AUTOMATED DIAGNOSIS OF MALIGNANT AND OTHER ABNORMAL CELLS BY FLOW-CYTOMETRY USING THE DIAGNOS1 PROGRAM SYSTEM

G.Valet[1], H.H.Warnecke[2] and H.Kahle[1]

1) Mildred-Scheel-Labor für Krebszellforschung, Max-Planck-Institut für Biochemie, D-8033 Martinsried, FRG
2) Frauenklinik Zentralkrankenhaus, D-8900 Augsburg, FRG

1. INTRODUCTION

The automated recognition of malignant cells and more generally of abnormal cells by flow-cytometry is of great interest for clinical diagnosis but also for research. Single parameter flow-cytometry such as the determination of DNA-aneuploidy alone is not sufficient to detect all malignomas because many malignant tumors do not contain a detectable number of DNA-aneuploid cells.

The approach of this work is to measure several functional cell parameters simultaneously. The data are exhaustively evaluated and the results are written into a self learning database. A multifactorial analysis of the data is then performed to achieve automated diagnosis.

2. MATERIAL AND METHODS

Freshly taken smears of human cervix uteri cells on cotton swabs were first smeared onto fat free glass slides. The cells were fixed with Merckofix spray (Merck, Darmstadt, FRG) and subsequently automatically Papanicolaou stained in a Varistain 24-3 staining apparatus (Shandon, Frankfurt, FRG). The stained slides were embedded with Eukitt and classified according to the Papanicolaou classification by an approved and experienced cytologist.

The cotton swabs were then immersed into 50ml plastic tubes with conical bottom containing 10ml cold HBS-buffer (0.15M NaCl, 10mM HEPES, pH 7.35). The tubes were closed and vigorously shaken about 20 times to loosen the cells from the cotton swab. The cotton swab was then removed, 30ml cold HBS-buffer were added and the cells were washed twice by 10min centrifugation at 200g with 40ml of cold HBS-buffer.

CLINICAL CYTOMETRY AND HISTOMETRY
ISBN 0-12-143240-8

The cell sediment was finally resuspended in 1ml HBS-buffer and stored at 0°C.

250µl of this cell suspension were stained for 5min at room temperature by the addition of 5µl of an ADB/PI staining cocktail (cyto-P-check, Paesel, Frankfurt, FRG). The dye cocktail contained 1mg/ml of the esterase and pH-indicator dye 1,4-diacetoxy-2,3-dicyano-benzene (1, 2) for the staining of vital cells and 2mg/ml of the dye propidium-iodide to stain the DNA of dead cells. The solvent of the cocktail was dimethyl-formamide. DCH stained, monodisperse latex particles of 5µm diameter (65 μm^3) were admixed to the cocktail and served as internal standard for cell volume and cell fluorescence.

The stained cell samples were measured for 2min in a FLUVO-METRICELL II flow cytometer (3) (HEKA-Elektronik, Lambrecht/Pfalz, FRG). The cell volume was determined electrically at a current of 0.23mA in HBS-buffer after hydro-

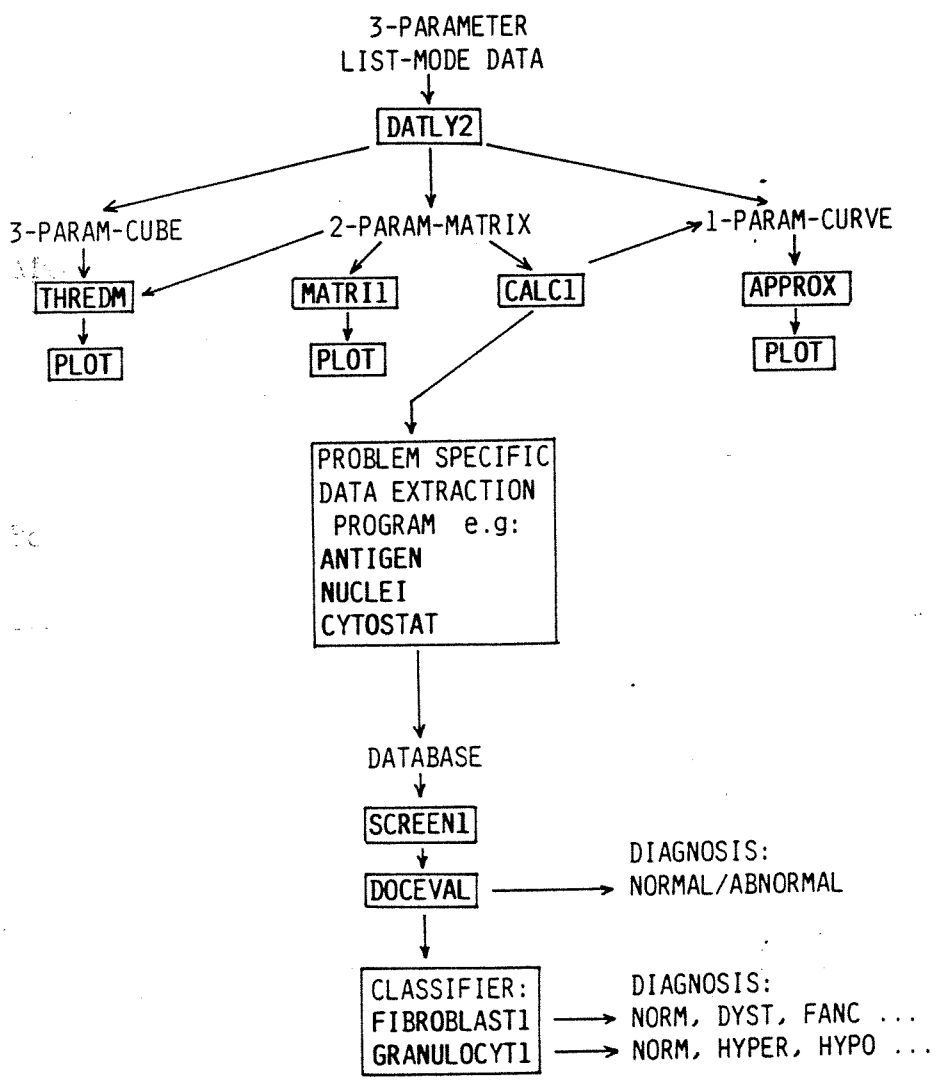

Fig.1 Schematic representation of the DIAGNOS1 program system for automated diagnosis of abnormal cells

dynamic focusing of the cells through the center of the cylindrical orifice of 80µm diameter and 100µm length. The fluorescence of both dyes was excited by a HBO-100 high pressure mercury arc lamp between 300 and 400nm. ADB is cleaved by intracellular esterases and the pH-sensitive dye DCH (2,3-dicyano-hydrochinon) is liberated. The blue DCH-fluorescence was collected simultaneously between 418 and 440nm (F1) and the green DCH fluorescence and red propidium iodide fluorescence between 500 and 700nm (F2). The F1/F2 ratio of the vital cells is a measure of the intracellular pH. The volume and fluorescence pulses of each cell were amplified by 2.5 decade logarithmic amplifiers. The maximum amplitude of each signal was digitized by 128 step analog-digital converters. The digitized pulses were collected on-line as list-mode data (4) on magnetic tape and analyzed with the Fortran programs of the DIAGNOS1 program system (fig.1).

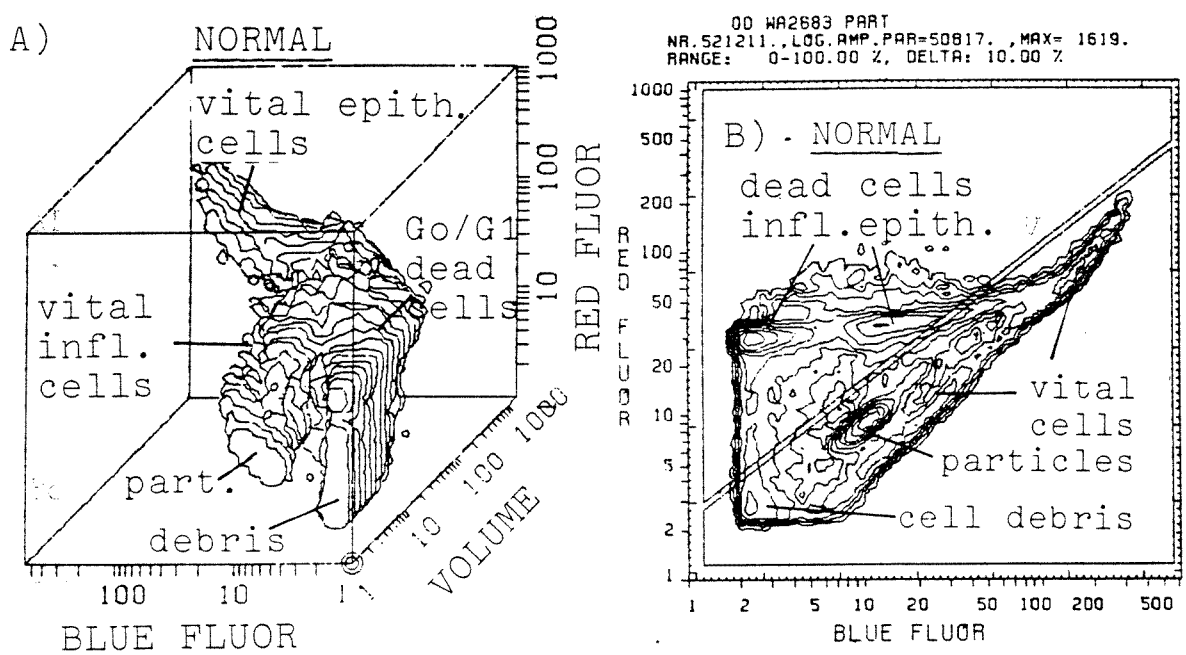

Fig.2a Cloud display (5) of an ADB/PI stained human cervix cell sample. The contour lines are plotted at 10% of the maximum logarithmic channel content (1640 cells). Vital and dead cells as well as standardization particles and cell debris can be distinguished. 50610 cells were measured.

Fig.2b Projection of the cube of fig.2a onto the blue/red fluorescence plain to distinguish between vital and dead cells. The logarithmic channel contents of the graph are standardized to the maximum channel content (1619 cells) as 100%. Contour lines are plotted at 10% linear intervals downwards. (infl. = inflammatory cells (lymphocytes, granulocytes), epith. = epithelial cells)

3. RESULTS

The DATLY2 program (fig.1) first generated the 3-parameter cube matrix (fig.2a) and the two parameter projections (fig.2b, 3a, 3b) from the list-mode data.

The contents, the mean values in x- and y-direction, the mean ratios y/x and the coefficients of variation of all mean values for each window of the two dimensional graphs were calculated by program CALC1 (fig.1) using either preset (fig. 2b,3a) or self-adjusting, "intelligent" windows (fig.3b). The first window was set by the program around the peak of the G0/G1 phase of the epithelial cells (fig.3b). The DNA thresholds were then set 2 classes above and below the peak cellular G0/G1 DNA content. The left volume boundary of the epithelial cell window was set 6 classes to the left of the G0/G1 peak. The DNA-boundaries of the G0/G1 phase of the inflammatory cells and the cell nuclei were adjusted in the same way as for the epithelial cells. The volume boundary between inflammatory cells and bare nuclei was preset at volume class 11 and remained unchanged. After completion of thresholding for the "leading" windows, the boundaries of all "dependent" windows were adjusted to touch the boundaries of the leading windows. All adjustments were made without operator interference.

The CALC1 program furthermore automatically identified DNA aneuploidy of the epithelial cells. Aneuploidy was indicated by the program when separate DNA peaks higher than 20% of the normal G0/G1 DNA peak were found in the DNA distribution of the epithelial cells. Cell samples with less than 200 cells in the DNA distribution of the epithelial cells were rejected from the diagnostic analysis. Near aneuploid tumor cells which formed shoulders within the G0/G1 DNA peak were not identified by the program but the DNA shoulder of such samples usually lead to a significant increase of the number of cells in the S+G2/M compartment of the epithelial cells which was recognized by the subsequent analysis. All values determined by the CALC1 program during the analysis of the two parameter histograms were sequentially written into the binary CALC1 result file which was subsequently used by the task specific data extraction program. The CALC1 result file contained data from the two parameter curves of fig.2a, 3a, 3b and further from a cell volume versus intracellular pH histogram (not shown).

Program VAGPHPI is the data extraction program for cervix cells. It automatically took 50 predetermined values for each cell sample from the CALC1 result file and appended them to the editable database file: VAGPHPI.DO2. On the

first run the VAGPHPI program had written a standardized header into the database to identify all data columns.

The program SCREEN1 (fig.1) subsequently learned automatically the differences between 100 radomly labelled normal cervix samples (Papanicolaou I, II) and 20 randomly labelled malignant samples (Papanicolaou IV,V) by evaluating the respective database records. In a first step the mean values and standard deviations of each of the 50 data columns were calculated for the normal and for the malignant samples and written into a reference file. In a second step the respective cumulative distributions of the normal and malignant samples were calculated for each data column of the database. These one parameter distributions were plotted by program APPROX (fig.1). The SCREEN1 program knew for each data column from the reference file whether the mean value of the malignant cells was below or above the corresponding mean of normal cells. If the mean of the malignant cells was below the mean of the normal cells, the data value at the 10% percentil of the cumulative distribution of the normal

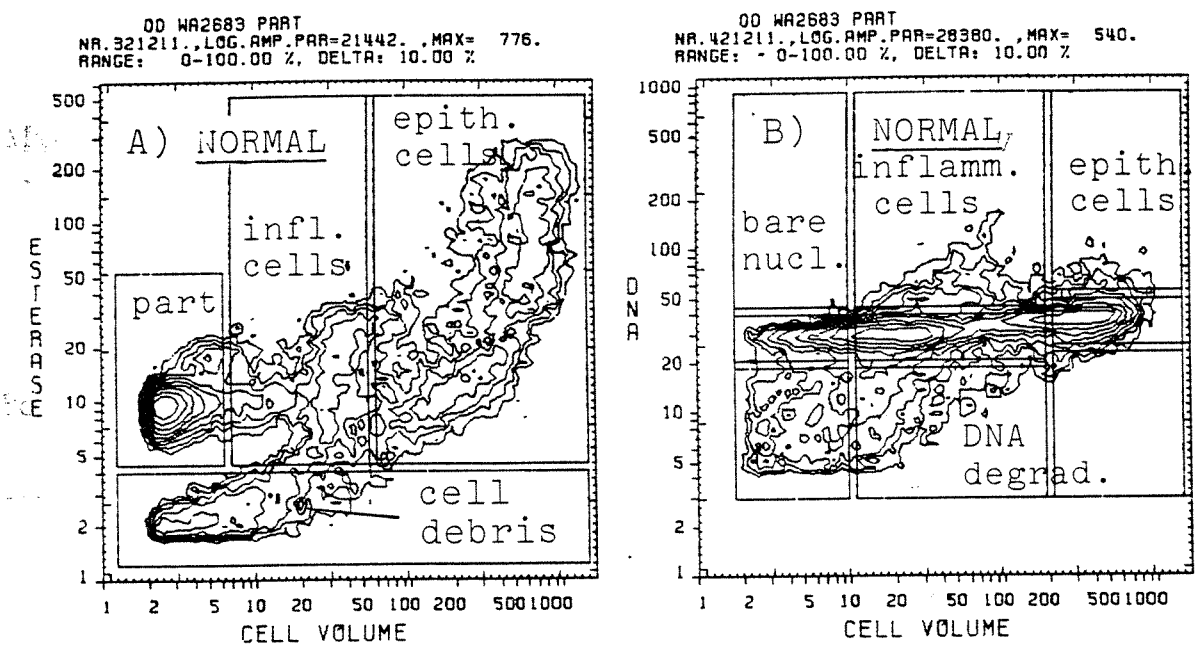

Fig.3a Volume versus esterase activity of the vital cells of fig.2b. Inflammatory and epithelial cells can be distinguished as well as standardization particles and cell debris. The maximum channel content is 776 cells and 21440 cells and particles were measured. One logarithmic volume class corresponds to 26μm^3.

Fig.3b Volume versus DNA of the dead cells of fig.2b. The DNA degradation, the G0/G1-phase and the S+G2/M phase compartments of bare nuclei, inflammatory and epithelial cells are distinguishable. The maximum of the histogram corresponds to 540 cells. A total of 28380 cells was measured.

cells was written as threshold into the reference file. Samples below the 10% percentil value were classified as abnormal. If the mean of the malignant cells for this data column was above the normal samples, then cell samples with values above the 90% percentil of the cumulative distribution were labelled as abnormal. The 0/100%, 1/99%, 2/98%, 5/95% or the 15/85% percentils can be also be used as thresholds (specificity) instead of the 10/90% percentils which are the default thresholds. The SCREEN1 program automatically uses aneuploidy and the five most distinctive data columns for the detection of abnormal samples (sensitivity). The most distinctive parameters of the SCREEN1 analysis besides aneuploidy were: Ratio of % epithelial cells in the S+G2/M phase divided by % inflammatory cells in the S+G2/M phase (+), esterase concentration (+), intracellular pH (-), volume of the G0/G1 phase epithelial cells (-) and the DNA staining value of the G0/G1 epithelial cells (+). The results of such an analysis usually provided sensitivities between 50% and 70% for cervix tumor cells at a specificity of 90% when single data columns were evaluated. This was too low for the efficient detection of malignant samples.

Program DOCEVAL improved this result significantly by performing a multifactorial analysis of the five most discriminative data columns of the database. In a first step all values of each database column were standardized to the mean value of the normal samples of the same data column. In a second step the values of the five most discriminative

SAMPLE	CODE	RECORD IDENTIFICATION		1234	DIAGNOSIS
1:	305912.	>00 WA1002C5	<	++	ABNORMAL
2:	305913.	>00 WA1002C5	<	++	ABNORMAL
3:	306204.	>00 WA1004C5	<	++	ABNORMAL
4:	306323.	>00 WA1012C5	<	++	ABNORMAL
5:	306325.	>00 WA1013C5	<	+	ABNORMAL
6:	306334.	>00 WA1015C5	<	++	ABNORMAL
7:	307758.	>00 WA1079C5	V<	+	ABNORMAL
8:	316571.	>00 WA2202C5	V<	++	ABNORMAL
9:	316856.	>00 WA2247C5	<	++	ABNORMAL
10:	317399.	>00 WA2335C5	<		
11:	317825.	>00 WA2376C5	<	+	ABNORMAL
12:	317929.	>00 WA2395C5	V<	+	ABNORMAL
13:	320447.	>00 WA2602C5	<	+	ABNORMAL
14:	320450.	>00 WA2605C5 AT 2	<	++	ABNORMAL
15:	320724.	>00 WA2623C5	<	++	ABNORMAL
16:	323347.	>00 WA2737C5	<	+	ABNORMAL
17:	325123.	>00 WA2836C5	V<	++	ABNORMAL
18:	325130.	>00 WA2843C5	<		
19:	325136.	>00 WA2847C5	V<	++	ABNORMAL
20:	325139.	>00 WA2850C5	<	++	ABNORMAL
21:	325252.	>00 WA2952C5	V<	+	ABNORMAL
22:	325253.	>00 WA2853C5	V<	++	ABNORMAL
23:	325649.	>00 WA2862C5	V<	+	ABNORMAL
24:	325650.	>00 WA2963C5	V<	+	ABNORMAL
25:	325653.	>00 WA2866C5	V<	++	ABNORMAL
26:	325654.	>00 WA2867C5	V<	++	ABNORMAL
27:	325990.	>00 WA2879C5	V<	+	ABNORMAL
28:	328195.	>00 WA2925C5	V<	++	ABNORMAL
29:	328287.	>00 WA2928C5	V<	++	ABNORMAL
30:	328289.	>00 WA2929C5	V<	++	ABNORMAL
31:	328290.	>00 WA2930C5	V<	++	ABNORMAL
32:	328291.	>00 WA2928C5	V<	++	ABNORMAL
33:	317023.	>00 WA2286C5	<		

Fig.4 DIAGNOS1 output of 33 cervix cancer samples (Pap.IV, V). 31 samples were correctly recognized as abnormal at a false positive rate of 12% (tab.1).

Tab.1

FRESH CERVIX UTERI SMEARS

morphological classification (Papanicolaou)	number of samples	DIAGNOS1: % correct recognition
cancers (IV,V)	33	91
dysplasias (IIID)	9	44
normals (I,II)	830	88

data columns were multiplied or divided for each cell sample depending on whether the mean of the malignant samples was higher or lower than the mean of the normal samples. All possible permutations were calculated. The resulting new distributions of multifactors of the normal samples were thresholded by the SCREEN1 program as before. Samples were labelled abnormal when their multifactors were below or above the 10% or the 90% percentils. The multifactorial analysis joined the five most discriminative parameters of each cell sample to single multifactors. The best multifactor was in most instances significantly more discriminative than any one of the previous five parameters. Aneuploidy and the most discriminative three multifactors were used for the final diagnosis. A DOCEVAL printout for cervix cancer samples is shown in fig.4.

4. DISCUSSION

The DIAGNOS1 program system was developped to improve the recognition of abnormal cells from databased flow cytometric results. The system recognized more the 97% colo-rectal cancer cells when carcino-embryonic antigen, DNA and cell volume were simultaneously measured on formaldehyde fixed patient cells (6). Immunostains provide specific information, but are time consuming due to several incubation steps and cell washings.

An important goal of this study was to investigate whether quickly staining functional dyes such as the intracellular esterase and pH dye ADB and the DNA dye propidium iodide (1, 2) were suitable for automated cancer cell recognition. Most functional stains enrich sufficiently within 1 to 15min after addition to the samples and no cell washings are necessary. 91% of the malignant cervix samples and 88% of the normal samples were correctly recognized (tab.1). Similar results were obtained for human cancers of the colon, stomach and lung (7). One of the possibilities

to further improve these results, is to measure several different functional cell stains successively. Once the data columns with the most significant differences between normal and malignant samples in each database are known, only these data columns of each database are joined within a new database which is similarly analysed with the programs SCREEN1 and DOCEVAL as the individual databases.

An interesting possibility is to use the available cell biochemical information in the database for automated cell classification by flow-cytometry. Fibroblast classification in hereditary disease and granulocyte functions during severe infection or following polytrauma seem of particular practical interest (fig.1). Another promissing application could be the automated interpretation of the results of immuno-monitoring assays. The DIAGNOS1 program system is a highly integrated but nevertheless modular software tool. The total diagnostic analysis is done with a single command which is: DIAGVAG in case of cervix cells. New procedures on the other hand can be easily developped and scientific work is facilitated because all results are databased.

Most of the cell biochemical measurements by flow-cytometry can be calibrated in molecular terms. This permits the direct comparison of cells from different organs which is important for biochemical cell pathology. A further advantage of the biochemical investigation of complex cell systems by flow-cytometry is that prognostic information for the patient may be obtained (7) which is often not possible from morphological methods. It is conceivable that many functional abnormalities of diseased cells can be better detected by flow-cytometry than by either morphology or macroscopic biochemistry. Especially the development of further functional stains seems of importance to detect minor differences of cellular metabolism.

1. Valet, G. et al. (1981) Naturwiss. 68, 265-266
2. Valet, G. et al. (1984) in: "Intracellular pH-determination" (Ed. G.Valet), Paesel, Frankfurt, pp. 1-20
3. Kachel, V. et al. (1977) J.Histochem.Cytochem. 25, 804-812
4. Benker, G. et al. (1980) in: Flow Cytometry IV, Universitetsforlaget, Oslo, pp. 116-119
5. Valet, G. (1980) in: Flow Cytometry IV, Universitetsforlaget, Oslo, pp. 125-129
6. Valet, G. et al. (1984) J.Clin.Chem.Clin.Biochem. 22, 935-942
7. Valet, G. et al. (1986), in: "Endokrin aktive maligne Tumoren", (Ed. D.Engelhardt, K.Mann), Springer Verlag, Heidelberg, in press

Clinical Cytometry and Histometry

Edited by

G. Burger

Institut für Strahlenschutz, Neuherberg, FRG

J. S. Ploem

University of Leiden, The Netherlands

and

K. Goerttler

Deutches Krebsforschungszentrum, Heidelberg, FRG

1987

ACADEMIC PRESS

Harcourt Brace Jovanovich, Publishers

London San Diego New York Berkeley
Boston Sydney Tokyo Toronto