Journal of Cellular Biochemistry 17G:199–204 (1993)

Standardization and Quality Control of Quantitative Microscopy in Pathology

Robert L. Becker, Jr., MD, PhD, LtCol, USAF, MC

Armed Forces Institute of Pathology, Department of Cellular Pathology, Washington, DC 20306-6000

Abstract Standardization and quality control of quantitative microscopy techniques are distinct but related concerns. The first deals with the great variety of quantitative methods, measured features, and even response variables used in investigation of biological or clinical processes. The latter deals with reproducibility of results from those investigations across time and test performance sites. Though distinct, efforts for standardization and quality control are inherently interactive.

Consensus on standard methods, instrumentation, and data analysis is hard to achieve in fields developing as rapidly as quantitative microscopy. Consensus is possible, however, on the issues that affect test performance and interpretation. For example, issues of specimen type, fixation, processing, and staining affect image cytometry just as they do flow cytometry. Raw data acquisition issues include area sampling rules and fidelity of optical and sensor systems (light wavelength, glare/stray light, lens aberrations, numerical aperture, depth of focus, scan precision, pixel spacing and depth, sensor linearity, and stability). Intermediate data issues are primarily related to image foreground/background segmentation techniques—automated versus manual, object-specific versus field-based. Data reduction and interpretation procedures also provide many roads for divergence from uniformity. Each of these issues must be considered in terms of its effect on comparability and utility of quantitative microscopy results.

Quality control for quantitative microscopy is as important as standardization for its use in research programs and with clinical specimens. The *sine qua non* of quality control is comparison of experimental results against a known "correct" value to estimate accuracy, and against other experimental results to estimate precision. Intralaboratory quality control often uses internal standards, but can also use analysis of separate specimens with feature values known to a specified precision. Such separate specimens can also be used for interlaboratory, or "survey," quality control efforts. In any of these settings, limits must be established by which to declare a test in or out of control. The proper values of those limits depend on the accuracy and precision required for confident use of test results for a specific purpose.

Standardization and quality control are challenging requirements for effective multicenter use of cytometry or any other technology to establish surrogate endpoints of disease progression.

Key words: Computer-assisted image analysis, densitometry, morphometry, ploidy, quality control

Quantitative microscopy includes an array of techniques with the common element of reducing images to numeric features (Table I). Some powerful quantitative techniques (*e.g.*, mitotic figure counts and stereology) require only modest enhancement of a surgical pathologist's microscope. Others (*e.g.*, planimetry and ploidy determination

Address correspondence to Robert L. Becker, Jr., LtCol, USAF, MC, Department of Cellular Pathology, Armed Forces Institute of Pathology, Washington, DC 20306-6000.

by densitometry) often rely on cameras and sophisticated computer assistance for image acquisition, processing, and analysis. All are applied with the goal of better diagnosis and prognosis through more objective, precise, and accurate characterization of images. It is important to note that numerical features are not congruent with the visual criteria used by pathologists. Quantitative data can capture image features which are inapparent to the eye. On the other hand, quantitation provides no guarantee that all desired diagnostic or prognostic information is encoded in evaluated images.

Though many biomarkers (stage, proliferative activity, steroid receptors, ploidy, HER-2/*neu*, p53, and others) [1–8] predict the course of invasive breast carcinoma, only histologic and cytologic features have been extensively applied to risk from mammary preneoplasia and preinvasive neoplasia. Diagnostic criteria for precancerous lesions are well established [9,10], but there is *variation* in their application by pathologists, and even by individuals viewing slides more than once [10–12]. Hence, small differences between groups of patients are hard to detect unless many patients are studied. Alternatively, larger differences might develop and be detected if studies proceed for much longer time spans. Both solutions are expensive in the context of chemoprevention trials. An alternative is to require one or more pathologists to evaluate each case more than once. We would expect grading precision to increase according to the square of the number of observations. However, obtaining multiple observations is expensive, and does not correct inaccuracies resulting from bias in specimen evaluation. Suppose nuclear grade assigned by a pathologist is inappropriately biased by the most atypical cells in a preparation. As a countermeasure, we might present randomized images from many cases and require assignment of grade to each image, then average those values for a final grade. Such maneuvers, however, carry their own risk of distorting the grading result.

Quantitative microscopy is important to define biomarkers that can serve as surrogate endpoints in trials on the development of infiltrating breast carcinoma, but only to the extent that it captures outcome-related image features with accuracy, precision, and economy. Standardization and quality control are the means to provide those attributes. As with most clinical laboratory tests, quantitative microscopy techniques are well-suited to analysis as a series of steps (Table II), each contributing to overall error. We will consider examples of such error sources in association with one quantitative microscopy technique, DNA densitometry.

TABLE I. Quantitative Microscopy Methods

Mitotic Figure Counts
Densitometry
DNA
Immunohistochemistry
Stereology
Morphometry
Object-oriented
Field-oriented

TABLE II. Some Factors Affecting DNA Densitometry Results

Specimen preparation
Fixation and embedding
Section thickness
Stain technique
Data acquisition
Camera linearity
Shading and apertures
Image resolution
Sampling rules
Image segmentation and feature extraction
Manual
(Semi)automatic
Feature of interest
Presence of aneuploidy
DNA index
Histogram class
Proliferation rate

METHODS

Cell Preparation

Nuclei were recovered from 80 μm paraffin sections of colonic carcinoma using a modified Hedley technique [13]. Nuclei in suspension

were stained with propidium iodide (PI) by the Coulter Prep reagent. Nuclei dried onto a glass slide were stained by the pararosaniline Feulgen method [14].

Flow Cytometry

DNA determinations from PI-stained nuclei were made using a Coulter EPICS flow cytometer equipped with an Argon laser and Biosense flow cell [13]. The fluorescence signal path included the following filters: 488 nm laser blocker, 515 nm long pass, 590 nm dichroic, and 610 nm long pass. The instrument was calibrated against Coulter DNA Check fluorescent beads.

Image Cytometry

Determinations of integrated optical density (IOD) from Feulgen-stained nuclei were made using a custom built instrument. The microscope was an upright Olympus Vanox model fitted with a 550 nm bandpass filter, 40× (NA 0.95, theoretical optical resolution 0.35 μm) and 100× (NA 1.4, theoretical optical resolution 0.24 μm) planapochromatic objectives. Microscope apertures, lighting (via neutral density filters), and xyz stage position were set by stepping motors under computer control via IEEE 488 and RS2-32C interfaces. The sensor was a Videk Megaplus CCD camera (square pixels, 100% fill, 8 bit digital output, estimated 7¾ bits precision for two averaged 50 ms exposures at room temperature). Horizontal and vertical interpixel distance was 0.26 μm for the 40× objective and 0.10 μm for the 100× objective. Camera output was captured by Univision controller and display cards in an MS-DOS ISA computer running Bioscan Optimas software and custom macros. Optimas represents foreground and background segmentation of images as closed polygons rather than bitmaps. Images were stored on disk at full resolution (1189 × 1024 pixels) with lossless compression. Coefficient of variation (CV) for image histograms was calculated from peak width at half height [15]. Shading correction was a pixelwise interpolation of each specimen image (usually averaged from two exposures) between darkfield and brightfield images (each averaged from four exposures) according to the formula:

$$SPEC_{CORR}=\left(\frac{SPEC_{RAW}-DARK}{BRIGHT-DARK}\right)\times 255$$

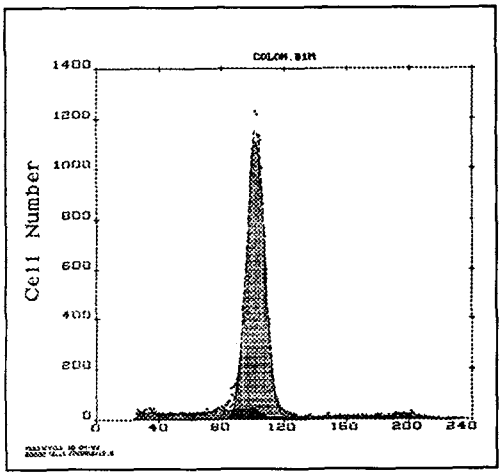

Fig. 1. Flow cytometry histogram of nuclei from colonic carcinoma. CV of the G_0/G_1 population was 5.6%. For comparison with image histograms of Figure 2.

RESULTS

The Videk sensor showed a very linear response against calibrated Wratten neutral density filters with densities from 0.1 OD to 1.0 OD. Pixels associated with the most heterochromatically stained nuclei (*i.e.*, lymphocytes in the colon preparation) rarely exceeded 0.6 OD. Hence, absorbance of the Feulgen preparation was well within the linear response range of the sensor. The camera signal was also quite reproducible, with a CV of 0.2% for IOD from repeated single exposures on one nucleus.

Results for varied image cytometry conditions were compared to those from baseline conditions that used optimal Koehler illumination, shading correction, full camera resolution, and manual separation of nuclei left conjoined by visually optimal pixelwise image segmentation according to gray value (Fig. 2a). The histogram for 800 tumor nuclei was consistent with diploidy, with a CV of 9.0%. Histogram quality from the same nuclei deteriorated substantially when shading correction was omitted (Fig. 2b), with a CV of 13.7%. Histogram quality was also sensitive to the setting of field and condensor apertures (Fig. 2c), with a CV of 11.7% for cells from the same area as before, but with both apertures

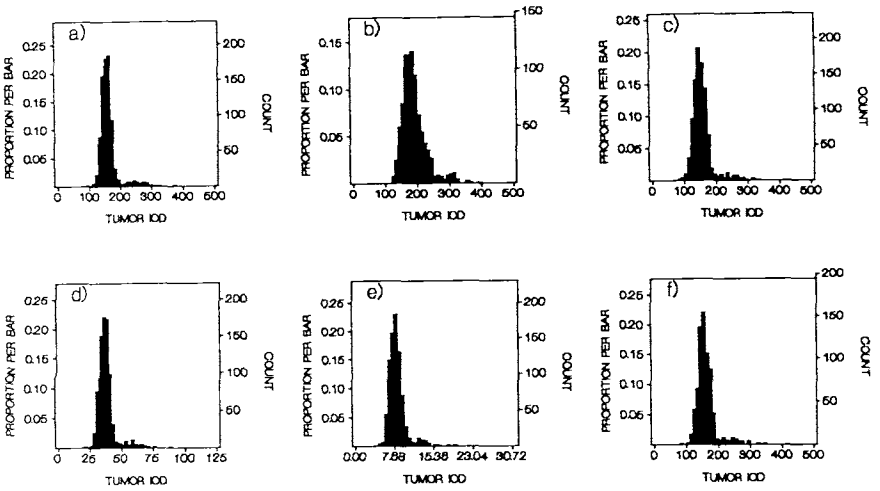

Fig. 2. Image cytometry histograms of nuclei from the same preparation as in Figure 1, but stained by the pararosaniline Feulgen method: (a) baseline histogram obtained with shading correction, Koehler illumination, full camera resolution, and manual splitting of abutting nuclei (CV 9.0%); (b) histogram from same cells as in panel (a) but without shading correction (CV 13.7%); (c) histogram from same slide area as in (a) but with field and condensor apertures fully open (CV 11.7%); (d) histogram from same cells as in (a) but with 2× resampling (reduction) of image and segmentation polygons (CV 10.0%); (e) as in (d) except with a 4× reduction from original image data (CV 12.7%); (f) as in (a) but excluding all nuclei that required manual splitting from neighbors (CV 8.7%).

left fully open. Flow cytometry of the colonic carcinoma (Fig. 1) yielded a histogram consistent with diploidy for 20,000 cells, with a CV of 5.6% for the G_0/G_1 population (CV was 0.84% for DNA Check beads and 1.88% for PI-stained chick erythrocytes).

The effect of varying image pixel spacing was tested by 2× and 4× reduction in image size via bilinear interpolation, matched by 2× and 4× shrinkage of the full-resolution segmentation polygons. The histogram CV increased to 10.0% and 12.7% with the 2× and 4× reductions, respectively (Fig. 2d, 2e). A histogram formed only from nuclei that were successfully segmented by gray value in baseline images (*i.e.*, without manual splitting) contained one-third fewer events, but had a CV of 8.7% (Fig. 2f).

There was substantial decrease in IOD for lymphocytes imaged two or more microns away from the z-position that yielded maximum IOD (Table III). A similar but less pronounced effect was found for tumor cells. For both cell types, IOD changed more per unit stage z-displacement above than below optimal focus position.

Visual setting of the gray value segmentation thresholds did not yield the best CV for repeat measurements from individual cells. In the example shown in Figure 3, setting a foreground window for gray values from 0 to 237 gave a visually satisfactory segmentation boundary, and repeat observations separated by xy displacement and manual refocusing yielded an IOD of 90.1 ± 5.4 (CV of 6.0%). Setting the foreground window for the same images at gray values 0–253 recruited many pixels surrounding the nucleus into the foreground. IOD was 97.5 ± 4.5 (CV of 4.6%).

DISCUSSION

Effects of fixation, processing, and staining on DNA densitometry results have been noted often [14,16]. Neutral buffered formalin is a generally satisfactory fixative, though it is not free of artifact [17]. Section thickness, so important in preparation of flow cytometry specimens from paraffin, is equally important for minimizing cut and overlapped nuclei in densitometry [14]. Frequently, such nuclei preclude obtaining high resolution DNA histograms from tissue, so touch preparations or nuclei from disaggregated tissue must be used instead. Even with the best preparations, histogram CV is usually higher with image cytometry than with flow cytometry, and IOD of

TABLE III. Effect of Focus on Integrated Optical Density of Feulgen-Stained Nuclei

Microns from best focus	−3.0	−2.0	−1.0	0.0	1.0	2.0	3.0
Percent of maximum IOD for lymphocytes	85%	91%	99%	100%	96%	86%	75%
Percent of maximum IOD for epithelial cells	93%	98%	98%	100%	96%	90%	86%

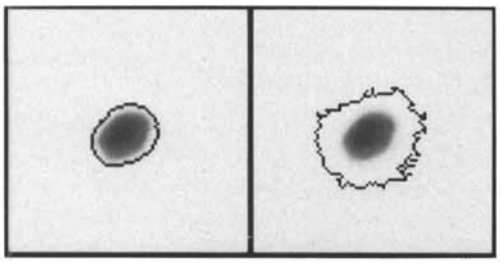

Fig. 3. Effect of varying foreground gray value segmentation window. Gray value lymphocyte image corrected for shading is the same in both panels. Segmentation window for left panel used gray values 0–237, window for right panel used gray values 0–253. Though the segmentation in left panel is visually more appropriate, segmentation on the right provided higher IOD and smaller CV in repeated measurements.

heterochromatic nuclei is often substantially lower than IOD of euchromatic nuclei with the same amount of DNA .

The data presented here demonstrate several sources of variation derived from instruments and algorithms requiring control in brightfield image cytometry. It is critically important to compensate for illumination inhomogeneity arising from the light source or dust in the light path; this is easy to do with today's image processing equipment. Adherence to principles of Koehler illumination significantly improves histogram quality. Higher resolution sensors also improve histogram CVs, if well-matched to the resolution of the microscope's optical elements. Digitized images capture the optical density of each nucleus over an image area greater than the corresponding size of the nucleus, in accordance with the microscope's point spread function. Nuclear boundaries determined by eye often exclude fringe pixels that, in aggregate, contain an important fraction of the IOD. This effect is even more serious if the nuclear image is not optimally focused.

Image analysis systems use various methods to minimize or compensate for image formation problems. Recent observations (A. Zetterberg, B. Palcic, J. Baak, personal communications) indicate that these problems can be overcome for DNA cytometry. Quality control targets from paraffin material may soon include data acquisition on at least one thousand nuclei per specimen with diploid G_0/G_1 CVs well under 5%, and no dependence of IOD on chromatin condensation. However, such targets can be adopted only if methods assuring high performance become widely known and practiced. Most of the issues affecting DNA densitometry apply to other image-based measurements as well. Features such as area and distance measurements [18], chromatin texture, and multispectral quantitation (unpublished data) are critically dependent on subtle aspects of image formation and analysis.

At this writing, there are no widely accepted standardization or quality control targets in place for image cytometry. In practice, such targets will develop as experience is gathered on clinical impact of the measurements. Targets for use of cytometry or any other measure of a surrogate endpoint biomarker for chemoprevention trials will be best constructed with knowledge of the minimum biomarker change that must be reliably detected for the purposes of the trial. Nonlabile proficiency testing specimens might be devised for some image-based tests. For example, stored digital images might be distributed by disk or network to allow comparison of image segmentation and feature extraction methods. However, some tests (including those with complex specimen preparation or labile product) will require distribution of unprocessed cells or tissue for adequate quality evaluation.

The opinions and assertions contained herein are the private views of the author and are not

to be construed as official or as representing the views of the Departments of the Army, Air Force, or Defense.

REFERENCES

1. Fisher ER, Redmond C, Fisher B, Bass G. (1990) Pathologic findings from the National Surgical Adjuvant Breast and Bowel Projects (NSABP). Prognostic discriminants for 8-year survival for node-negative invasive breast cancer patients. Cancer 65:2121–2128.
2. Dressler LG, Seamer L, Owens MA, Clark GM, McGuire WL. (1987) Evaluation of a modeling system for S-phase estimation in breast cancer by flow cytometry. Cancer Res 47:5294–5302.
3. Dressler LG, Eudey L, Gray R, Tormey DC, McGuire WL, Gilchrist KW, Clark GM, Osborne CK, Mansour EG, Abeloff MD. (1992) Prognostic potential of DNA flow cytometry measurements in node-negative breast cancer patients: Preliminary analysis of an intergroup study (INT 0076). Monogr Natl Cancer Inst 11:167–172.
4. Abandowitz HM, Ow KT, Hardy D, Keightley DD, Sarfaty GA, Nash AR. (1987) Relationship between flow cytometric parameters, steroid receptors, and menopausal status in breast cancers. Oncology 44: 24–29.
5. Auer G, Caspersson TO, Wallgren AS. (1980) DNA content and survival in mammary carcinoma. Analyt Quant Cytol 2:161–165.
6. Auer GU, Askensten U, Erhardt K, Fallenius A, Zetterberg A. (1987) Comparison between slide and flow cytophotometric DNA measurements in breast tumors. Anal Quant Cytol Histol 9:138–146.
7. Slamon DJ, Godolphin W, Jones LA, Holt JA, Wong SG, Keith DE, Levin WJ, Stuart SG, Udove J, Ullrich A. (1989) Studies of the HER-2/*neu* proto-oncogene in human breast and ovarian cancer. Science 244: 707–712.
8. Allred DC, Clark GM, Elledge R, Fuqua SA, Brown RW, Chamness GC, Osborne CK, McGuire WL. (1993) Association of p53 protein expression with tumor cell proliferation rate and clinical outcome in node-negative breast cancer. J Natl Cancer Inst 85: 200–206.
9. Dupont WD, Parl FF, Hartmann WH, Brinton LA, Winfield AC, Worrell JA, Schuyler PA, Plummer WD. (1993) Breast cancer risk associated with proliferative breast disease and atypical hyperplasia. Cancer 71:1258–1265.
10. Connolly JL, Schnitt SJ. (1993) Benign breast disease. Cancer 71:1187–1189.
11. Schnitt SJ, Connolly JL, Tavassoli FA, Fechner RE, Kempson RL, Gelman R, Page DL. (1992) Interobserver reproducibility in the diagnosis of ductal proliferative lesions using standardized criteria. Am J Surg Pathol 16:1133–1143.
12. Rosai J. (1991) Borderline epithelial lesions of the breast. Am J Surg Pathol 15:209–221.
13. Becker RL, Venzon D, Lack EE, Mikel UV, Weiss SW, O'Leary TJ. (1991) Cytometry and morphometry of malignant fibrous histiocytoma of the extremities: Prediction of metastasis and mortality. Am J Surg Pathol 15:957–964.
14. Mikel UV, Fishbein WN, Bahr GF. (1985) Some practical considerations in quantitative absorbance microspectrophotometry. Preparation techniques in DNA cytophotometry. Analyt Quant Cytol Histol 7:107–118.
15. Shapiro HM. (1935) "Practical Flow Cytometry." New York, NY: Alan R. Liss, Inc., p 202.
16. Mikel UV, Becker RL. (1991) A comparative study of quantitative stains for DNA in image cytometry. Analyt Quant Cytol Histol 13:253–260.
17. Becker RL, Mikel UV. (1990) Interrelation of formalin fixation, chromatin compactness and DNA values as measured by flow and image cytometry. Analyt Quant Cytol Histol 12:333–341.
18. Young IT. (1988) Sampling density and quantitative microscopy. Analyt Quant Cytol Histol 10:269–275.