References and Notes

1. S. Brenner, *Br. Med. Bull.* **29**, 269 (1973); Y. Hotta and S. Benzer, *Nature (Lond.)* **240**, 527 (1972); K. Ikeda and W. D. Kaplan, *Proc. Natl. Acad. Sci. U.S.A.* **67**, 1480 (1970); C. Kung and R. Eckart, *ibid.* **69**, 93 (1972); E. R. Macagno, V. Lopresti, C. Levinthal, *ibid.* **70**, 57 (1973).
2. P. Rakic and R. Sidman, *J. Comp. Neurol.* **152**, 133 (1973).
3. D. Bentley and R. R. Hoy, *Sci. Am.* **231**, 34 (1974).
4. J. S. Edwards and J. Palka, *Proc. R. Soc. Lond. Ser. B Biol. Sci.* **185**, 83 (1974).
5. B. Mendenhall and R. K. Murphey, *J. Neurobiol.*, in press.
6. R. K. Murphey, in *Intracellular Staining Techniques in Neurobiology*, S. Kater and C. Nicholson, Eds. (Springer-Verlag, Berlin, 1973), p. 135.
7. W. Gnatzy and K. Schmidt, *Z. Zellforsch. Mikrosk. Anat.* **122**, 190 (1971); K. Schmidt and W. Gnatzy, *ibid.* **126**, 206 (1972); H. Sihler, *Zool. Jarhb. Abt. Anat. Ontog. Tiere* **45**, 519 (1924).
8. E. D. Petrovskaya, *Zh. Evol. Biokhim. Fiziol.* **5**, 337 (1969); J. Palka and M. Schubiger, in preparation.
9. J. Palka and J. S. Edwards, *Proc. R. Soc. Lond. Ser. B Biol. Sci.* **185**, 105 (1974).
10. R. K. Murphey and J. Palka, *Nature (Lond.)* **248**, 249 (1974); I. Parnas and D. Dagan, *Adv. Insect Physiol.* **8**, 95 (1971); J. Iles, *J. Exp. Biol.* **56**, 647 (1972).
11. R. K. Murphey, J. Palka, J. S. Edwards, *Am. Zool.* **13**, 1302 (1973); R. K. Murphey, B. Mendenhall, J. Palka, J. S. Edwards, *J. Comp. Neurol.*, in press.
12. A. Globus and A. B. Scheibel, *Exp. Neurol.* **18**, 116 (1967); R. W. Guillery, *J. Comp. Neurol.* **149**, 423 (1973); W. H. Jones and D. B. Thomas, *J. Anat.* **96**, 375 (1962); J. P. Kelley and W. M. Cowan, *Brain Res.* **42**, 263 (1972).
13. C. Blakemore, in *The Third Neurosciences Study Program*, F. O. Schmidt and B. R. Worden, Eds. (MIT Press, Cambridge, Mass., 1973); p. 105; D. H. Hubel and T. N. Wiesel, *J. Physiol. (Lond.)* **206**, 419 (1970); T. N. Wiesel and D. H. Hubel, *J. Neurophysiol.* **26**, 1003 (1963); J. D. Pettigrew, *J. Physiol. (Lond.)* **237**, 49 (1974).

14. I thank A. Raymond for the electron microscopy and for technical assistance, J. Kallo for dye fills of some neurons and W. Kutsch and J. Altman for criticism of the manuscript. Support provided by PHS grant NS 09074 to D.B. and NSF grant GB-38359 to the Electron Microscope Laboratory, University of California, Berkeley.

23 September 1974

Pattern Recognition May Resolve Management of Breast Cancer: Limited Mastectomy versus Radical Mastectomy

In the 7 October 1974 issue of the *Journal of the American Medical Association* two surgeons debated the methods of management of breast cancer. Crile (*1*) took the position that limited mastectomy was the method of choice, while Anglem (*2*) took the position that management of breast cancer should be by radical mastectomy. Both surgeons cited their own studies as well as those of others to support their respective positions.

However, the data base structures cited by the two surgeons were apparently not the same, and as a result it is not possible to compute survival probabilities which are conditioned on different assumptions. I now propose a method for organizing the data base and then evaluating the outcome of different treatments. The method can be applied to existing data as well as to data that may be collected in the future.

Pattern recognition theory and practice is based on classes, features of these classes, and the statistical distribution of these features for the respective classes. In the problem of breast cancer survival there are two basic classes (a class will be generally denoted C).

C_{RY} = survival for Y years after radical mastectomy

C_{LY} = survival for Y years after limited mastectomy

Additional classes can be defined in terms of other treatments or features.

The features of the classes will be inferred from the discussion by the two surgeons and other studies: f_1, cancer stage; f_2, patient age; f_3, histopathology; f_4, lymph node involvement; f_5, treatment; f_6, family history of breast cancer; f_7, immunological status; f_8, patient's cause of death; f_9, location of primary lesion; f_{10}, duration of disease at diagnosis; f_{11}, pre- or postmenopause; f_{12}, associated pregnancy; f_{13}, antibody to breast cancer antigen; and f_{14}, tumor growth rate.

The patient's feature vector may be denoted as

$$\mathbf{f} = [f_1, f_2, \ldots, f_{14}]$$

a column vector. Let the joint probability distribution of $\mathbf{f}$ be $p(\mathbf{f} \mid C_{RY})$ and $p(\mathbf{f} \mid C_{LY})$ for the two classes, respectively.

Given a patient with a particular feature vector $\mathbf{f}$, what is the probability of the occurrence of C_{RY} or C_{LY}?

To answer this question, we must ask what is the probability of C_{RY} or C_{LY} for a patient with breast cancer before we look at $\mathbf{f}$ for that patient? These a priori probabilities will be denoted $p(C_{RY})$ and $p(C_{LY})$, respectively. Then, the probability of C_{RY} or C_{LY} for a patient after looking at his feature vector, called an a posteriori probability, will be denoted $p(C_{RY} \mid \mathbf{f})$ and $p(C_{LY} \mid \mathbf{f})$ for the two respective classes. For the latter notation, the $\mid$ in $p(C_{RY} \mid \mathbf{f})$ is used to mean given: that is, probability of C_{RY} given $\mathbf{f}$.

The well-known Bayes theorem relates the a posteriori probability to the a priori probability as

$$p(C_{RY}|\mathbf{f}) = \frac{p(\mathbf{f}|C_{RY})\ p(C_{RY})}{p(\mathbf{f})}$$

For class C_{RY} (1)

$$p(C_{LY}|\mathbf{f}) = \frac{p(\mathbf{f}|C_{LY})\ p(C_{LY})}{p(\mathbf{f})}$$

For class C_{LY} (2)

where

$$p(\mathbf{f}) = p(\mathbf{f}|C_{LY})\ p(C_{LY}) + p(\mathbf{f}|C_{RY})\ p(C_{RY}) \quad (3)$$

A standard pattern recognition problem is, given "training samples" for class C_{RY} and "training samples" for class C_{LY}, to estimate the respective probability distributions $p(\mathbf{f} \mid C_{RY})$ and $p(\mathbf{f} \mid C_{LY})$. This is not an easy task. First, a model for the probability distribution must be assumed a priori; or in the language of pattern recognition, a "family structure" must be assumed for the probability distribution. Patrick considers such estimation of probability distribution in detail (*3*). There is not one currently available computer procedure that can be used to construct the required estimated probability density functions $p(\mathbf{f} \mid C_{RY})$ and $p(\mathbf{f} \mid C_{LY})$. Any procedure assumes some structure about the functional form of p, whether it is a multivariate Gaussian assumption at one extreme or a "nonparametric" structure at the other extreme (*4–6*). Up to now, analyses of breast cancer data have been dependent on the formation of one-dimensional or two-dimensional probability distributions; for example, with f_1 (that is, the cancer stage) or with the two features f_1 and f_3 (that is, histopathology). Probability distributions might have been constructed for f_1 and f_3, with the use of different values of f_2 (that is, patient age). Another problem is how to use a priori knowledge that two features are statistically dependent.

Thus, given the number of training samples N_R for class C_{RY} and the number of training samples N_L for class C_{LY}, we can estimate $p(\mathbf{f}|C_{RY})$ and $p(\mathbf{f}|C_{LY})$. Once these are estimated and a priori probabilities $p(C_{RY})$ and $p(C_{LY})$ specified, then the a posteriori probabilities $p(C_{RY} \mid \mathbf{f})$ and $p(C_{LY}) \mid \mathbf{f})$ can be calculated for any patient $\mathbf{f}$, where $p(\mathbf{f})$ is calculated from Eq. 3.

However, difficulties can be anticipated. First, what are the a priori probabilities $p(C_{RY})$ and $p(C_{LY})$? Might one assume $p(C_{RY}) = p(C_{LY}) = 1/2$? Did Crile (*1*) and Anglem (*2*) cause a priori

probabilities (by implication) to depend on patient age or cancer stage, when properly age and stage are part of the feature vector **f** and should have their effect by adding to the dimensionality of $p(\mathbf{f}|C_{RY})$ and $p(\mathbf{f}|C_{LY})$, respectively?

Many of the studies cited by Crile (*1*) and Anglem (*2*) involve different parts of the feature vector **f**. Thus, one study might provide training samples for one part of $p(\mathbf{f} \mid C)$ while another study might provide training samples for another part of $p(\mathbf{f} \mid C)$.

To conclude, any national effort to evaluate future breast cancer data or to reevaluate past breast cancer data should follow guidelines established by a recognized group of experts from the surgical community and the pattern recognition–statistician community. This group of experts should (i) determine a set of features to be included in **f** and the values which these features can have; (ii) agree on various a priori structures to be used in obtaining the estimated probability distributions $p(\mathbf{f} \mid C_{RY})$ and $p(\mathbf{f} \mid C_{LY})$, and agree on a priori relationships among the features in **f**; (iii) establish a computer data bank for patient samples and guidelines for this data to be accessed; and (iv) agree on values for $p(C_{RY})$ and $p(C_{LY})$. (v) Thus, estimates will be obtained for $p(\mathbf{f} \mid C_{RY})$ and $p(\mathbf{f} \mid C_{LY})$ for Y (years) = 5, 10, 15 (for example).

EDWARD A. PATRICK
Department of Electrical Engineering, Purdue University, West Lafayette, Indiana 47907, and *Indiana University School of Medicine, Indianapolis*

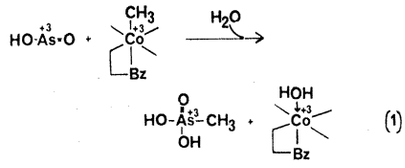
References

1. G. Crile, Jr., *J. Am. Med. Assoc.* **230**, 95, 106 (1974).
2. T. J. Anglem, *ibid.*, pp. 99, 108.
3. E. A. Patrick, *Fundamentals of Pattern Recognition* (Prentice-Hall, Englewood Cliffs, N.J., 1972).
4. ———, L. L. Shen, F. P. Stelmack, "Introduction to the theory of medical consulting and diagnosis," Proceedings of the 1973 National Computer Conference.
5. ———, *IEEE (Inst. Electr. Electron. Eng.) Trans. Syst. Man Cybern.* **4**, 1 (1974).
6. L. B. Lusted, *New Engl. J. Med.* **284**, 416 (1971).

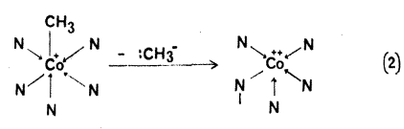
21 October 1974; revised 13 December 1974

The Methylation of Arsenic Compounds

In Wood's excellent article on the subject of the biological methylation of metals in the environment (*1*) and in a series of other papers (*2*) he, and his co-workers, have made considerable use of the concept of formal oxidation numbers. However, this concept must be used with considerable caution. In the proposed methylation of arsenite to methylarsonic acid (*3*) by cobalamin, the formal oxidation number of arsenic remains 3+ on either side of the equation. In order that the oxidation number of cobalt remains unchanged (Eq. 2) the methyl group must be transferred as a methyl carbanion. However, in order for the oxidation number of arsenic to remain unchanged, the methyl group must have a formal oxidation number of 1+. This is inconsistent with the methyl carbanion transfer mechanism.

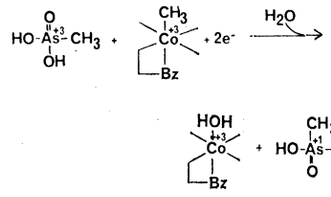


(1)



(2)

In the equation proposed for the methylation of methylarsonic to dimethylarsinic acid (*3*), because the methyl group is assigned a formal oxidation



(3)

number of 1+, the arsenic atom on the right-hand side of the equation ends up with a formal charge of 1+. In order to account for the change in the formal oxidation number of arsenic, two electrons are added to the left-hand side of Eq. 3. What is happening here can be explained by the displacement of OH^- by CH_3^-. This is consistent with Wood's methyl carbanion transfer mechanism (Eq. 2). However, neither an oxidation, nor a reduction, is taking place and no electron transfer is involved. The failure to recognize that the change in the oxidation number arises from the formal application of a set of rules, and not to electron transfer, could lead the unwary reader to seek the assistance of a biological electron transport system when none need be invoked.

In Wood's article (*1*), figure 3, which outlines a proposed biological cycle for arsenic, makes liberal use of the concept of formal oxidation numbers. As has been pointed out in the foregoing examples, the failure to recognize that the oxidation numbers are artificial can result in misleading implications.

RALPH A. ZINGARO
KURT J. IRGOLIC
Department of Chemistry, Texas A & M University, College Station 77843

References and Notes

1. J. M. Wood, *Science* **183**, 1049 (1974).
2. ———, F. S. Kennedy, C. G. Rosen, *Nature (Lond.)* **220**, 173 (1968); J. M. Wood, F. S. Kennedy, R. S. Wolfe, *Biochemistry* **7**, 1707 (1968); N. W. Penley, D. G. Brown, J. M. Wood, *ibid.* **9**, 4302 (1970); J. M. Wood, *Adv. Environ. Sci. Technol.* **2**, 39 (1971); B. C. McBride and R. S. Wolfe, *Biochemistry* **10**, 4312 (1971); M. J. Wolin, E. A. Wolin, R. S. Wolfe, *Biochem. Biophys. Res. Commun.* **15**, 420 (1964); B. C. McBride and R. S. Wolfe, *Bacteriol. Proc.* (1969), p. 130.
3. J. M. Wood, in *Sixth International Conference on Organometallic Chemistry*, M. D. Rausch and S. A. Gardner, Eds. (Univ. of Massachusetts Press, Amherst, 1973), P9.

15 April 1974; revised 30 July 1974

I would like to thank Ralph Zingaro and Kurt Irgolic for pointing out some of the problems that can arise from application of the concept of formal oxidation numbers. However, inorganic chemists consistently write about the reduction of mercuric ion to mercurous ion, but when I apply the rule of formal oxidation numbers, I discover that the formal oxidation state of each atom of mercury in mercurous ion is still 2+. I have met distinguished inorganic chemists who are adamant about the application of formal oxidation states, and an equal number who are adamant about valence. Thus it makes it difficult for a biochemist like myself to find consistency. I don't know where Zangaro and Irgolic stand, but perhaps they could urge their colleagues to get themselves sorted out! I thank them for raising the issue, and I give no arguments.

J. M. WOOD
Freshwater Biological Institute, University of Minnesota, P.O. Box 100, Navarre 55392

6 December 1974