Computer-aided analyses in diagnostic histopathology

IVOR R. H. KRAMER

M.D.S., F.D.S.R.C.S., F.F.D., R.C.S.I., F.R.C.Path.

Department of Pathology, Institute of Dental Surgery

ALMOST every branch of pathology has derived benefit from greater objectivity and greater quantitation. However, diagnostic histopathology remains obstinately subjective and non-quantitative. Although a well developed art, it is heavily dependent on experience acquired over a long period of time. Even so, experienced pathologists will sometimes reach confident but differing diagnoses on the same set of sections.

It was because of dissatisfaction with current methods in diagnostic histopathology that the present attempt was made to introduce greater objectivity into the diagnostic process, and, at the same time, to gain a greater insight into present, largely subjective, methods.

The area chosen for this work was the group of lesions sometimes referred to as the 'white lesions' of the oral mucosa. This group includes the ill defined conditions of leukoplakia, and non-diagnostic keratosis, together with lichen planus and certain other conditions that are more clear cut and easily recognizable. This group was chosen for several reasons: firstly, many biopsies are available from disorders of this type; secondly, differential diagnosis is well known for its difficulty and subjectivity; thirdly, certain of these white lesions are known to be precancerous but it is difficult to predict which particular cases are most likely to proceed to squamous cell carcinoma.

By conventional methods, when the pathologist examines a section he observes a complex visual pattern. By processes that are partly conscious and partly subconscious he interprets this pattern in terms of cells and intercellular materials, of numbers and spatial relationships, and he relates these observations to his concept of the normal. Having decided, subjectively and non-quantitatively, that an abnormality exists, he now compares this abnormality with his memory store of information and experience, to determine whether the observations make a meaningful pattern that will lead to a diagnosis.

How, then, could the computer help in this process? Despite all of the recent progress in automatic processing and recognition of visual data, the pattern presented by the conventional histological section is far too complex for automated analysis. Therefore, it is not yet possible to eliminate completely the human observer. However, it is possible for a microscopist, with relatively little training, to note histological features according to a carefully constructed set of definitions, and to record these features without any attempt at diagnostic interpretation. For example, in the context of mucosal lesions, criteria can be defined for each component of the epithelial changes and for density of inflammatory cell infiltration (using, if necessary, photographic standards). For each histological feature or variable, the human observer is required to record a simple 'yes' or 'no', or to give a roughly quantitative answer according to defined criteria. Now, for each case this record provides a fairly complete set of data which hopefully characterize the histological appearances, and these data can be fed to the computer for analysis.

The present study was based on a retrospective survey of biopsies from 248 cases. Of these, there were 235 consecutive cases that had been diagnosed by conventional methods as keratosis, leukoplakia or lichen planus. To these were added thirteen cases of invasive carcinoma: these were added as markers, and the purpose will be seen later. For each of the 248 cases, carefully calibrated observers recorded seventy histological variables or grades without knowledge of the clinical data (other than the site from which the biopsy had been taken) and without knowledge of the diagnosis that had been made. Later, these seventy variables were simplified to forty-one variables and converted to the binary coding used in the computer analyses. Detailed descriptions of the histological features used were given by Kramer *et al.* (1970a).

The first type of analysis employed was cluster analysis, and the basic principle can be illustrated by an analogy.

Suppose that, in 100 individuals, the height and weight were measured. If the relationship between these two parameters were plotted, the result might

FIG. 1. Three possible results of plotting height against weight. In (a) there is no obvious relationship. In (b) there is a linear relationship. In (c) it is apparent that the individuals fall into two groups or 'clusters' (from Kramer, 1969).

take a variety of forms: the plot might look like Fig. 1a from which it would appear that height and weight were unrelated, or it might look like Fig. 1b which suggests a direct relationship. If the plot looked like Fig. 1c, although in addition to the indication that the parameters were related, it would be noticed that the results formed two groups or clusters, suggesting that the population sampled contained two different kinds of individual.

This clustering occurs because, in the space defined by the parameters chosen, the individuals lie close together—in other words, the individuals in one cluster have similar characteristics (in relation to the parameters measured) and the characteristics of one cluster are different from those of the other clusters. Furthermore, account can be taken of the distance between clusters as an indication of the amount of difference between one cluster and another: in many ways, this is like two-dimensional chromatography.

If only two parameters are used, it is easy to plot the position of each individual, and to obtain some information about clustering by simple inspection of the two-dimensional diagram. If three parameters were used, such as height, weight and age, a three-dimensional model could be constructed, or a diagram that conveyed the impression of three dimensions. However, if it is necessary to use more than three parameters it is no longer possible to plot the results in real space, and the term 'hyper-space' is used.

The computer can locate individuals in hyper-space, or multidimensional space, using almost any number of parameters, and this is the technique known as cluster analysis. In its output, the computer allocates individuals to clusters, it can express mathematically how much variation there is within each cluster, and the distance or difference between one cluster and another.

In the present study, the computer was supplied with data on the forty-one variables for each of the 248 cases. It was programmed to analyse the similarities and differences between each case and every other case, to place together in a group or cluster those cases which were similar to one

another, and to do this in such a way that within-cluster similarities are maximized, and the between-cluster similarities are minimized. The program also provided for instruction regarding the number of clusters to be formed.

The cluster analysis program used has been described elsewhere (Kramer *et al.*, 1970a).

Subjected to the cluster analysis were 187 cases originally diagnosed as belonging to the 'leukoplakia' or non-diagnostic keratosis group, forty-eight diagnosed as lichen planus, and the thirteen marker cases of invasive carcinoma.

In the first analysis, the computer was programmed to form firstly two clusters and then three clusters. The cases placed by the computer in each cluster were then analysed to see what diagnoses had been made on them by conventional methods (Table 1). There was some discrimination between the clusters: for example, in the three cluster analysis, the computer had placed all the carcinomas in cluster 1, the majority of the leukoplakia-keratosis group were placed in cluster 2, and most of the lichen planus cases in cluster 3. However, each cluster is 'impure', in that it contains a mixture of cases of differing diagnoses.

TABLE 1. The 248 biopsies divided into two and three clusters by the computer on the basis of the recorded histological features. Each cluster formed by the computer has been analysed to show the original histological diagnoses

Original diagnosis	Number of biopsies	Two cluster analysis		Three cluster analysis		
		1	2	1	2	3
Carcinoma	13	13	—	13	—	—
Leukoplakia/keratosis	187	21	166	22	154	11
Lichen planus	48	4	44	1	18	29

The computer was now used to re-analyse all the cases, this time to four, five, six or seven clusters, and the results of the six and seven cluster analyses are shown in Table 2.

In the seven cluster analysis, out of the original mixture of 248 cases, cluster 2 contains 102 cases, of

TABLE 2. The same cases as in Table 1, now divided by the computer into six and seven clusters

Original diagnosis	No. of biopsies	Six cluster analysis						Seven cluster analysis						
		1	2	3	4	5	6	1	2	3	4	5	6	7
Carcinoma	13	12	—	—	1	—	—	11	—	—	—	—	—	2
Leukoplakia/ keratosis	187	11	110	6	20	28	12	9	101	1	21	27	12	16
Lichen planus	48	—	2	34	2	7	3	—	1	33	—	7	3	4

which all but one are of the leukoplakia-keratosis group, and cluster 3 contains thirty-four cases, of which all but one are lichen planus.

It is evident, therefore, that on the basis of the coded information about the histological features of each case, this program is capable of a surprising degree of separation of cases into their various diagnostic groups—surprising, in view of the fact that each histological feature was given equal importance, there being no weighting of the histological changes thought to be most important in making a diagnosis.

Another feature of this analysis relates to the cases that the computer has put into the same cluster as the main group of carcinomas.

This was a retrospective survey, and it was known that, out of the original 187 cases in the leukoplakia-keratosis group, 4.8% developed a malignancy. Reference to the six-cluster analysis (Table 2) shows that eleven cases were placed with the carcinomas; of these eleven, 36% have become malignant. Clearly, therefore, the cluster analysis is tending to select the leukoplakias that are most likely to become malignant, and is separating these from the main bulk of cases. No reference has been found to the application of this type of computer-aided analysis to the problems of precancerous lesions of the cervix, or other sites; it seems likely that such a study would be profitable.

The selectivity of the cluster analysis might be further improved if a weighting factor could be determined for each variable, so that more importance is accorded to some histological features than to others. This might be done subjectively, on the basis of experience, or it can be done objectively by the computer using the technique of discriminant analysis. In discriminant analysis the computer is supplied with data on two or more separate groups of cases and is programmed to determine how best to discriminate between these stated groups.

The program written for this purpose produces, for each histological feature, a weighting factor: the application of all these weighting factors produces the best possible separation or discrimination between the groups. The program also yields a mathematical assessment of the degree of separation thus achieved,

and of the statistical significance of this separation (Kramer *et al.*, 1970b).

Another part of the present study shows how this discriminant analysis works. There were sixty cases of the more severe type of keratosis or 'leukoplakia' and at the time this analysis was done it was known that of these sixty, eight had become malignant. Could computer analysis have helped determine, in advance, which of the original sixty leukoplakias were most likely to become malignant?

The cases were split into two groups, comprising the eight that later became malignant, and the fifty-two that did not. Discriminant analysis of the histological features of these two groups produced a non-overlapping separation that was significant at the 5% level (Kramer *et al.*, 1970b).

As mentioned earlier, discriminant analysis calculates a weighting factor for each histological variable, in order to discriminate between two groups as efficiently as possible. The implications are clear. In his diagnostic processes, every pathologist regards some tissue changes as more important than others. For example, in assessing the chances of malignant change in a stratified squamous epithelium, mild spongiosis will be regarded as of little importance, acanthosis as more important, and abnormal mitoses of greater importance still. Discriminant analysis gives a quantitation, an objectively calculated value, to these concepts of 'less important' and 'more important', and it can bring to light discriminators that were not previously recognized as valuable.

For example, in the discriminant analyses between the leukoplakias which did not become malignant and those which did, the computer gave the variables listed in Table 3 as the eight most important characteristics separating the future malignancy group from the others. Of these eight, seven would have been obvious—that is to say, they accorded with generally accepted knowledge. However, it had not been realized that the presence of Russell bodies had prognostic significance in this context. It was of interest that the discriminant analysis also indicated that, in this group of lesions, the intactness or otherwise of the basement membrane was of little prognostic significance (it should be noted that the basement membrane referred to here is the PAS-

FIG. 2. Polar vector diagrams showing the histological characteristics of the first three clusters in the seven-cluster analysis. (a) consists mainly of 'severe' leukoplakias and carcinoma, (b) consists mainly of keratosis cases, and (c) consists mainly of cases diagnosed as lichen planus (from Kramer, El-Labban and Sonkodi, 1974).

TABLE 3. The eight most important variables identified by the computer in the discrimination between the leukoplakias that subsequently developed carcinoma and those that did not

Abnormal mitoses in the stratum spinosum
Disturbance in the polarity of the basal cell layer
Abnormal mitoses in the basal cell layer
Nuclear hyperchromatism in the epithelium
Russell bodies in the lamina propria
Enlarged nucleoli in the stratum spinosum
Epithelial cell pleomorphism
Intra-epithelial keratinization

positive structure immediately subjacent to the epithelium. Commonly, statements are made about the state of the 'basement membrane' as seen in sections stained with HE: such statements usually imply an impression of the architecture of the dermoepidermal junction, rather than the state of a particular structure or tissue component).

Reference was made earlier to the diagnostic purity of certain of the clusters: in other words, certain clusters consisted almost entirely of cases that, by conventional methods, had been placed into a single diagnostic group. These clusters were formed entirely by the computer from analysis of the histological data, without reference to the diagnosis made on each case by conventional methods. Thus, a diagnostically pure cluster might represent a group of cases typical of one of the diagnostic groups. Further analyses can now be performed on such clusters, to determine what is the typical histological pattern defining the cluster, and expressing quantitatively the various features of the pattern.

If this typical histopathological pattern is to be presented quantitatively, a convenient method is to plot the results in the form of polar vector graphs, in which each histological variable is on a radius of a circular graph, and its quantitative importance in

defining the characteristics of the cluster is indicated by the distance from the centre. For example, Fig. 2a shows the computer-calculated 'finger-print' of the histopathology of typical 'severe' leukoplakia and carcinoma, Fig. 2b shows the characteristics of a cluster composed mainly of keratosis cases, and Fig. 2c shows the characteristics of lichen planus. Identification of each variable in these diagrams is given elsewhere (Kramer, El-Labban and Sonkodi, 1974). It can be seen that the histological fingerprints of the different diagnostic groups have very different forms. It is now possible to take any new case, and calculate how closely it approximates to any of these characteristic patterns.

From this short account of work described more fully elsewhere, it will be clear that these are very early steps in this approach to less subjective diagnosis, and clearly the computer will not replace the histopathologist in the foreseeable future. However, at this early stage, there have been occasions when the computer has suggested that the diagnosis made by conventional methods was wrong, and when these cases were reviewed it often appeared that the computer was probably right. It seems reasonable to suppose that further studies of this type will help to improve present diagnostic and prognostic methods—to make them more of a science and less of an art.

Acknowledgments

I am indebted to my co-authors in the original papers on which this summary is based; to Mr Michael Clarke for his invaluable help with the computer analyses; to the Editors of *Annals of the Royal College of Surgeons of England* and the *British Journal of Cancer* for permission to reproduce tables and figures originally published in those journals, and to the British Empire Cancer Campaign for Research for financial support.

References

- KRAMER, I.R.H. (1969) Precancerous conditions of the oral mucosa. *Annals of the Royal College of Surgeons of England*, **45**, 340.
- KRAMER, I.R.H., LUCAS, R.B., EL-LABBAN, N. & LISTER, L. (1970a) A computer-aided study on the tissue changes in oral keratoses and lichen planus, and an analysis of case groupings by subjective and objective criteria. *The British Journal of Cancer*, **24**, 407.
- KRAMER, I.R.H., LUCAS, R.B., EL-LABBAN, N. & LISTER, L. (1970b) The use of discriminant analysis for examining the histological features of oral keratoses and lichen planus. *The British Journal of Cancer*, **24**, 673.
- KRAMER, I.R.H., EL-LABBAN, N.G. & SONKODI, S. (1974) Further studies on lesions of the oral mucosa using computer-aided analyses of histological features. *The British Journal of Cancer*, **29**, 223.