, save that the enzyme profile can be used to predict response. The assumptions arise upon imposing a particular mathematical structure for the function \( h_i \). In general, the latter will depend upon unknown parameters that must be estimated by the data, and a suitable approximation as to its mathematical form will emerge through intensive exploratory data analysis. The function \( h_i \) admits of an interpretation in terms of a "log odds," in the sense that the natural logarithm of the ratio

\[
\text{Probability } i\text{th patient responds} = \text{Probability } i\text{th patient does not respond}
\]

is precisely \( h_i \). The problem then is to find a suitable functional form for the log odds. Equation A corresponds exactly to the logistic distribution (13), in which the 3 selected enzymes will play the roles of regression variables for the binary clinical outcome.

With a view to approximating the function \( h_i \), we have plotted the enzyme activity levels taken 2 at a time for each patient using distinguishing marks for the responders and nonresponders. Because there are 11 enzymes, there are 55 such graphs, although as pointed out earlier not every patient had complete information on all 11 enzymes. The result of this endeavor was the observation that, whenever LDH was present, the observations exhibited a certain polarity in which the responders clustered separately from nonresponders. This behavior was not without minor violation, but no marked clustering was evident in graphs in which LDH did not appear. It is evident on the basis of these data that LDH plays a prominent role in separating responders from nonresponders. In an attempt to keep the functional form for \( h_i \) as simple as possible and in view of the relative importance of LDH, an additive model was considered, namely,

\[
h_i = \beta_0 + \beta_1 \text{LDH}_i + g_i
\]

Here \( \beta_0 \) and \( \beta_1 \) are unknown parameters that must be estimated, LDH denotes the enzyme activity level for the \( i \)th patient, and \( g_i \) is as yet an unspecified function of the remaining enzymes. A variety of choices for \( g_i \) was investigated, with the additional constraint that it contain no more than 1 unknown parameter. This would guard against

<table>
<thead>
<tr>
<th>Patient</th>
<th>LDH</th>
<th>ICD</th>
<th>PGM</th>
<th>Clinical outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10.20</td>
<td>0.81</td>
<td>0.46</td>
<td>NR*</td>
</tr>
<tr>
<td>2</td>
<td>5.60</td>
<td>0.28</td>
<td>0.02</td>
<td>R</td>
</tr>
<tr>
<td>3</td>
<td>3.50</td>
<td>0.51</td>
<td>0.03</td>
<td>R</td>
</tr>
<tr>
<td>4</td>
<td>9.60</td>
<td>0.70</td>
<td>0.08</td>
<td>R</td>
</tr>
<tr>
<td>5</td>
<td>14.50</td>
<td>1.39</td>
<td>0.64</td>
<td>R</td>
</tr>
<tr>
<td>6</td>
<td>16.10</td>
<td>1.23</td>
<td>1.27</td>
<td>R</td>
</tr>
<tr>
<td>7</td>
<td>16.00</td>
<td>0.91</td>
<td>0.17</td>
<td>R</td>
</tr>
<tr>
<td>8</td>
<td>7.40</td>
<td>0.21</td>
<td>0.58</td>
<td>NR</td>
</tr>
<tr>
<td>9</td>
<td>9.40</td>
<td>0.63</td>
<td>0.35</td>
<td>NR</td>
</tr>
<tr>
<td>10</td>
<td>8.90</td>
<td>0.50</td>
<td>0.10</td>
<td>NR</td>
</tr>
<tr>
<td>11</td>
<td>7.40</td>
<td>0.02</td>
<td>0.07</td>
<td>NR</td>
</tr>
<tr>
<td>12</td>
<td>0.50</td>
<td>0.59</td>
<td>0.11</td>
<td>NR</td>
</tr>
<tr>
<td>13</td>
<td>8.90</td>
<td>0.24</td>
<td>0.02</td>
<td>R</td>
</tr>
<tr>
<td>14</td>
<td>16.65</td>
<td>0.48</td>
<td>0.05</td>
<td>R</td>
</tr>
<tr>
<td>15</td>
<td>25.17</td>
<td>1.63</td>
<td>0.08</td>
<td>R</td>
</tr>
<tr>
<td>16</td>
<td>28.58</td>
<td>1.01</td>
<td>0.37</td>
<td>R</td>
</tr>
<tr>
<td>17</td>
<td>22.11</td>
<td>0.50</td>
<td>0.43</td>
<td>R</td>
</tr>
<tr>
<td>18</td>
<td>13.81</td>
<td>0.79</td>
<td>0.54</td>
<td>NR</td>
</tr>
<tr>
<td>19</td>
<td>2.57</td>
<td>0.25</td>
<td>0.03</td>
<td>NR</td>
</tr>
<tr>
<td>20</td>
<td>68.00</td>
<td>3.07</td>
<td>0.71</td>
<td>R</td>
</tr>
<tr>
<td>21</td>
<td>62.66</td>
<td>2.09</td>
<td>0.54</td>
<td>R</td>
</tr>
<tr>
<td>22</td>
<td>7.94</td>
<td>0.40</td>
<td>0.10</td>
<td>R</td>
</tr>
<tr>
<td>23</td>
<td>4.98</td>
<td>0.85</td>
<td>0.04</td>
<td>R</td>
</tr>
<tr>
<td>24</td>
<td>20.92</td>
<td>0.94</td>
<td>0.30</td>
<td>R</td>
</tr>
<tr>
<td>25</td>
<td>17.00</td>
<td>2.28</td>
<td>1.28</td>
<td>R</td>
</tr>
<tr>
<td>26</td>
<td>9.15</td>
<td>0.82</td>
<td>0.16</td>
<td>NR</td>
</tr>
<tr>
<td>27</td>
<td>1.28</td>
<td>0.04</td>
<td>0.04</td>
<td>NR</td>
</tr>
<tr>
<td>28</td>
<td>6.05</td>
<td>0.40</td>
<td>0.40</td>
<td>NR</td>
</tr>
<tr>
<td>29</td>
<td>9.90</td>
<td>0.16</td>
<td>0.13</td>
<td>NR</td>
</tr>
<tr>
<td>30</td>
<td>20.76</td>
<td>1.53</td>
<td>0.06</td>
<td>R</td>
</tr>
<tr>
<td>31</td>
<td>8.53</td>
<td>0.85</td>
<td>0.08</td>
<td>R</td>
</tr>
<tr>
<td>32</td>
<td>9.01</td>
<td>0.75</td>
<td>0.09</td>
<td>R</td>
</tr>
<tr>
<td>33</td>
<td>12.38</td>
<td>1.35</td>
<td>0.13</td>
<td>R</td>
</tr>
<tr>
<td>34</td>
<td>4.49</td>
<td>1.09</td>
<td>0.15</td>
<td>R</td>
</tr>
<tr>
<td>35</td>
<td>30.74</td>
<td>1.65</td>
<td>0.32</td>
<td>R</td>
</tr>
<tr>
<td>36</td>
<td>3.17</td>
<td>0.14</td>
<td>0.13</td>
<td>NR</td>
</tr>
<tr>
<td>37</td>
<td>8.79</td>
<td>0.73</td>
<td>0.05</td>
<td>NR</td>
</tr>
</tbody>
</table>

* NR, nonresponder; R, responder.

The abbreviations used are: GPI, glucosephosphate isomerase; LDH, lactate dehydrogenase; ICD, NADP-isocitrate dehydrogenase; PK, pyruvate kinase; G6PD, glucose-6-phosphate dehydrogenase; PGM, phosphoglucomutase.
overfitting a model to data (i.e., producing a remarkably good predictor of response by including enough parameters to accommodate most of the variability in the data, the danger of which is its failure to predict on a future set of observations). We would therefore need to estimate at most 3 unknown constants on the basis of 37 observations. In order to specify a choice for the function \( g_1 \), we examined the ratio of one enzyme activity to another. Excluding LDH there are 45 such possible ratios. However, in a large number of cases aspartate aminotransferase, hexokinase, and 6-phosphogluconate dehydrogenase were not assayed, and the remaining 7 enzymes (G6PD, PGM, \( \alpha \)-glycerolphosphate dehydrogenase, glutamate dehydrogenase, PK, GPI, and ICD) give rise to 21 ratios to examine. Of these, 2 produced spatial clustering of responders and nonresponders, namely, ICD/PGM and G6PD/PGM; while these were remarkably similar in the patterns they yielded, the former exhibited more distinct clustering of responders from nonresponders. We therefore took

\[
g_1 = \beta_0 (\text{ICD}/\text{PGM}),
\]

giving the full working model

\[
h_i = \beta_0 + \beta_1 \text{LDH} + \beta_2 (\text{ICD}/\text{PGM}),
\]

and we investigated the following choices for the log odds:

\[
h_i = \beta_0 \quad \text{(C)}
\]

\[
h_i = \beta_0 + \beta_1 \text{LDH} \quad \text{(D)}
\]

\[
h_i = \beta_0 + \beta_2 (\text{ICD}/\text{PGM}) \quad \text{(E)}
\]

\[
h_i = \beta_0 + \beta_1 \text{LDH} + \beta_2 (\text{ICD}/\text{PGM}) \quad \text{(F)}
\]

For completeness, we include the model that appeared in the earlier report.\(^3\) It takes the form

\[
h_i = \beta_0 + \beta_1 \text{LDH} + \beta_2 (\text{PK} \times \text{GPI} \times \text{ICD}) \quad \text{(G)}
\]

and was considered because of the finding of Hilf et al. (10) that, among responders, the enzyme activity levels of PK, GPI, and ICD were significantly elevated. The product of these 3 quantities would then have the effect of further exaggerating the differences between responders and nonresponders. The model that performed best, in the sense of maximizing the joint probability of all the observations and in which the presence of each parameter was judged significant (\( p < 0.02 \) for each parameter), was Equation F, the full working model. The unknown constants were estimated by the method of maximum likelihood with the use of an iterative technique. The estimates and their approximate standard deviations and correlations are given in Table 2. Hence, from Equations A and B, the estimated probability that the \( i \)th individual will respond is given by

\[
\hat{p}_i = \frac{e^{\beta_0 + \beta_1 \text{LDH} + \beta_2 (\text{ICD}/\text{PGM})}}{1 + e^{\beta_0 + \beta_1 \text{LDH} + \beta_2 (\text{ICD}/\text{PGM})}}
\]

the symbol over \( p \), indicating that it is an estimate of the true probability. These estimated probabilities have been evaluated for all 37 individuals in the study, and they appear in Chart 1 adjacent to the corresponding point that locates each individual with respect to her LDH level in the horizontal axis and her ICD/PGM ratio on the vertical axis. Superimposed on Chart 1 are the estimated quartile lines of equiprobability. For example, every point on the 0.50 line corresponds to an (hypothetical) individual with an estimated 50% chance of responding. If the 0.50 line is selected for separating responders from nonresponders, then 5 individuals would be misclassified; 2 predicted not to respond who in fact did and 3 predicted to respond who in fact did not. These errors of misclassification obviously change if a different cutoff point is selected. For purposes of assigning subsequent treatment, the choice of a cutoff point is crucial and ought to depend upon the risks attendant to the 2 types of errors.

We have also shown (Table 3) the value of \(-2 \log_e \chi^2\)
In this paper we are concerned with the assessment of the breast cancer patient’s chance of response to cytotoxic chemotherapy. Such regimens as Cytoxan/methotrexate/5-fluorouracil have been reported to induce objective responses in 50 to 60% of patients with advanced disease (2, 4), and it appears that establishing a reliable predictive index would be of considerable value in therapeutic decision making. Such an approach has gained wider publicity for estimation of hormone dependence or responsiveness (11) and analysis of breast carcinomas for presence of estrogen receptors has become commonplace in the management of the patient (14). However, it appears that the absence of estrogen receptors is a more reliable predictor of nonresponse than is the presence of receptor for response to hormonal therapy. To improve the latter, measurement of progesterone receptors has been proposed and preliminary data are quite encouraging (15).

The statistical model used here is appropriate whenever the prediction of binary clinical outcome, e.g., response versus nonresponse, may depend upon 1 or more identifiable characteristics of the patient. These characteristics may be either of the continuous measurement type (e.g., an enzyme activity) or of discrete type (e.g., menopausal status). Conditional upon the individual characteristics of each patient, the model estimates (retrospectively) each individual’s chance of responding. In this respect logistic regression has more to offer than the more classic discriminant function approach. The main distinction between the 2 methods, however, is that the former regards the measured prognostic variables as fixed, whereas the latter treats them as random variables that jointly follow a multivariate normal distribution.

In predicting response to cytotoxic chemotherapy, the logistic model used here utilizes as prognostic factors, enzyme activity data measured on the primary tumor. We have found that LDH activity plays a major role in separating responders (high LDH values) from nonresponders (low LDH values).

The ratio ICD/PGM in conjunction with LDH improves the predictive value over the null model (with none of the enzymes used) considerably, there being an almost 50% reduction in units of log likelihood at the expense of 2 additional parameters. Other enzymes may be useful in further improving the prediction for response, but it would be premature to assess their inclusion in a model based on the 37 observations at hand. Many more observations would be needed in order to assess the relative importance of one enzyme to another insofar as predicting for response is concerned.

The data on the 37 patients reported here represent accrual and follow-up over a 5-year period at this institution. Current studies involving 4 additional institutions are underway to augment the sample size and to test the predictive validity of the model.

ACKNOWLEDGMENTS

We wish to acknowledge the continuing collaboration of Dr. R. A. Cooper, Jr., and Dr. R. A. Orlando who performed the pathological analyses of these
tissues, the project nurse S. Kinsella for her efforts in obtaining tissue and patient follow-up records, and the technical assistance of S. Gibson for the enzyme activity assays. We thank Dr. G. C. Escher and Dr. M. E. Sears for their efforts in evaluating the cases presented here.

APPENDIX 1: DEFINITIONS OF RESPONSE

**Complete Response (CR)**. Complete disappearance of all measurable lesions with no new bone lesions and calcification of osteolytic lesions.

**Partial Response (PR)**. A decrease of 50% or greater in the sum of the products of the 2 largest perpendicular diameters of all measurable lesions and/or partial calcification of osteolytic lesions (regression or no change in osteoblastic lesions).

**No Response (NR)**. A decrease of less than 50% or an increase of less than 25% over original measurements of the sum of the products of the 2 largest diameters of all measurable lesions with no change of the osseous lesions.

**Progression (PR)**. An increase of 25% or greater over original measurements in the sum of the products of the 2 largest perpendicular diameters of all measurable lesions and/or the occurrence of new lesions and/or progression of osteolytic lesions.

REFERENCES


A Statistical Model for Predicting Response of Breast Cancer Patients to Cytotoxic Chemotherapy

Michael L. Feldstein, Edwin D. Savlov and Russell Hilf


Updated version
Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/38/8/2544

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.