

# Identifying and Quantifying Apoptosis: Navigating Technical Pitfalls

Megan M. Garrity, B.S., Lawrence J. Burgart, M.D., Darren L. Riehle, B.A., C.T., Eunice M. Hill, B.S., C.T., Thomas J. Sebo, M.D., Ph.D., Thomas Witzig, M.D.

*Department of Laboratory Medicine and Pathology (MMG, LJB, DLR, EMH, TJS) and Department of Internal Medicine (TW), Mayo Clinic, Rochester, Minnesota*

**Apoptosis or programmed cell death is often altered in malignancies and is frequently determined by the terminal transferase-mediated nick end labeling technique (TUNEL). However, commercially available protocols can produce high background and false-positive staining, which renders the distinction between apoptosis and necrosis difficult. In an attempt to develop a rapid and reproducible method for detecting and quantifying apoptosis, we coupled optimization of the Apoptag Plus Peroxidase *In Situ* Apoptosis Detection kit with quantitative histomorphometric computer imaging software using the Bacus Laboratories Incorporated Slide Scanner (BLISS). Multiple (200–350) unique 40× images were scanned using the BLISS system and downloaded into the WebSlide Browser program, creating a permanent, scanned record of the area assessed. The stored images were counted, with the final analysis simultaneously taking into account cells that were immunohistochemically positive and the histology of the surrounding cells to reduce the possibility of false positive and negative staining. In addition, cells with equivocal staining can be simultaneously reviewed by other technologists with networked WebSlide Browser access to the same images. Our data show that the advantages offered by the BLISS imaging software greatly reduce the potential drawbacks of using the TUNEL method as a sole means of quantification.**

**KEY WORDS:** Apoptosis, Digital image analysis, Image morphometry, Necrosis, TUNEL assay.

*Mod Pathol* 2003;16(4):389–394

Apoptosis is a physiologic form of cell death that serves as one aspect of tissue growth and size regulation (1). The inhibition of apoptosis is believed to play a role in carcinogenesis in one of two ways: (1) it may allow for the unchecked accumulation of genetic alterations (2, 3) and (2) it may lead to unbalanced proliferation of tumor (4). As such, it has become an area of intense interest in oncologic research.

The morphologic criteria for identifying cells undergoing apoptosis are well established and include cytoplasmic condensation, loss of cell-cell contact, and cell shrinkage. This is separate and distinct from necrosis, an alternate form of cell death that involves cell swelling and rupture with associated surrounding tissue damage (5–7). Although cells undergoing apoptosis can be identified using these morphologic criteria on a standard hematoxylin and eosin (H&E) slide, studies have shown that using this method alone may underestimate the rate of apoptosis by 2-fold to 3-fold (8). As a result, there is increased interest in finding a reliable, reproducible method for identifying and quantifying apoptosis in whole tissue sections.

Apoptosis is most commonly identified in tissue using terminal deoxynucleotidyl transferase-mediated dUTP nick end labeling (TUNEL). TUNEL works on the principle that DNA strand breaks occur during apoptosis. Terminal deoxynucleotidyl transferase (Tdt) catalyzes the labeling of these breaks with deoxyuridine triphosphates (dUTPs), which are then detected by immunoperoxidase techniques (8–10). Several kits are commercially available for TUNEL detection. However, reports have indicated incidences of high background and false-positive staining, making interpretation difficult (6, 8, 9, 11).

This high rate of false-positive staining can potentially arise from several factors. First, initial fix-

Copyright © 2003 by The United States and Canadian Academy of Pathology, Inc.

VOL. 16, NO. 4, P. 389, 2003 Printed in the U.S.A.

Date of acceptance: January 28, 2003.

Supported in part by a grant from the National Cancer Institute (CA78899).

Address reprint requests to: Megan Garrity, B.S., Mayo Clinic, Department of Laboratory Medicine and Pathology, 200 First Street SW, Rochester, MN 55905; e-mail: garrity.megan@mayo.edu.

DOI: 10.1097/01.MP.0000062657.30170.92

ation of tissue sections that is too extensive, incomplete, or delayed can lead to nonspecific staining. Second, the need for unmasking after formalin fixation and paraffin embedding, i.e., through protease addition, can produce artificial strand breaks that are unable to be immunohistochemically differentiated from true apoptosis. Finally, necrosis also results in DNA strand breaks and as such represents a target for Tdt-mediated dUTP labeling (1, 8, 11, 12). Although compensating for deficits in initial tissue fixation in retrospective studies is problematic, addressing the issue of overall optimization remains crucial if the role of apoptosis in carcinogenesis is to be delineated. In addition, many investigators have suggested that proper interpretation of TUNEL results requires a second confirmatory methodology (12–14).

Initially we compared two commercially available TUNEL kits to determine the ease of optimization and the reproducibility across serial tissue sections. We then investigated two digital image morphometry systems to see whether either could yield reproducible, reliable values for the rate of apoptosis while also incorporating multiple methodologies and technical checks for quantifying apoptosis. The two methodologies evaluated were (1) morphometric area analysis (MAA), in which the apoptotic index is reported as a percentage of area stained positive *versus* total area stained using the Cell Analysis System (CAS 200, Bacus Laboratories, Inc., Lombard, IL) and (2) morphometric cell count (MCC) analysis, in which the apoptotic index is reported as the total number of apoptotic cells *versus* the total number of cancer cells using the Bacus Laboratories Incorporated Slide Scanner (BLISS, Bacus Laboratories, Inc.). We evaluated both of these methods for quantifying apoptosis as compared with semiquantitative visual scoring of 1, 2, or 3+ to determine how precisely each technique can discriminate between rates of apoptosis.

## MATERIALS AND METHODS

### Tissue Samples

Formalin-fixed, paraffin-embedded tissue blocks were obtained and serial 4- $\mu\text{m}$  sections were processed for use in the TUNEL assays. Both tonsil and colorectal adenocarcinoma tissues were used for initial TUNEL optimization. Colorectal adenocarcinoma tissue was used for subsequent development of the quantification assays. The use of all samples was approved by the institutional review board of the Mayo Clinic.

### TUNEL Optimization

Serial sections of tonsil tissue were cut from several formalin-fixed, paraffin-embedded blocks.

Multiple sections were stained using the Apoptag Plus Peroxidase *In Situ* Apoptosis Detection Kit (Intergen Company, Purchase, NY) or the *In Situ* Cell Death Detection Kit, AP (Roche Diagnostics Corp, Indianapolis, IN). Studies were done to evaluate the effects of proteinase K, Tdt concentration, secondary substrate (anti-digoxigenin-peroxidase or anti-fluorescein antibody AP conjugated) incubation and counterstaining on overall reproducibility, sensitivity, and interpretation of stain. For each kit, proteinase K concentration was varied from 5–30  $\mu\text{g}/\text{mL}$  (5, 10, 15, 20, 25, and 30  $\mu\text{g}/\text{mL}$ ) in phosphate buffered saline (PBS) and incubated from 5–20 minutes (5, 10, 15, and 20 minutes) at each concentration. After optimizing the time and concentration for proteinase K (25  $\mu\text{g}/\text{mL}$  for 20 minutes), samples were incubated at both room temperature and 37° C to determine any other substantive differences. Tdt concentration was varied from the recommended dilution to half of that. Anti-digoxigenin-peroxidase or anti-fluorescein antibody AP secondary substrate (Apoptag Plus Peroxidase *In Situ* Apoptosis Detection Kit and *In Situ* Cell Death Detection Kit, AP respectively) was incubated at both room temperature and 37° C. Finally, because digital imaging systems delineate positive and negative cells using the difference in wavelength between chromogen and counterstain, we tested the effects of diluted hematoxylin, fast red, and methyl green counterstaining on overall readability with the CAS 200 and BLISS.

To test for reproducibility and sensitivity under optimal staining conditions, 10 serial sections of tonsil tissue were cut from each of two blocks. Five of each were then stained using both the Apoptag Plus Peroxidase *In Situ* Apoptosis Detection kit (Apoptag) or the *In Situ* Cell Death Detection kit, AP (ISCD-AP). Similar studies were done on colorectal adenocarcinoma sections to confirm reproducibility and sensitivity before development of quantification techniques.

### TUNEL Assay for Development of Quantification Assay

Thirty-five colorectal adenocarcinomas were stained in duplicate using the Apoptag kit previously optimized for formalin-fixed, paraffin-embedded whole tissue sections. Formalin-fixed, paraffin-embedded tissue slides were rehydrated using xylene to alcohol washings, followed by a hydrogen peroxide-methanol quench. The samples were treated with 25  $\mu\text{g}/\text{mL}$  of proteinase K at 37° C for 20 minutes. After washing and incubation with equilibration buffer for 5 minutes, Tdt was diluted 1:3.9 (14  $\mu\text{L}$  of Tdt enzyme in 40  $\mu\text{L}$  of reaction buffer) and incubated on the slides for 1 hour at 37° C. After applying stop solution for 15 minutes and

washing, the samples were incubated with anti-digoxigenin peroxidase conjugate at 37° C for 30 minutes. Slides were developed with a 1:20 dilution of diaminobenzidine (3,3'-diaminobenzidine) substrate, counterstained with methyl green, dehydrated, and coverslipped.

### Morphometric Area Analysis

A board-certified cytotechnologist who had been specially trained to perform digital image analysis evaluated areas representing both normal and neoplastic tissue that were marked by a pathologist. The CAS 200 image analyzer was used to determine the amount of cellular area stained with 3,3'-diaminobenzidine (positive for apoptosis) at a wavelength of 620 nm and the amount of cellular area counterstained with methyl green or hematoxylin (negative for apoptosis) at a wavelength of 500 nm within the same selected area. Values were quantified using the Quantitative Proliferation Index software (Bacus Laboratories, Inc.) Briefly, the apoptotic rate was calculated as the quotient of 3,3'-diaminobenzidine-positive area over the total area (3,3'-diaminobenzidine positive + methyl green negative), times 100 (15).

### Morphometric Cell Count

Between 200 and 350 40× histologic images, pre-selected by a pathologist, were captured using the BLISS system. These images were viewed with the WebSlide Browser (Bacus Laboratories, Inc.) program, and the digital images were automatically stored for later retrieval on a computer screen. The BLISS device scans and digitizes these bright-field images as a series of nonoverlapping tiles. Consequently, with this device, there is no risk of accidentally skipping over areas of tumor or, conversely, of counting fields more than once. Apoptotic bodies (3,3'-diaminobenzidine positive) were counted across all of the captured 40× images. The total number of intact carcinoma cells was established by counting these cells (counterstained in methyl green) in a random collection of 10% of the captured 40× images. This was then extrapolated across the entire area scanned. The apoptotic index was calculated by dividing the total number of apoptotic bodies by the total number of intact carcinoma cells and multiplying by 100.

## RESULTS

### Assay Optimization

Each variable (proteinase K, antibody incubation, and counterstaining) was analyzed for both tonsil and colorectal adenocarcinoma whole-tissue sections. Staining indicated that optimization of the concentration, temperature, and duration of proteinase K incubation was most crucial for both kits to achieve high sensitivity. In addition, decreasing the concentration of the Tdt enzyme and increasing the incubation temperature of the anti-digoxigenin in the Apoptag kit greatly enhanced its overall sensitivity and specificity.

Despite optimization attempts with the ISCDD-AP kit, overall reproducibility and sensitivity remained low, as shown in Table 1. The Apoptag kit had good reproducibility, as shown by the low standard deviations. The Apoptag kit did show an affinity for staining some intensely necrotic areas and attempts to optimize this greatly reduced sensitivity. However, this pattern of staining was easily discernible from truly apoptotic staining as there was no “bleeding” of the chromogen into these areas (Fig. 1A), as seen with the ISCDD-AP kit (Fig. 1B). In addition, the sensitivity toward fixation, as seen by patchy areas of intense, nonspecific staining, seemed higher in the ISCDD-AP kit, making retrospective studies more challenging.

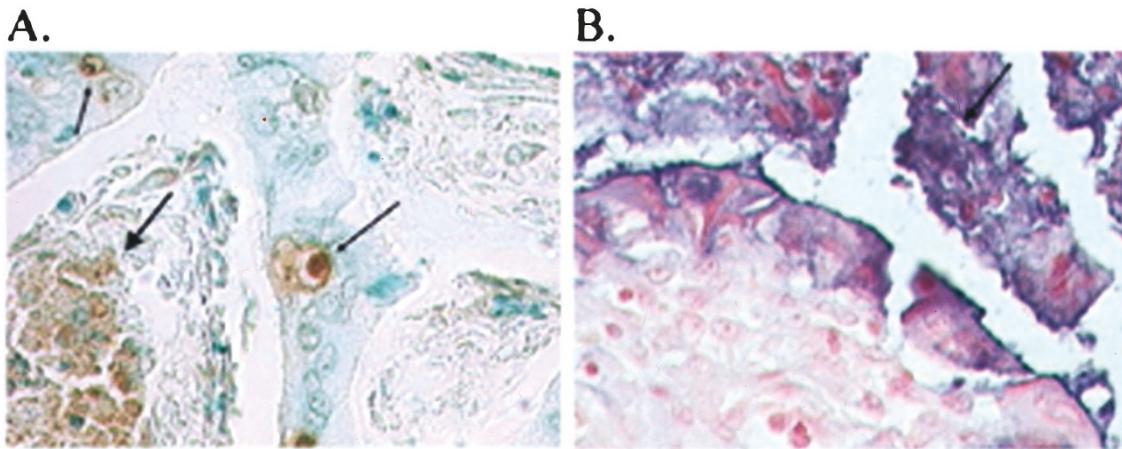
The reliance of the CAS 200 on wavelengths favored the use of methyl green as a counterstain. The reading of stains produced by the ISCDD-AP kit was not possible with the CAS 200 because the difference between the wavelengths for positive and negative cells was minimal, making thresholding between the two cell populations very difficult. Similarly, given the rare event nature of apoptosis and the range of 3,3'-diaminobenzidine intensities as a cell progresses through apoptosis, diluted hematoxylin was difficult to consistently interpret with both the BLISS and CAS 200. Because the intent was to compare the use of MCC with that of MAA, methyl green was used as the counterstain.

### Quantification Methodology

As a measure of reproducibility, a comparison of the standard deviation between the apoptotic index of colorectal adenocarcinomas obtained using both the MAA and the MCC analysis showed that over

**TABLE 1. Reproducibility and Sensitivity**

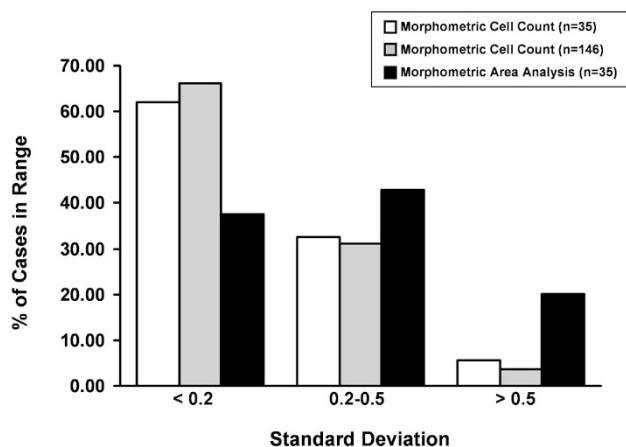
Area	Apoptag					Avg ± SD	ISCDD-AP					Avg ± SD
	1	2	3	4	5		1	2	3	4	5	
1A	0.72	0.74	0.72	0.71	0.72	0.722 ± 0.011	0.4	0.25	0.16	0.46	0.41	0.336 ± 0.126
1B	0.52	0.57	0.58	0.51	0.53	0.542 ± 0.031	0.15	0.36	0.17	0.52	0.23	0.286 ± 0.154
2A	0.32	0.37	0.4	0.38	0.45	0.384 ± 0.047	0.48	0.28	0.17	0.5	0.27	0.340 ± 0.144
2B	0.43	0.5	0.51	0.45	0.47	0.472 ± 0.033	0.32	0.23	0.15	0.39	0.42	0.302 ± 0.112



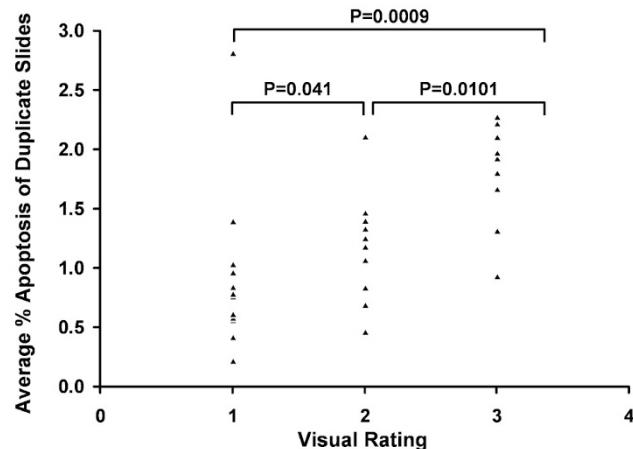
**FIGURE 1.** Necrosis *versus* apoptosis. Staining of intensely necrotic areas (*heavy arrow*) using the Apoptag kit was distinct from that of true apoptotic cells (*thin arrows*; A). In contrast, there was generalized, nonspecific staining with the ISCDD-AP kit (B).

60% of values obtained using the MCC method fell within 0.2 standard deviations across duplicate samples. This compares with less than 40% using the MAA method. Less than 10% of values obtained using the MCC method fell outside of the 0.5 standard deviation range *versus* 20% using the MAA method. The average variation between duplicate samples was 0.189 using the MCC method, *versus* 0.585 using the MAA method. This low standard deviation using the MCC method remained the same through subsequent samples ( $n = 146$ ) (Fig. 2).

In addition, a comparison of the average percentage apoptosis found using the MCC method to an overall visual scoring of the case yielded a correlative trend (Fig. 3). The MAA method failed to yield a similar positive correlation (Fig. 4). The only statistically significant difference using the MAA method occurred between the 1 and 3 categories.



**FIGURE 2.** Digital imaging (MCC/MAA) reproducibility. A comparison of the standard deviation of the values obtained across duplicate samples showed that MCC ( $n = 35$ ) has better overall reproducibility than MAA ( $n = 35$ ). This low standard deviation was maintained through subsequent sample analysis ( $n = 146$ ).



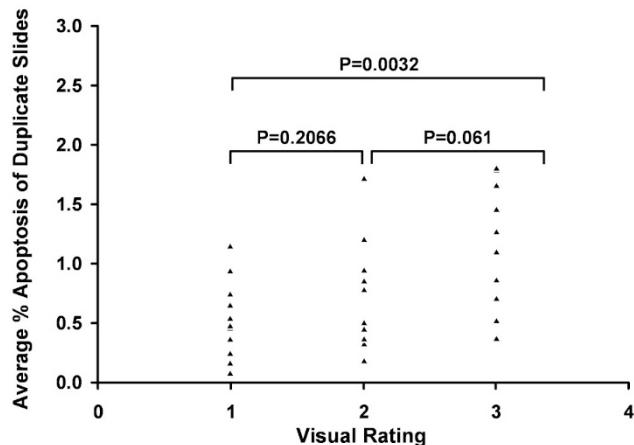
**FIGURE 3.** MCC *versus* visual score. Values obtained using the MCC method showed a correlative trend with a semiquantitative visual score. Using the Wilcoxon rank sum test, a statistically significant difference ( $P < .05$ ) was seen between all visual scores.

The MCC method, however, provided a statistically significant difference across each visual rating (1 to 2, 2 to 3, and 1 to 3), indicating a stronger ability to delineate percentage apoptosis.

## DISCUSSION

Cell death has become an area of increasing interest in oncologic research. There are two known types of cell death, apoptosis and necrosis. Apoptosis is considered to be a highly conserved process that occurs in response to internal physiologic disturbances and allows for a more controlled single cell deletion of aberrant cells. In contrast, necrosis occurs in response to external stimuli and involves a corresponding inflammatory response that leads to the destruction of tissue adjacent to the necrotic cell (5–7, 16).

Disruption of apoptosis has been implicated in the unchecked accumulation of genetic alterations



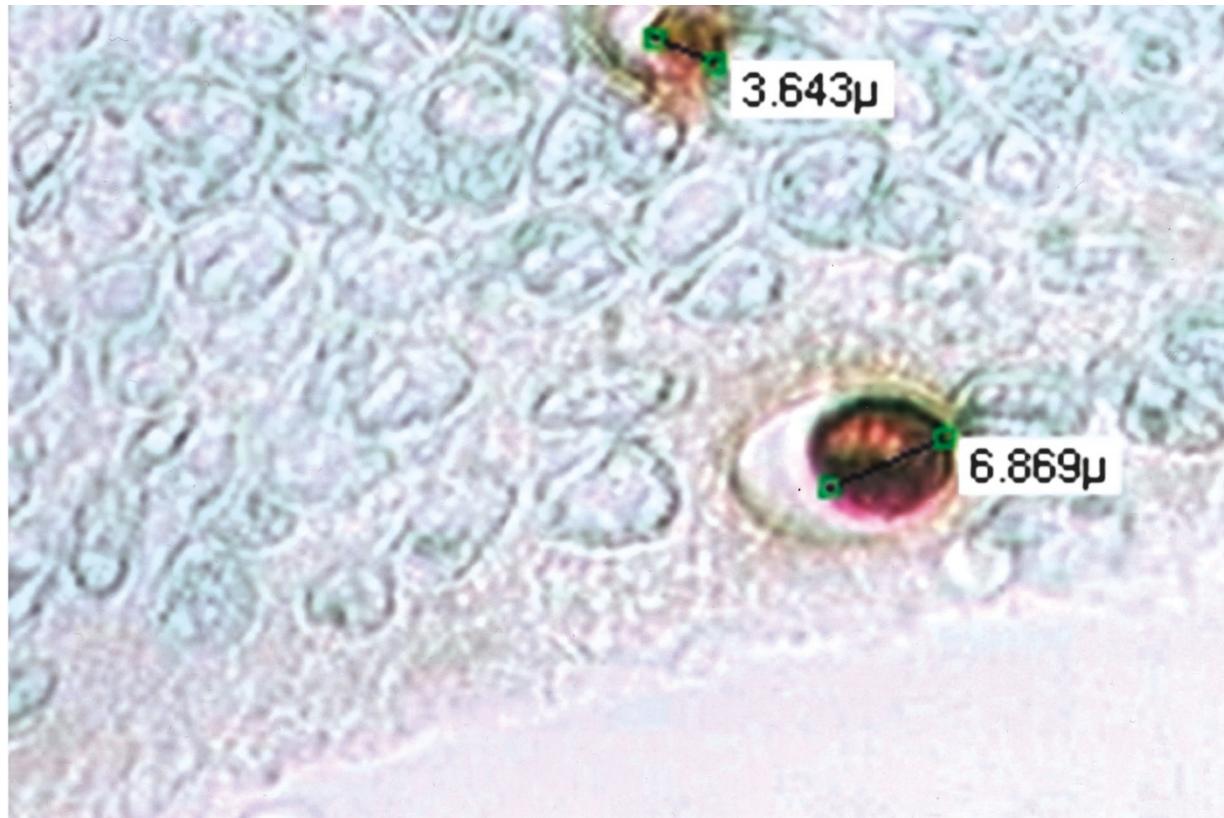
**FIGURE 4.** MAA versus visual score. Values obtained using the MAA method did not correlate with the semiquantitative visual score. Using the Wilcoxon rank sum test, the only statistically significant difference ( $P < .05$ ) occurred between the visual score of 1+ to 3+.

or proliferation of tumor cells, thus leading to carcinogenesis (16, 17). Because it may follow an identifiable pathway, modulation of apoptosis has potential applications to chemotherapy and chemoprevention for various cancers (5). However, quantification and subsequent evaluation of this process remains problematic (6, 8, 9, 11). Developing a reliable, reproducible means for

identifying and quantifying apoptotic cells remains a crucial step in determining the overall significance of this process in carcinogenesis.

Two methods frequently used for quantifying apoptosis are counting morphologically positive cells on an H&E stain or evaluating the number of TUNEL-positive cells per 1000–2000 cells (4, 9, 16). Both methods appear to have drawbacks. Studies have shown that without the use of immunohistochemical markers to identify apoptotic cells, that is, using standard H&E stains, apoptosis rates can be underestimated (8, 16). Similarly, by only evaluating 1000–2000 cells to determine the apoptotic rate, the total number of  $40\times$  areas considered may be low and may not accurately reflect the amount of apoptotic cells in the overall tissue section.

In this study, we evaluated two currently available means of quantification for reproducibility and overall correlation to visual scoring. Although the use of the CAS 200 for image analysis on other immunohistochemical markers such as MIB-1 is well documented (18, 19), we found that it was less useful in determining apoptotic rates in whole tissue sections. In contrast, MCC analysis appears to be an effective means for assessing percentage apoptosis for several reasons. First, the MCC method showed both reproducibility across duplicate stains



**FIGURE 5.** Nuclear size change. The process of apoptosis changes the nuclear size of the cell. Therefore, depending on the stage of apoptosis, the area that MAA assigns a cell varies (in this case by approximately one half), altering the calculated apoptotic rate. In contrast, MCC would count this as two events.

and a positive correlation to an overall qualitative score, whereas the MAA did not. This may be due to the reliance of MAA on the measurement of nuclear area, the alteration of which is one hallmark of apoptosis. Conversely, MCC uses a cell count, an actual measure of event occurrence, which is unaffected by such changes (Fig. 5).

Second, the TUNEL assay used to identify apoptotic cells for quantification has been subject to some debate. Our research confirms that optimization is an essential step toward improving the reliability of TUNEL. Our laboratory found that the greatest effects on overall optimization came from the duration, concentration, and temperature of the proteinase K unmasking step. In addition, the adjustment of the concentration of Tdt enzyme and temperature for the anti-digoxigenin peroxidase in the Apoptag kit greatly decreased background and increased differential staining between necrotic areas and truly apoptotic areas. We were unable to achieve such results with the ISCDD-AP kit.

Even with optimization, TUNEL may not detect early stages of apoptosis in which morphological changes are still evident or may not adequately differentiate necrotic cells. Because of these limitations in identifying and quantifying apoptosis, many have suggested that an alternate means of evaluation be used in conjunction with TUNEL (6, 9, 12–14). We have shown that the coupling of MCC with TUNEL allows for such dual evaluation. In addition to being able to quickly read positive, brown-stained cells off the scanned 40× images, we were also able to evaluate the remaining cells for morphologic changes indicative of apoptosis. MAA, which relies solely on the color of the cell for designation of whether that cell is apoptotic, does not allow for such complementary assessment.

Finally, 200 to 350 unique and distinct 40× images are scanned in and evaluated with MCC. This allows for the assessment of a larger area of tissue and therefore may yield a more representative apoptotic rate for the entire tissue section. It also allows for the simultaneous assessment of questionable areas through networked access. In addition, MCC creates a permanent scanned electronic record of the area evaluated. This allows for easy reference back to cases and evaluation of identical areas by multiple individuals.

TUNEL optimization of multiple variables is time consuming and appears to be necessary for each new study. For instance, we have found that reoptimization is essential when dealing with small biopsy specimens *versus* whole tissue sections. However, once optimal conditions have been set for a given study, these do not appear to change, and sections can be quickly stained. To ensure maximum accuracy, we also recommend doing samples in duplicate if possible. MCC may be challenging at

times, but it does seem to provide an advantage over traditional quantification methods. Therefore, we feel that MCC analysis coupled with optimized TUNEL provides a strong, reproducible means of evaluating percentage apoptosis.

## REFERENCES

1. Cummings MC, Winterford CM, Walker NI. Apoptosis. Am J Surg Pathol 1997;21:88–101.
2. Sinicrope FA, Ruan SB, Cleary KR, Stephens LC, Lee JJ, Levin B. bcl-2 and p53 oncogene expression during colorectal tumorigenesis. Cancer Res 1995;55:237–41.
3. Bedi A, Pasricha PJ, Akhtar AJ, Barber JP, Bedi GC, Giardiello FM, et al. Inhibition of apoptosis during development of colorectal cancer. Cancer Res 1995;55:1811–6.
4. Baretton GB, Diebold J, Christoforis G, Vogt M, Muller C, Dopfer K, et al. Apoptosis and immunohistochemical bcl-2 expression in colorectal adenomas and carcinomas. Cancer 1996;77:255–64.
5. Kerr JFR, Winterford CM, Harmon BV. Apoptosis: its significance in cancer and cancer therapy. Cancer 1994;73:2013–26.
6. Willingham MC. Cytochemical methods for the detection of apoptosis. J Histochem Cytochem 1999;47:1101–10.
7. Leist M, Nicotera P. The shape of cell death. Biochem Biophys Res Commun 1997;236:1–9.
8. Jerome KR, Vallan C, Jaggi R. The TUNEL assay in the diagnosis of graft-versus-host disease: caveats for interpretation. Pathology 2000;32:186–90.
9. Cuello-Carrion FD, Ciocca D. Improved detection of apoptotic cells using a modified *in situ* TUNEL technique. J Histochem Cytochem 1999;47:837–9.
10. Alison M. Identifying and quantifying apoptosis: a growth industry in the face of death. J Pathol 1999;188:117–8.
11. Labat-Moleur F, Guillermet C, Lorimier P, Robert C, Lantuejoul S, Brambilla E, et al. TUNEL apoptotic cell detection in tissue sections: critical evaluation and improvement. J Histochem Cytochem 1998;46:327–34.
12. Grasl-Kraupp B, Ruttkay-Nedecky B, Koudelka H, Bukowska K, Bursch W, Schulte-Hermann R. *In situ* detection of fragmented DNA (TUNEL assay) fails to discriminate among apoptosis, necrosis, and autolytic cell death: a cautionary note. Hepatology 1995;21:1465–8.
13. Darzynkiewicz Z, Juan G, Li X, Wojciech G, Murakami T, Traganos F. Cytometry in cell necrobiology: analysis of apoptosis and accidental cell death (necrosis). Cytometry 1997;27:1–20.
14. Collins RJ, Harmon BV, Gobe GC, Kerr JFR. Internucleosomal DNA cleavage should not be the sole criterion for identifying apoptosis. Int J Radiat Biol 1992;61:451–3.
15. Sebo T. Laboratory medicine and pathology: digital image analysis. Mayo Clin Proc 1995;70:81–2.
16. Koike M. Significance of spontaneous apoptosis during colorectal tumorigenesis. J Surg Oncol 1996;62:97–108.
17. Kasagi N, Gomyo Y, Shirai H, Tsujitani S, Ito H. Apoptotic cell death in human gastric carcinoma: analysis by terminal deoxynucleotidyl transferase-mediated dUTP-biotin nick end labeling. Jpn J Cancer Res 1994;85:939–45.
18. Sebo TJ, Cheville JC, Riehle DL, Lohse CM, Pankratz VS, Myers RP, et al. Predicting prostate carcinoma volume and stage at radical prostatectomy by assessing needle biopsy specimens for percent surface area and cores positive for carcinoma, perineural invasion, Gleason score, DNA ploidy and proliferation, and preoperative serum prostate specific antigen: a report of 454 cases. Cancer 2001;91:2196–204.
19. Cheng L, Pisansky TM, Sebo TJ, Leibovich BC, Ramnani DM, Weaver AL, et al. Cell proliferation in prostate cancer patients with lymph node metastasis: a marker for progression. Clin Cancer Res 1999;5:2820–3.