Methodological Findings and Considerations in Measuring Colorectal Epithelial Cell Proliferation in Humans¹

Roberd M. Bostick,² Lisa Fosdick, Tamera J. Lillemoe, Paula Overn, James R. Wood, Patricia Grambsch, Patricia Elmer, and John D. Potter

Department of Public Health Sciences, Epidemiology, The Bowman Gray School of Medicine, Wake Forest University, Winston-Salem, North Carolina 27157 [R. M. B.]; Divisions of Biostatistics [L. F., P. G.] and Epidemiology [P. E.], School of Public Health, University of Minnesota, Minneapolis, Minnesota 55414; Abbott Northwestern Hospital, Minneapolis, Minnesota 55407 [T. J. L., P. O.]; Digestive Healthcare, PA, Minneapolis, Minnesota 55404 [J. R. W.]; and Fred Hutchinson Cancer Research Center, Seattle, Washington 98104 [J. D. P.]

Abstract

The methodological issues for measuring colorectal epithelial cell proliferation, an intermediate endpoint for studies of colon neoplasia, in epidemiological studies are deceptively numerous and complex, with few methodological data available. Accordingly, during our experience with measuring colorectal epithelial cell proliferation from nearly 500 participants attending over 1300 study visits over a 6-year period, we recorded data on a variety of measurement variations. Methods investigated included rectal biopsy technique, general histological and labeling procedures (including the tritiated thymidine, 5-bromodeoxyuridine (BrdUrd), and the proliferating cell nuclear antigen (PCNA) immunohistochemical techniques used to label S-phase cells in colonic crypts in rectal biopsy specimens), biopsy scoring procedures, and summary scoring methods. Findings include that the PCNA technique was the simplest, most economical, and least time-consuming. The BrdUrd labeling failure rate was 15% versus <1% for PCNA. The percentage of labeled cells (labeling index) was highest using PCNA in biopsies processed without prior incubation, intermediate using PCNA in biopsies processed with prior incubation as for BrdUrd, and lowest using BrdUrd. The percentage of labeled cells that were in the upper 40% of the crypt (Φu) was higher using BrdUrd than PCNA; visit-to-visit correlations were higher using PCNA (r = 0.51 versus 0.35), and visit-to-visit variability was lower and between-person variability was higher using PCNA. Intra- and inter-rater reliabilities for the techniques were comparable (PCNA intra-rater r = 0.93, inter-rater r = 0.92). The PCNA technique, compared to the BrdUrd technique, is more feasible and reliable, provides a more accurate estimate of the labeling index, and cell proliferation measures determined with PCNA have statistical properties that are generally more favorable for detecting differences in clinical trials. Thus, the PCNA technique may be preferable to techniques requiring incubation of biopsies. Other methodological findings lead us to recommend that, for larger studies measuring colorectal epithelial cell proliferation on outpatient rectal biopsies, biopsies should be taken 10 cm above the anus using a flexible, preferably jumbo cup, endoscopic forceps through a rigid sigmoidoscope, and histological sections should be 3 μm thick taken 50 μm apart.

Introduction

Cancer of the colon and rectum is now the second most common cause of cancer death in the United States (1). Studies of colon cancer using colon cancer, or even its adenoma precursor, as an endpoint in clinical trials require extremely large sample sizes, prolonged follow-up, or both; consequently, such studies are extremely expensive, and few have been undertaken. Furthermore, maintaining prolonged complex interventions, such as dietary interventions, in clinical trials is difficult and personnel intensive. From observational epidemiological studies, a few consistent associations appear to be emerging; however, by and large, the results of these studies have been inconsistent (2). In such studies, investigations of diet-disease associations are severely hampered by the homogeneity of dietary constituents within populations, the multitude of dietary factors and their interactions, and the lack of accuracy of current dietary measures (2, 3). Thus, the development of colon cancer intermediate endpoints or precursors that can be measured easily, the use of which would require relatively small sample sizes and interventions of short duration (months), would be an important advance in understanding colon cancer etiology and assessing interventions to reduce colon cancer risk (3–7).

Colorectal epithelial cell proliferative kinetics have been measured to assess risk of colon cancer and to test the efficacy of potential chemopreventive agents. There are no prospective data that allow prediction of risk from labeled biopsies, but the overall evidence is strong and relatively consistent. Several studies (8–22) have reported that, compared to patients at low risk for colon cancer, patients with colon cancer (8–17) and patients in every category known to be at higher risk for colon cancer [those with a history of sporadic adenoma (8–11, 13–17), familial polyposis (12, 18), ulcerative colitis (8, 19, 20), or a family history of colon cancer (12, 13, 21), and the elderly (9, 22)], on average, exhibit in their normal-appearing mucosa both an increased colonic epithelial cell proliferation rate and an extension of the colon crypt proliferative zone from the lower (basal) 60% of the crypt to include the upper (luminal) 40% of the crypt. In patients with previous colon cancer or sporadic...
adenomas, these changes also predict adenoma recurrence (23, 24). In large bowel tumors in humans, an upward shift in the proliferative zone is found in colon cancers and adenomas but not in hyperplastic polyps (25). As reviewed elsewhere (26–28), proliferative changes in normal-appearing mucosa have been shown to be a consequence of both cancer-initiating and cancer-promoting agents; proliferative changes both precede and accompany colonic neoplasms in rodents given chemical carcinogens, and a high fat diet produces proliferative changes in both rodents and humans. Animal experimental evidence and preliminary evidence in humans strongly suggest that these two proliferation abnormalities (hyperproliferation and upward shift of the proliferation zone) are reversible biomarkers or precursors for colon neoplasia (26–29). In humans, the two proliferation abnormalities appear to be independent variables (15, 30), and rectal biopsy findings on both measures reflect those throughout the colon (17, 31).

Colon crypt epithelial cell proliferation has primarily been measured by two different categories of laboratory methods: incorporation techniques and identification of endogenous cell cycle markers. Two common incorporation techniques have been used: labeling S-phase colorectal epithelial cells with tritiated thymidine (12–14, 16–24, 28) or with BrdUrd1 (25, 31–36). In the S-phase of the cell cycle, the nucleus is gathering nucleotides, such as thymidine, and using them to synthesize DNA in preparation for cell division (proliferation). Radioactive labeled thymidine (tritiated thymidine) is incorporated in place of normal thymidine and is identified through microautoradiography. BrdUrd, a thymidine analogue that is also incorporated in place of normal thymidine, is identified through immunohistochemical techniques.

The second basic method of identifying cell proliferation, identification of endogenous cell cycle markers, has largely involved the PCNA technique, another immunohistochemical technique. This technique identifies a cell protein, proliferating cell nuclear antigen (the auxiliary protein of DNA polymerase δ), that begins to increase in late G1 of the cell cycle and reaches a peak during S phase (37–39). Although production stops after S-phase, the protein has a long half-life and can be identified in cells into G2 (39, 40).

There are few data available to aid investigators in selecting which method of identifying cell proliferation to use or to help in avoiding the numerous pitfalls in measuring colorectal epithelial cell proliferation in humans. To address this lack of information, we kept records on various aspects of measuring colorectal epithelial cell proliferation in humans for approximately 1300 patient visits (1280 during randomized trials) on nearly 500 individuals (447 randomized to trials) over a 6-year period. Methods investigated were rectal biopsy procedures; general histological procedures, labeling procedures, including the tritiated thymidine, BrdUrd, and the PCNA immunohistochemical techniques; biopsy scoring procedures in general; and PCNA immunohistochemistry scoring procedures in particular. Cost effectiveness issues were also explored.

Materials and Methods

Sources of Data

Data were kept during the methods development phases and the clinical phases of the following four randomized controlled clinical trials using colorectal epithelial cell proliferation as an end point:

Trial 1: A pilot randomized, double-blind, placebo-controlled clinical trial in sporadic adenoma patients (n = 21) to test the efficacy of calcium supplementation in reducing colorectal epithelial cell proliferation as measured by the tritiated thymidine technique (28). Biopsies (n = 42) to measure colorectal epithelial cell proliferation parameters were obtained at baseline and at 2 months of follow-up.

Trial 2: A full-scale randomized, double-blind, placebo-controlled, three-armed parallel group clinical trial in sporadic adenoma patients (n = 193) to test the efficacy of calcium supplementation in reducing colorectal epithelial cell proliferation as measured by the BrdUrd and the PCNA techniques (41). Biopsies (n = 729) were obtained at baseline and at 1, 2, and 6 months of follow-up.

Trial 3: A pilot randomized, double-blind, placebo-controlled clinical trial in ulcerative colitis patients (n = 32) to test the efficacy of calcium supplementation in reducing colorectal epithelial cell proliferation as measured by the BrdUrd and the PCNA techniques. Biopsies (n = 63) were obtained at baseline and at 2 months of follow-up.

Trial 4: A full-scale randomized, controlled clinical trial in sporadic adenoma patients (n = 201) to test the efficacy of a high vegetable and fruit diet in reducing colorectal epithelial cell proliferation as measured by the BrdUrd and the PCNA techniques. Biopsies (n = 381) were obtained at baseline and at 12 months of follow-up.

Trial 1 preceded trials 2–4, and methodological issues related to this trial have been reported extensively elsewhere (28); these data are not repeated here, but a few comments relating our experience from it (and its use of tritiated thymidine) to the experience with BrdUrd and PCNA are provided. Trials 2 and 3 were conducted simultaneously, and trial 4 began approximately midway into trials 2 and 3.

Although our most interesting findings are from our comparisons of BrdUrd and PCNA labeling procedures, the following methods descriptions are presented sequentially from the biopsy procedures to the calculation of the measures of cell proliferation. Statistical methods developed by our group are reported elsewhere (30, 42), and methods using variance components to determine optimal sample sizes and cell proliferation scoring rules will be reported separately. The underlying principles in the development and comparisons of our biopsy procedures, general histological procedures, labeling procedures, and biopsy scoring procedures were practicality, feasibility, economics, reliability, and statistical characteristics for use in large epidemiological observational and intervention studies.

Biopsy Procedures

Procedures Development. A particular consideration in the development of our approach was the need for procedures that would be feasible and reliable for large studies in which biopsies are obtained during the course of normal busy clinical days by multiple practicing physicians with varying levels of dedication to the study. Thus, the biopsy procedure needed to be safe, inexpensive, quick, and to consistently yield biopsies with rectal mucosa containing adequate numbers of scorable crypts on which the labeling procedure was successful. Flexible sigmoidoscopy was rejected a priori because it was judged to be too time-consuming and expensive for large-scale observational epidemiological studies and clinical trials. For trial 1, biopsies were taken 2–3 cm above the level of the anus with jumbo cup flexible endoscopic forceps through an anoscope.

---

1 The abbreviations used are: BrdUrd, 5-bromodeoxyuridine; PCNA, proliferating cell nuclear antigen; DAB, 3,3′-diaminobenzidine; LI, labeling index; CV, coefficient of variation.
For trial 2, in the methods development phase, biopsies were taken with standard-sized flexible endoscopic forceps, jumbo cup flexible endoscopic forceps, and rigid sigmoidoscopy forceps. Biopsies were also taken 2-3 cm above the level of the anus through an anoscope and 10 cm above the level of the anus through a rigid sigmoidoscope.

Procedures during Clinical Trials. Biopsies for all trials were taken from normal-appearing rectal mucosa "non-prep," that is, without any enemas or bowel-cleansing preparations. During the clinical phase of trial 2, all biopsies were taken 10 cm above the anus through a rigid sigmoidoscope, initially using rigid sigmoidoscopy forceps, and subsequently using jumbo cup flexible endoscopic forceps mounted on a semirigid rod. For trials 3 and 4, all biopsies were taken 10 cm above the anus through a rigid sigmoidoscope using jumbo cup flexible endoscopic forceps mounted on a semirigid rod. For the clinical phases of all trials, the goal was to procure three macroscopically adequate biopsies.

Procedures Comparisons. The comparisons of biopsy procedures reported herein are based on a clinical series rather than on a randomized study.

General Histological Methods
Procedures Development. The particular goals in the development of our general histological procedures were to: (a) handle and cut biopsies so as to avoid scoring the same crypts twice and, because consistently obtaining an adequate number of scorable crypts was a challenge, to do so while maximizing tissue yield; (b) cut the biopsies so as to avoid, for the slide scorer, eye strain and ambiguity in assessing which cells in a crypt should be counted; and (c) obtain consistently adequate numbers of scorable crypts on which the labeling procedure was successful. Histological sections of biopsies were examined to determine a range of crypt widths and to assess labeling and scoring characteristics. During the pretrial development for trial 2, biopsies were incubated and labeled with BrdUrd, either full-sized as taken or divided into variously sized pieces. Preventing loss of PCNA antigen sites by the use of ethanol rather than formalin as a fixative for BrdUrd or PCNA has been demonstrated by others (36).

Procedures during Clinical Trials. Biopsies for BrdUrd and PCNA were fixed in 70% ethanol for at least 12 h and then processed for 5 h in a tissue processor as follows: 80% ethanol for 30 min at 0°C, 95% ethanol for 30 min at 0°C two times, 100% ethanol for 30 min at 0°C two times, xylene for 30 min at 0°C two times, and low temperature paraffin for 15 min at 55°C four times. For BrdUrd, these steps took place after incubation for BrdUrd uptake. For PCNA, incubation was not required. Two biopsy specimens per patient per visit were embedded in one paraffin block. Biopsies were embedded, positioned, and cut with a microtome so that crypts were longitudinally sectioned from base to lumen. Section lengths were 3 μm thick and taken 50 μm apart (taking sections 50 μm apart precluded scoring the same crypt more than once while maximizing tissue yield). Sections were then mounted on microscope slides. For each biopsy visit, four slides, each containing five section levels, were prepared. Prior to immunohistochemical detection of BrdUrd or PCNA, slides were deparaffinized at room temperature in the standard way with the following reagents and times: xylene for 5 minutes three times, 100% ethanol for 2 minutes two times, 95% ethanol for 2 minutes two times, 80% ethanol for 2 minutes two times, and then distilled water for 2 minutes once. For BrdUrd, slides were hydrolyzed in 2.0 N HCl for 1.5 h.

Procedures Comparisons. The general histological procedures comparisons reported herein are based on a clinical series rather than on a randomized study, and the reported results, in some instances, were subjective.

Labeling Methods
Procedures Development. For the BrdUrd incubation procedure, a variety of shaking devices, shaking speeds, and other shaking procedures were used. Incubation with and without hyperbaric oxygen was also examined. For PCNA, a variety of combinations of different brands and concentrations of primary, secondary, and tertiary immunohistochemistry agents were investigated, as were numerous other variations in immunoperoxidase procedures. The goal was to have nuclei that stained dark brown (i.e., only those that stained dark brown and not those that stained a lighter brown) correspond to any nuclei stained brown by an optimally done BrdUrd procedure; this, as will be addressed further below, was equivalent to a goal to attain the highest level of PCNA detection without background stain.

Procedures during Clinical Trials. The technique used for labeling with tritiated thymidine has been described elsewhere (28). After developmental investigations described above (see also "Labeling Methods" in "Results"), the following BrdUrd and PCNA techniques were used in trials 2, 3, and 4. The original plan for the studies was to label S-phase cells by incorporation of BrdUrd and subsequent identification by immunohistochemical methods. This protocol required that immediately after the biopsies were obtained, they were to be placed in a minimum essential medium for suspension cultures (SMEM). Within 15 min, the biopsies were stretched out flat, lumen side up on bibulous paper; transferred to a vial containing the culture medium with 0.01775 mg/ml BrdUrd; and placed under a 95% O2/5% CO2 atmosphere at two atmospheres of pressure on a rocker platform for 2 hours of incubation. At the end of the 2 hours, the biopsy vials were placed in ice to stop the reaction, and the biopsies were fixed in 70% ethanol. Despite apparently successful developmental and pilot work, the BrdUrd procedure labeling failure rate during the clinical phases of trials 2 and 3 proved unacceptable, and immunohistochemical detection of PCNA was added (described further below) and run on the same biopsies. Both the BrdUrd and the PCNA with incubation as for BrdUrd (hereinafter referred to as PCNA/BrdUrd) procedures were run on different slides from all biopsies from the two studies. In the PCNA/BrdUrd procedure, all of the steps between the biopsy orientation on bibulous paper and its placement in 70% ethanol were unnecessary; however, to ensure uniform handling of all biopsies throughout the studies, all of the steps were retained. (Although the PCNA/BrdUrd procedure is not a recommended procedure, it is included in this paper because of the insights it provided into the problems introduced by incubation independent of the antibody being tested for.) For trial 4, where sufficient tissue was available, samples were split three ways for analysis by three methods: BrdUrd, PCNA/BrdUrd, and PCNA without incubation (the latter hereinafter referred to as PCNA). The general procedures for subjecting slides to immunohistochemical analysis for BrdUrd or PCNA were that a monoclonal antibody to BrdUrd or PCNA was applied, followed by further immunoperoxidase technique steps to color the resulting antigen-antibody complexes concentrated in the nuclei of proliferating cells, thereby permitting visual identification of nuclei of cells in S phase and, in the case of PCNA (40, 43, 44);
near S phase. The detailed methods are given in the following two paragraphs.

Slides subjected to immunohistochemical analysis for BrdUrd were prepared as follows. Slides were deparaffinized (as discussed above), hydrolyzed in 2.0 N HCl for 1.5 h, placed in 0.1 M borax (Sigma B-0127) to stop the reaction, and rinsed in PBS. The slides were then mounted on a Sequenza stainer (Shandon) and received three 5-min washes with PBS (pH 7.4). Three drops of a blocking agent, 5% normal goat serum (Signet kit), were applied for 20 min. The slides received three 5-min washes with PBS and then three drops of anti-BrdUrd monoclonal antibody (Becton Dickinson 7580) were added for an additional 60 min. The slides then received three 5-min washes with PBS, and three drops of the linking agent (Signet kit #4) were applied for 20 min. The slides again received three 5-min washes with PBS, and three drops of the labeling agent (Signet kit #5) were added to the slides for 20 min. Following three more 5-min washes with PBS, the slides were dismounted from the Sequenza and put in a Coplin jar with PBS. A fresh DAB working solution (0.4 g Sigma D-0127 DAB and 10 ml of distilled water) was prepared, and the slides were stained with it for 5 min. The slides received two 5-min washes with distilled water and then were counterstained with Harris hematoxylin for 3 s and rinsed in distilled water twice more. Slides were then dehydrated (80% ethanol for 1 minute, then 95% ethanol for 1 minute, then 100% ethanol for 1 minute, and then xylene for 1 minute), and Permount and coverslips were applied.

Slides subjected to immunohistochemical analysis for PCNA were prepared as follows. Slides were deparaffinized (as discussed above) and rinsed in PBS. The slides were then mounted on a Sequenza stainer (Shandon) and received three 5-min washes with PBS (pH 7.4). Three drops of a blocking agent, 5% normal horse serum (Vector S-2000; diluted in PBS), were applied for 20 min, then three drops of PC-10 clone PCNA-antibody (Oncogene NA03; diluted 1:200 with PBS) were added for an additional 45 min. The slides then received three 5-min washes with PBS and three drops of the secondary antibody, rabbit anti-mouse IgG (DAKO E413; diluted 1:200 with PBS), were applied for 45 min. The slides again received three 5-min washes with PBS, and the endogenous peroxide was blocked with 0.3% H2O2 for 10 min. Following three more 5-min washes with PBS, three drops of the tertiary agent, avidin-biotin complex (Vector Standard kit PK-6100 at half strength), were applied for 45 min. The slides received three more 5-min washes with PBS and then were developed by immersing them with the chromogen DAB (Sigma D-9015) in a 1 mg/ml solution with 0.003% H2O2 for 2–5 min while monitoring the chromogenic development using light microscopy. Slides were then rinsed with two changes of distilled water, counterstained with Harris’s hematoxylin (hematoxylin for 1 min, rinse in tap water, two quick dips in 1% acid alcohol, rinse in tap water, two dips in ammonia water, and rinse in tap water), dehydrated (80% ethanol for 1 min, then 95% ethanol for 1 min, then 100% ethanol for 1 min two times, then 100% ethanol/xylene for 1 min, and then xylene for 1 min), and Permount and coverslips were applied.

Procedures Comparisons. The tritiated thymidine, BrdUrd, and PCNA methods were compared for feasibility, ease of use, and costs. These comparisons were primarily subjective. The BrdUrd and PCNA techniques were also compared for: (a) level of labeling success and ease of attaining a consistent level of labeling success; (b) susceptibility of the measures of levels of cell proliferation rate (LI) and distribution of labeled cells in colon crypts (Φh) to extraneous influences such as biopsy forceps and size of biopsies; (c) levels of measures of cell proliferation rate (LI) and distribution of labeled cells in colon crypts (Φh); (d) intra- and inter-rater scoring reliability; and (e) statistical properties, including coefficients of variation of the two measures, visit-to-visit correlation, and components of variance. These comparisons are based on a clinical series rather than on a randomized study.

Other Considerations. Because PCNA has not been validated previously in human populations, measures of cell proliferation characteristics between ulcerative colitis and sporadic adenoma patients were compared using the PCNA technique using data from trials 2 and 3, which were conducted concurrently. We have reported elsewhere (41) that PCNA labeling can be used to detect modulation of colorectal epithelial cell proliferation kinetics by chemopreventive agents (primary results of trial 2).

Biopsy Scoring Methods

Procedures Development. Biopsy scoring procedures development has been described previously in connection with trial 1 (28).

Procedures during Clinical Trials. Colon crypts longitudinally sectioned from base to lumen were analyzed. The total number of cells and number of labeled cells for each crypt scored were counted. The cell count position of each BrdUrd or PCNA-labeled cell within a crypt was recorded. The scoring for trials 2, 3, and 4 were the same. The scoring for BrdUrd and PCNA were identical in all but one respect; for BrdUrd, a cell was considered labeled if it had any detectable amount of labeling, and for PCNA, two grades of labeling intensity were recorded as described below. The scoring procedures for PCNA with or without incubation were identical. A scoreable crypt was defined as an intact crypt extending from the muscularis to the lumen. Only whole crypts were counted: crypts with only one intact column were not used. The crypt base must have been touching the muscularis mucosa, or if the muscularis was missing or the biopsy curled, the base must have extended down as far as those of the surrounding crypts. Crypts must have extended to the lumen but did not have to be fully open if the crypt column could still be clearly followed to the lumen. Likewise, if other areas (such as the base) had more than one row of cells, the crypt could be counted if one continuous column could be followed. Crypts with cell loss of greater than two cells (as a result of handling or cutting) were not used. Crypts did not have to be perfectly U-shaped or symmetrical.

Countable cells were defined as crypt cells in line in a single column of nuclei that extended from base to lumen. A cell slightly out of line in the lumen was counted if it appeared to have been dislodged from the gap between the cells immediately above it or below it in the column and if it was overlapped by those same cells immediately above or below it. Small gaps (one cell width) in the crypt column were not counted as cells. A structure that could not definitely be identified as a cell nucleus versus a nonspecific fragment was not counted. The last cell in the crypt column was considered to be the last in-line cell before turning the corner at the luminal surface.

An unlabeled cell was defined as a cell with a blue nucleus. For BrdUrd, a labeled cell was defined as a cell with a nucleus that had any degree of brown color. For PCNA, a weakly labeled cell (a cell in late G1 or early S) was defined as a cell with a nucleus that was light brown in color (1–2+ intensity). This was distinguished from any background stain that occasionally occurred. A strongly labeled cell (a cell in
S-phase) was defined as a cell with a nucleus that was dark brown in color (3–4" intensity). These cells tended to stand out prominently on the slides. Whenever there was any doubt whether a cell was unlabeled or weakly labeled, it was scored as unlabeled. If there was any doubt whether a cell was weakly labeled or strongly labeled, it was scored as weakly labeled.

Before scoring, control slides were checked, and the patient’s slides were scanned for biopsy adequacy. If the biopsy was adequate and the staining procedure appeared successful, scoring began. The slide reader began with the first BrdUrd or PCNA-stained slide on the patient. The slide was oriented in a standardized fashion, and the section levels on the slides were viewed in sequence using light microscopy. Scoring was begun at the first complete crypt found. The reader moved through a section level from left to right until a scorable crypt was located.

Once a scorable crypt was located, the biopsy number, slide number, section level number, and crypt number were entered on a lap top computer containing a scoring data entry program. The crypt was always counted under ×400. Once focused, focusing up and down to find a cell was not allowed but was allowed to distinguish whether a cell was labeled or not. Cells within a crypt were counted by beginning counting at the top right of the crypt and continuing down and around to the top left. While counting, a number was entered into a computer data entry program that permitted analysis of the cell count position and other characteristics of each cell. Each cell was identified as an unlabeled cell, a weakly labeled cell, or a strongly labeled cell. The crypt base center cell was also visually identified. (The scorer was also asked to identify which column they would have scored first if they were following an alternate protocol used in other laboratories. In the alternate protocol, the scorer counts one column, then counts the second column from the crypt base center cell upwards to the first labeled cell. The total number of cells in the second column is then estimated as being the same as that in the first column.)

The goal in our protocol was to score a minimum of 8-10 columns, then estimate as being the same as that in the first column. Whenever there was any doubt whether a cell was weakly labeled or strongly labeled, it was scored as weakly labeled.

Procedures Comparisons. The following scoring or cell proliferation calculation issues were investigated: (a) whether it makes an important difference in clinical trials if crypts with no labeled cells are counted; (b) whether it makes any difference if the crypt base center cell is selected visually or mathematically (the cell that is exactly one-half of the way, by cell count, from the top cell in one column to the top cell in the other column); (c) whether it makes any difference if biopsies with less than five scorable crypts are used in analyses; (d) whether it makes any difference if all cells in a crypt are counted versus all cells in one column are counted and the cell count is estimated as twice the cell count in the one counted column; (e) whether, for PCNA, the S-phase or the all labeled cell LI and Ph, is preferable; and (f) whether it is necessary to calculate both an LI and a Ph.

Statistical Methods

χ² tests were used to compare biopsy forceps with respect to biopsy quality, labeling failure, and complications. Because LI and Ph were not normally distributed, all analyses for these two variables were done using natural logarithm transformed data (41). Comparisons of the labeling methods were done using paired t tests, and correlations presented are Pearson product moment correlations. Intra- and inter-rater reliability were calculated as intraclass correlation coefficients. Comparisons between adenoma and ulcerative colitis patients were made using multivariate ANOVA adjusting for age and sex. ANOVA on log-transformed variables from clinical trial control groups was used to assess the components of variance.

Results

Biopsy Procedures. Taking biopsies 2–3 cm above the anus sometimes produced inadequate biopsies as a result of obtaining squamous epithelium rather than columnar epithelium, and some patients experienced pain during the procedure. These problems were eliminated when the protocol was to obtain 5 biopsies 10 cm above the anus.

Biopsies obtained by the standard-sized flexible endoscopic forceps tended to be the smallest and most shallow; those obtained by the rigid sigmoidoscopic forceps tended to be the largest and deepest; and those obtained by the jumbo cup flexible endoscopic forceps tended to be intermediate in both respects. Individual biopsy pinches taken with the standard-sized flexible endoscopic forceps tended to be easy to handle, but because they contained fewer crypts, a greater number of biopsy pinches were required to get the same number of scorable crypts as with other biopsy forceps. The larger number of biopsies that had to be handled presented difficulty in timely placement of oriented biopsies in incubation for BrdUrd (not a major concern for PCNA). Biopsies insufficiently deep to see the bases of crypts also were a more frequent occurrence than with other forceps. On the other hand, biopsies taken with rigid
Measuring Colonic Epithelial Cell Proliferation

Biopsy forceps effects

<table>
<thead>
<tr>
<th>Biopsy Quality</th>
<th>Rigid sigmoidoscopic</th>
<th>Jumbo flexible</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of biopsies</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>No crypts</td>
<td>375</td>
<td>698</td>
</tr>
<tr>
<td>&lt;8 crypts</td>
<td>28</td>
<td>7.5</td>
</tr>
</tbody>
</table>

Labeling failure

<table>
<thead>
<tr>
<th>PCNAPCNA/BrdUrd</th>
<th>No. of biopsies</th>
<th>%</th>
<th>No. of biopsies</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>BrdUrd</td>
<td>375</td>
<td>685</td>
<td>113</td>
<td>30.1</td>
</tr>
<tr>
<td>PCNA/BrdUrd</td>
<td>354</td>
<td>590</td>
<td>15</td>
<td>4.2</td>
</tr>
<tr>
<td>PCNA*</td>
<td>2</td>
<td>346</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Complications

<table>
<thead>
<tr>
<th>No. of biopsies</th>
<th>Bleeding</th>
<th>Pain</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of biopsies</td>
<td>Mild</td>
<td>Severe</td>
</tr>
<tr>
<td>378</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>713</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

Bleeding

<table>
<thead>
<tr>
<th>No. of biopsies</th>
<th>Pain</th>
<th>Moderate</th>
<th>Severe</th>
</tr>
</thead>
<tbody>
<tr>
<td>378</td>
<td>6</td>
<td>0.3</td>
<td>0.1</td>
</tr>
<tr>
<td>713</td>
<td>3</td>
<td>0.3</td>
<td>0.0</td>
</tr>
</tbody>
</table>

P for rigid sigmoidoscopic versus jumbo cup flexible endoscopic forceps; χ² test.

Immunohistochemistry for BrdUrd uptake.

PCNA/BrdUrd, immunohistochemistry for PCNA after incubation as for BrdUrd.

Immunohistochemistry for PCNA.

sigmoidoscopic forceps were more difficult to handle because they tended to curl, thus necessitating that they be cut into several smaller pieces to eliminate the curling. Biopsies taken with the jumbo cup flexible endoscopic forceps proved simplest to handle, because only two to three biopsy pinches were required, curling was not enough of a problem to require cutting them into more than two smaller pieces, and the biopsies were almost always of sufficient depth.

Biopsy quality (presence of scorable crypts) was comparable between biopsies obtained by rigid sigmoidoscopic forceps and those by jumbo cup flexible endoscopic forceps (Table 1). Overall, approximately 4% of biopsies contained no crypts, and approximately 8% contained fewer than eight crypts (less than optimal for statistical purposes, but still evaluable (28)). Biopsies obtained by the flexible forceps compared to the rigid forceps resulted in a significantly lower proportion that had labeling failure for BrdUrd (9.3% versus 30.1%) and for PCNA/BrdUrd (1.6% versus 4.2%); biopsies obtained by the flexible forceps and analyzed for PCNA had no labeling failures.

Biopsy complications occurred with the rigid forceps. Of 378 biopsies obtained using rigid sigmoidoscopic forceps, 12 (3.2%) resulted in mild bleeding that required no treatment, three (0.8%) resulted in moderate bleeding that required a few minutes pressure on the biopsy site and observation in the office for 1 h, and 3 (0.8%) required several minutes pressure and/or cautery of the biopsy site as well as observation for greater than 1 h. Of 713 biopsied with jumbo cup flexible endoscopic forceps, one (0.1%) had mild bleeding that required no treatment; and none had moderate or severe bleeding. Of 20 patients biopsied with standard-sized flexible endoscopic forceps (data not shown), none experienced a bleeding complication of any degree. Reports of pain were uncommon and were comparable by biopsy forceps used. Overall, mild discomfort was reported at 0.8% of biopsy visits, and moderate pain was reported at 0.2% of biopsy visits. No patient reported severe pain. No patients sustained a perforation or developed an infection as a result of any of the biopsy procedures.

General Histological Procedures. Taking histological sections 3 μm thick produced slides with the best uptake and labeling and ease of biopsy scoring. Taking the sections 50 μm apart precluded scoring the same crypt more than once while maximizing tissue yield.

General histological procedures had substantial effects on labeling. BrdUrd labeling was highly susceptible to "edge labeling," i.e., labeling of cells only in crypts at the margins of the biopsies. This edge labeling was effected by the biopsy forceps used (see above), the size and the depth of the biopsies (larger biopsies and thicker biopsies resulted in more with only edge labeling), the shaking device used, the shape of the biopsy container used during incubation, shaker speed, the length of incubation, and whether hyperbaric conditions were used during incubation. These factors, to a lesser extent (see below), also affected PCNA/BrdUrd. Biopsy size did not apparently affect PCNA, and the incubation factors, of course, did not apply.

PCNA labeling was affected by the PCNA-antibody clone, the brands of primary, secondary, and tertiary immunohistochemical agents used, the concentrations of antibody used, as well as numerous other variations in the immunohistochemical procedures used. These types of influences were found also to be operative for the immunohistochemical analysis for BrdUrd. Labeling Procedures. Because the tritiated thymidine and BrdUrd procedures required orientation and initiation of incubation within 15 min of biopsy, it was necessary to perform biopsies in a central location near the cell proliferation laboratory. Multiple laboratories were not feasible. Biopsies obtained for PCNA without incubation required only orientation and placement in 70% ethanol, a task that could be easily performed by a variety of personnel. Because incubation within 15 min of biopsy is required for tritiated thymidine and BrdUrd, trained laboratory personnel had to be available for these techniques. Recruitment rates are important in all epidemiological studies, and retention of participants is important in clinical trials. In our studies, we found that many potential participants preferred to come in for visits at the beginning of the work day, over lunch, or at the end of a work day. Biopsies requiring incubation took 2 or more hours of handling. To accommodate these factors, specialized research cell proliferation laboratory personnel had to be available 10–11 h a day to handle biopsies obtained during the course of an 8-h clinic day. Laboratory personnel did not need to be available to handle biopsies to be processed for PCNA; furthermore, these biopsies could be handled by clinic staff and could even have been mailed to a remote laboratory.

The tritiated thymidine procedure proved to be exceedingly time-consuming, and in our experience, not well suited for use in large-scale epidemiological observational studies or clinical trials. The BrdUrd technique took significantly less time to perform than the tritiated thymidine technique but took longer than the PCNA technique by the length of time required for the BrdUrd incubation procedure. Costs (primarily personnel-associated) were directly related to the length of the differ-

Table 1 Biopsy forceps type and biopsy quality, labeling failure, and complications

<table>
<thead>
<tr>
<th>Biopsy forceps effects</th>
<th>Biopsy forceps type</th>
<th>Rigid sigmoidoscopic</th>
<th>Jumbo flexible</th>
<th>χ²</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of biopsies</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>No crypts</td>
<td>375</td>
<td>698</td>
<td>11</td>
<td>38</td>
</tr>
<tr>
<td>&lt;8 crypts</td>
<td>28</td>
<td>7.5</td>
<td>61</td>
<td>8.7</td>
</tr>
</tbody>
</table>

PCNA/BrdUrd

<table>
<thead>
<tr>
<th>No. of biopsies</th>
<th>%</th>
<th>No. of biopsies</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>BrdUrd</td>
<td>375</td>
<td>685</td>
<td>113</td>
</tr>
<tr>
<td>PCNA/BrdUrd</td>
<td>354</td>
<td>590</td>
<td>15</td>
</tr>
<tr>
<td>PCNA*</td>
<td>2</td>
<td>346</td>
<td>0</td>
</tr>
</tbody>
</table>

Complications

<table>
<thead>
<tr>
<th>No. of biopsies</th>
<th>Bleeding</th>
<th>Pain</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of biopsies</td>
<td>Mild</td>
<td>Severe</td>
</tr>
<tr>
<td>378</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>713</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

P for rigid sigmoidoscopic versus jumbo cup flexible endoscopic forceps; χ² test.

Immunohistochemistry for BrdUrd uptake.

PCNA/BrdUrd, immunohistochemistry for PCNA after incubation as for BrdUrd.

PCNA labeling was affected by the PCNA-antibody clone, the brands of primary, secondary, and tertiary immunohistochemical agents used, the concentrations of antibody used, as well as numerous other variations in the immunohistochemical procedures used. These types of influences were found also to be operative for the immunohistochemical analysis for BrdUrd. Labeling Procedures. Because the tritiated thymidine and BrdUrd procedures required orientation and initiation of incubation within 15 min of biopsy, it was necessary to perform biopsies in a central location near the cell proliferation laboratory. Multiple laboratories were not feasible. Biopsies obtained for PCNA without incubation required only orientation and placement in 70% ethanol, a task that could be easily performed by a variety of personnel. Because incubation within 15 min of biopsy is required for tritiated thymidine and BrdUrd, trained laboratory personnel had to be available for these techniques. Recruitment rates are important in all epidemiological studies, and retention of participants is important in clinical trials. In our studies, we found that many potential participants preferred to come in for visits at the beginning of the work day, over lunch, or at the end of a work day. Biopsies requiring incubation took 2 or more hours of handling. To accommodate these factors, specialized research cell proliferation laboratory personnel had to be available 10–11 h a day to handle biopsies obtained during the course of an 8-h clinic day. Laboratory personnel did not need to be available to handle biopsies to be processed for PCNA; furthermore, these biopsies could be handled by clinic staff and could even have been mailed to a remote laboratory.

The tritiated thymidine procedure proved to be exceedingly time-consuming, and in our experience, not well suited for use in large-scale epidemiological observational studies or clinical trials. The BrdUrd technique took significantly less time to perform than the tritiated thymidine technique but took longer than the PCNA technique by the length of time required for the BrdUrd incubation procedure. Costs (primarily personnel-associated) were directly related to the length of the differ-

---

Downloaded from cebp.aacrjournals.org on March 31, 2017. © 1997 American Association for Cancer Research.
ent procedures; thus, the most economical procedure was the PCNA procedure.

For investigators wishing to begin measuring colorectal epithelial cell proliferation, learning time may also be a factor in choosing a labeling technique. Fig. 1 shows our "learning curve" for biopsy labeling success in sporadic adenoma patients over time by labeling method. Labeling success with BrdUrd began at 0%, rose gradually to about 85% after 5 months of experience, and thereafter was subject to substantial fluctuations (primarily drops) in the success rate. PCNA/BrdUrd labeling was nearly 100% successful from the start, but as with the BrdUrd, variation was noted month to month. PCNA labeling without incubation was 100% successful throughout its use. As noted in Table 1, labeling success for BrdUrd was more affected by nonlaboratory factors such as biopsy forces. (Of further note is that, in ulcerative colitis patients, BrdUrd labeling success was 0%, whereas PCNA/BrdUrd labeling success was 100%.)

Differences were also noted in cell proliferation values measured by the different labeling techniques. Comparison of cell proliferation measurements from the various methods are presented in Table 2. The table shows data from the 583 person-visits from trials 2 and 4 in which both BrdUrd and PCNA/BrdUrd were successfully used; the 196 person-visits from trial 4 in which PCNA/BrdUrd and PCNA were used; and the 133 person-visits from trial 4 in which BrdUrd and PCNA were used.

As seen in Table 2, the mean number of cells counted per crypt were highest when using the PCNA technique, lowest when using the BrdUrd technique, and intermediate when using the PCNA/BrdUrd technique. As a result of this analysis, representative slides were reviewed. In all cases, crypts selected for scoring were again judged to be complete crypts with no evidence of missing cells, and re-scoring resulted in confirmation of the original readings. Comparisons of the S-phase LI and $\Phi_h$ by the BrdUrd and the PCNA/BrdUrd techniques reveal no substantial differences. However, comparisons of the S-phase LI and $\Phi_h$ by the PCNA/BrdUrd and the PCNA techniques revealed higher LI and lower $\Phi_h$ values by PCNA. These latter findings are mirrored in the "all-labeled cell" measures (defined in "Materials and Methods" in "Cell Proliferation Calculation Methods"), except that those for the all-labeled cell $\Phi_h$ are not statistically significant. In general, the correlations of measures by the three labeling methods are higher between BrdUrd and PCNA/BrdUrd than between PCNA/BrdUrd and PCNA. In the analyses presented in Table 2, the comparisons presented were made using "all biopsies" [all biopsies obtained after reaching fairly consistent BrdUrd labeling success (see Fig. 1)]. These analyses were repeated (data not shown) using only "best biopsies" (same as "all biopsies" except only biopsies with at least five scorable crypts that contained labeled cells). The "best biopsy" comparisons were to investigate whether, when done optimally, the different measurement methods provide different results; the "all biopsy" comparisons were to investigate whether, in a "real world setting," the different measurement methods provide more widely different results. The various analyses based on the two different biopsy inclusions/exclusions produced results that showed a slight tendency toward higher LI and $\Phi_h$ values when "best biopsies" data were analyzed for the PCNA/BrdUrd method, but results otherwise were not substantially different.

Intra- and inter-rater reliability for scoring rectal biopsies for measures of colorectal epithelial cell proliferation, by labeling method, are shown in Table 3. No clear pattern is apparent, and overall, the scoring reliability is comparable across labeling methods.

Comparisons of statistical properties of colorectal epithelial cell proliferation measurements obtained by different labeling methods are shown in Table 4. Several features are noted: (a) the CV for the LI is lower than that for the $\Phi_h$, regardless of labeling method; (b) the CV for the LI tends to be lower for BrdUrd and higher for PCNA, but the reverse is true for the $\Phi_h$; (c) for the two PCNA methods, the CVs for the $\Phi_h$ are lower by the S-phase-labeled cell scoring method than by the all-labeled cell scoring method; (d) using data from the placebo control "natural history" groups, in the ANOVA of the BrdUrd versus the PCNA/BrdUrd methods, a greater proportion of the variance in measures of LI and $\Phi_h$ over time (2 months) was between-person using the PCNA/BrdUrd method. In the ANOVA of the PCNA/BrdUrd versus the PCNA methods, a greater proportion of the variance in measures of LI and $\Phi_h$ over 12 months was between-person using the PCNA/BrdUrd method. In the ANOVA of the PCNA/BrdUrd versus the PCNA methods, a greater proportion of the variance in measures of LI and $\Phi_h$ over 12 months was between-person using the PCNA/BrdUrd method; (e) Two-month visit-to-visit correlations for LI and $\Phi_h$ using BrdUrd versus the PCNA/BrdUrd methods were higher using the PCNA/BrdUrd method. Twelve-month visit-to-visit correlations for LI and $\Phi_h$ using PCNA/BrdUrd versus the PCNA methods were higher using the PCNA method.

Two questions about the use of the PCNA labeling method are whether it can be used to distinguish between patients with different conditions, and whether it can be used to detect treatment effects. A comparison of mean multivariate-adjusted (ANACOVA) measures of cell proliferation in sporadic adenoma patients (from trial 2) versus ulcerative colitis patients (from trial 3) using PCNA/BrdUrd showed qualitatively higher in all, and substantially and statistically higher values for most, measures in ulcerative colitis patients (S-phase LI: 4.3 versus 6.7, $P < 0.001$; all-labeled cell LI: 16.1 versus 32.9, $P < 0.0001$; S-phase LI: 5.6 versus 7.8, $P = 0.21$; and all-labeled cell $\Phi_h$: 10.1 versus 18.5, $P < 0.0001$). In the full-scale calcium trial in sporadic adenoma patients (study 2), substantial, statistically significant reductions in the $\Phi_h$ in the calcium groups versus the placebo groups were found (41).
Measuring Colonic Epithelial Cell Proliferation

Table 2 Comparisons of colorectal epithelial cell proliferation kinetics measured by BrdUrd versus PCNA/BrdUrd, PCNA/BrdUrd versus PCNA, and BrdUrd versus PCNA on "all biopsies".

| Proliferation measure | Mean | SD | Mean | SD | Correlation coefficient | p/
|-----------------------|------|----|------|----|-------------------------|---
| (n = 583)             |      |    |      |    |                         |   |
| No. cells/crypt       | 115  | 14 | 131  | 16 | 0.70                    | <0.001 |
| S-phase LI*           | 4.0  | 1.9| 3.9  | 2.3| 0.26                    | 0.007 |
| S-phase O\(h\)_0      | 5.9  | 6.7| 5.8  | 6.3| 0.26                    | 0.19 |
| (n = 196)             |      |    |      |    |                         |   |
| PCNA/BrdUrd           |      |    |      |    |                         |   |
| No. cells/crypt       | 120  | 21 | 137  | 16 | 0.31                    | <0.001 |
| S-phase LI            | 3.1  | 1.9| 4.6  | 2.9| 0.20                    | <0.001 |
| S-phase O\(h\)        | 6.7  | 7.7| 3.0  | 3.9| 0.14                    | <0.001 |
| All labeled cell LI*  | 10.2 | 7.0| 16.5 | 9.2| 0.21                    | <0.001 |
| All labeled cell O\(h\)_0 | 8.4 | 7.2| 7.8  | 6.2| 0.09                    | 0.66 |
| (n = 133)             |      |    |      |    |                         |   |
| PCNA                  |      |    |      |    |                         |   |
| No. cells/crypt       | 117  | 14 | 135  | 15 | 0.39                    | <0.001 |
| S-phase LI            | 4.1  | 1.9| 4.6  | 3.0| 0.24                    | 0.75 |
| S-phase O\(h\)        | 8.4  | 7.6| 2.8  | 3.9| 0.19                    | <0.001 |

* Immunohistochemistry for BrdUrd uptake.
* Immunohistochemistry for PCNA after incubation as for BrdUrd.
* Immunohistochemistry for proliferating cell nuclear antigen.
* All biopsies after reaching a fairly consistent BrdUrd labeling success rate.
* Pearson product moment correlation coefficient calculated on raw scale data for no. of cells/crypt and on natural logarithm transformed data for LI and O\(h\)_0.
* From paired t test comparison on raw scale data for no. of cells/crypt and on natural logarithm transformed data for LI and O\(h\)_0.

Table 3 Intra- and inter-rater reliability for scoring rectal biopsies for measures of colorectal epithelial cell proliferation by labeling method.

<table>
<thead>
<tr>
<th>Labeling method</th>
<th>Intra-rater coefficient*</th>
<th>Inter-rater coefficient*</th>
</tr>
</thead>
<tbody>
<tr>
<td>BrdUrd</td>
<td>0.99</td>
<td>0.79</td>
</tr>
<tr>
<td>S-phase LI*</td>
<td>0.96</td>
<td>0.79</td>
</tr>
<tr>
<td>S-phase O(h)_0</td>
<td>0.87</td>
<td>0.61</td>
</tr>
<tr>
<td>PCNA/BrdUrd</td>
<td>0.88</td>
<td>0.65</td>
</tr>
<tr>
<td>S-phase LI</td>
<td>0.92</td>
<td>0.76</td>
</tr>
<tr>
<td>S-phase O(h)</td>
<td>0.90</td>
<td>0.78</td>
</tr>
<tr>
<td>All-labeled cell LI*</td>
<td>0.93</td>
<td>0.92</td>
</tr>
<tr>
<td>All-labeled cell O(h)_0</td>
<td>0.76</td>
<td>0.63</td>
</tr>
<tr>
<td>PCNA</td>
<td>0.93</td>
<td>0.92</td>
</tr>
<tr>
<td>S-phase LI</td>
<td>0.76</td>
<td>0.63</td>
</tr>
<tr>
<td>S-phase O(h)</td>
<td>0.85</td>
<td>0.92</td>
</tr>
<tr>
<td>All-labeled cell LI</td>
<td>0.83</td>
<td>0.87</td>
</tr>
</tbody>
</table>

* Intraclass correlation coefficient.
* Immunohistochemistry for BrdUrd uptake.
* LI (proportion of crypt cells that are labeled); only S-phase cells scored as "labeled."
* Proportion of labeled crypt cells that are in the upper 40% of the crypt; only S-phase cells scored as "labeled."
* Immunohistochemistry for PCNA after incubation as for BrdUrd.
* LI (proportion of crypt cells that are labeled); all-labeled cells were scored as "labeled."
* Proportion of labeled crypt cells that are in the upper 40% of the crypt; only S-phase cells were scored as "labeled."
* Proportion of labeled cells that are in the upper 40% of the crypt; only S-phase cells were scored as "labeled."
* Immunohistochemistry for PCNA.

Biopsy Scoring Procedures. One scoring protocol issue that varies among laboratories is whether to count crypts with no labeled cells. We calculate that encountering crypt sections with no proliferating cells is possible but should be a rare event. In any case, no substantial or statistically significant differences were noted with inclusion or exclusion of these crypts. A second scoring protocol issue is whether all cells in a crypt should be counted or whether it is sufficient to count the cells on one side of the crypt and to double this value to get the crypt total cell count. We found no substantial or statistically significant differences in these two counting protocols whether we doubled the side the slide reader chose to double or the one they did not choose. Assigning crypt base center cell designation by visual means versus by mathematical determination also had no impact on cell proliferation measurements.

Cell Proliferation Measures. For 545 biopsies labeled with PCNA/BrdUrd, the correlations between the S-phase and the all-labeled cell LI and O\(h\)_0 were 0.84 and 0.65, respectively. The correlations between the LI and the O\(h\)_0, as determined by S-phase and all-labeled cell scoring were 0.08 and 0.33, respectively.

Discussion

These findings indicate that the PCNA labeling method for measuring colorectal epithelial cell proliferation in humans may be preferable to that of the incubation-based labeling methods for use in larger epidemiological observational and interventional studies. The method is easier to learn, easier to perform, takes less time to perform, and is more economical than the other methods. The lack of an incubation requirement provides for more flexibility than any incubation-based method in the design and implementation of larger observational epidemiological studies and clinical trials. Labeling failure occurs less often and is less variable. Cell proliferation measures are less susceptible to hard-to-control extraneous factors such as variations in biopsy size. Finally, the statistical properties of cell proliferation measures determined with PCNA labeling are generally more favorable for detecting differences in clinical trials; a greater proportion of the variability is between-person,
I

(57.5%) 55 3.3 (100%) 7.4 (82.9%)
-1

S-phase LI 21 1.1 (34.4%) 2.7 (42.5%) 41 -0.1 0.2

S-phase LI 21 .35 0.51 41 -0.1 0.2

S-phase LI 21 .04 0.01 41 -0.2 0.2

All-labeled cell 

PCNA/BrdUrd 0h

-1

and labeling an endogenous cell protein; in both procedures, the biopsies were incubated. The PCNA and PCNA/BrdUrd procedures were identical in every respect except that the PCNA/BrdUrd biopsies underwent incubation as for BrdUrd; in both procedures, the same endogenous cell protein was labeled. The quality of results for PCNA/BrdUrd were intermediate between BrdUrd and PCNA. BrdUrd, on balance, produced the most attenuated LI estimates and the most unwanted variability. That it performed worse than the PCNA/BrdUrd in these regards was apparently because of the variability in BrdUrd uptake. That PCNA/BrdUrd performed worse in these regards than PCNA was apparently because of variability induced by incubation. Subjectively, slides processed for PCNA without incubation were very clear, with all brown staining confined to nuclei. Slides processed for PCNA after incubation for which the corresponding BrdUrd slide had only edge labeling contained some cells with brown stain in the cytoplasm, suggesting that these cells died during incubation and that PCNA leaked from the nuclei into the cytoplasm. These results and observations indicate that: (a) incubation results in the death of most cells of the upper 40% of the crypt; only S-phase cells were scored as "labeled." (b) Incubation results in variability in these cells died during incubation and that PCNA leaked from the nuclei into the cytoplasm. These results and observations indicate that: (a) incubation results in the death of most cells of the upper 40% of the crypt; only S-phase cells were scored as "labeled." (b) Incubation results in variability in these cells died during incubation and that PCNA leaked from the nuclei into the cytoplasm. These results and observations indicate that: (a) incubation results in the death of most cells of the upper 40% of the crypt; only S-phase cells were scored as "labeled." (b) Incubation results in variability in these
correlate better with the BrdUrd measures than with the PCNA without incubation.

PCNA can be used to distinguish between groups of patients with different conditions that predict risk of colon cancer. In our study, it distinguished differences in both the LI and \( \Phi_s \) between sporadic adenoma and ulcerative colitis patients. In another study of 80 patients, it distinguished differences in the LI but not the \( \Phi_s \) between patients with colonic neoplasms and persons without adenoma or other high-risk condition (46). PCNA can be used to test treatment effects of nutritional interventions in sporadic adenoma patients (41). The biopsy scoring reliability using PCNA is comparable to that using BrdUrd. The PCNA strongly labeled and all-labeled cell counting methods provide parallel, highly correlated results with similar statistical and reader reliability properties. For these reasons, although the strongly labeled counting method provides an estimate of S-phase cells similar to that of the BrdUrd method (when the PCNA strong labeling is titrated to correspond to BrdUrd labeling), and the all-labeled cell method provides an estimate of cells in and around S-phase (a larger number), the two counting methods can be considered inter-changeable for ranking patients according to their cross-sectional estimates of cell proliferation rate and distribution of crypt proliferative zone.

Thus, it is appropriate to use the PCNA labeling technique in epidemiological studies, and it is more reliable and feasible than other labeling methods for use in these studies. It is less susceptible to visit-to-visit artifactual variability, while providing greater ability to distinguish between individuals. The primary disadvantage to the PCNA technique compared to the BrdUrd technique is the extra training and inconvenience related to the different intensities of staining; however, this disadvantage, which on its face originally appeared would be a major disadvantage, has proved to be minor and pales in comparison to the disadvantage of the errors in cell proliferation measurement induced by incubation procedures for the incubation-based methods such as the BrdUrd technique.

One enigma from our labeling methods comparison data was the finding of a greater mean number of cells per crypt for PCNA, an intermediate number for PCNA/BrdUrd, and a lower number for BrdUrd. This phenomenon has been noted by others (47). A greater mean number of cells per crypt was seen for PCNA versus BrdUrd in normal patients in one small study (47), but no difference was seen in normal or resected cancer patients in another small study (48). One hypothesis for our findings was that surface and other cells were being sloughed during incubation; however, a review and re-scoring of representative slides did not indicate this as the source. Furthermore, incubation-induced sloughing not would not account for why PCNA/BrdUrd crypt cell counts were intermediate. We are unable to supply a satisfactory explanation.

Other conclusions are: (a) for studies depending on biopsy procurement by multiple physicians (a necessity for many larger studies), rectal biopsy protocols specifying that biopsies be taken from 10 cm above the level of the anus through a sigmoidoscope, rather than from less than 10 cm or through an anoscope, result in a greater proportion of adequate biopsies; (b) flexible sigmoidoscopy is not deemed as economical or practical as rigid sigmoidoscopy for procuring rectal biopsies for large studies in which the endoscopy is not part of usual clinical care; (c) the use of flexible endoscopy forceps is generally safe; however, rigid biopsy forceps cause excess bleeding in some patients and are not recommended. Jumbo cup and standard-sized flexible endoscopic forceps are equally safe; however, the jumbo cup forceps may produce fewer inadequate biopsies, important if time and other considerations limit the number of biopsies that can be obtained; (d) using biopsy protocols such as those described herein, rates of obtaining biopsies with no scorable crypts can be up to 5%, and with less than 8–10 crypts, up to 9%, and should be factored into study sample size considerations; (e) taking histological sections every 50 \( \mu \)m precludes scoring a colon crypt twice while maximizing tissue yield; (f) differences in scoring protocols used in different laboratories, such as whether to count crypts with no labeled cells, or whether to count whole or hemicrypts, has negligible impact on cell proliferation measurements; however, without extensive inter-reader standardization, even when using the same scoring protocol, reader effects are substantial (28); (g) the cell proliferation rate (LI) and the distribution of labeled cells within colon crypts (\( \Phi_s \)) are not highly correlated and should both be measured/calculated. This last point has received more rigorous statistical analysis elsewhere (30, 42) and has been noted empirically by other investigators (15).

A few points about limitations of the present study and a few caveats based on the cell proliferation literature are important to keep in mind: (a) although our study is based on 60–70-fold more samples and is more comprehensive in the number of issues examined relative to other published methodological studies of colorectal epithelial cell proliferation, many of the findings are those of a series of experiences rather than those of randomized comparisons; (b) although we tried innumerable variations of our presented protocols, there may be others that are better and that would have produced different results in comparing the basic methods (i.e., tritiated thymidine, BrdUrd, and PCNA). Indeed, although we adhere to a protocol for a given study, our methods continue to evolve, and we have implemented automated immunostaining, image analysis for scoring, the use of the Ki-67 (mib 1) antibody, and other changes that are presently being tested and subjected to the type analysis described in the present paper. Because our own data collection on these new changes are not yet complete and there are few data in these regards published by others, we cannot yet recommend them; (c) the mib 1 antibody to Ki-67, another endogenous cell cycle-related protein, has been used to assess cell proliferation with immunohistochemical methods (49), and this was not addressed in our study. This antibody appears to eliminate having to deal with different degrees of uptake; however, a report that variability in labeling is introduced by even small variations in how long the tissue to be analyzed was left in fixative (49) could pose a significant problem for large epidemiologic studies and needs further study; and (d) perhaps most importantly, colorectal epithelial cell proliferation as assessed by the PCNA method has not been adequately investigated regarding whether it can distinguish groups of persons who are adenoma-formers from those who are at low risk for colonic neoplasms. This is also true of the mib 1 method and may be considered true of all methods.

In summary, more work is needed to determine whether colorectal epithelial cell proliferation is a valid biomarker of risk or intermediate end point for colonic neoplasia in general and by the specific methods of assessing it (e.g., PCNA) in particular. Based on the current state of knowledge, we recommend, that for larger studies measuring colorectal epithelial cell proliferation from outpatient rectal biopsies in humans, the rectal biopsies be taken 10 cm above the anus using a flexible, preferably jumbo cup, endoscopic forceps through a rigid sigmoidoscope. We also recommend that histological sections be 3 \( \mu \)m thick taken 50 \( \mu \)m apart, that the PCNA technique be used to identify S-phase cells, and that both the LI and \( \Phi_s \) be measured using either the S-phase or the all-labeled cell scoring
protocols. Investigators need to be aware that small differences in general histological and PCNA laboratory protocols can produce large consequences, and that standardizing biopsy scorers is a time-consuming essential component of any cell proliferation study.

Acknowledgments

We thank Dr. Michael Wargovich for advice on the BrdUrd technique; Janine Einspahr and Dr. Dennis Ahnen for advice on the PCNA labeling technique; Colleen Forster for help with developmental work to adapt the PCNA labeling technique for large-scale studies; Mark Boldt and Patricia Winkels for advice and work on biopsy procurement and initial handling methods; Bryan Randall for development of the computer software necessary to score biopsy slides processed by the PCNA technique; and the physicians of Digestive Healthcare, PA for advice and work on biopsy procurement methods and implementing epidemiological studies at the University-private community interface that measure colorectal epithelial cell proliferation.

References

Cancer Epidemiology, Biomarkers & Prevention

Methodological findings and considerations in measuring colorectal epithelial cell proliferation in humans.


Updated version Access the most recent version of this article at:
http://cebp.aacrjournals.org/content/6/11/931

**E-mail alerts** Sign up to receive free email-alerts related to this article or journal.

**Reprints and Subscriptions** To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

**Permissions** To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.