Results of a randomized trial on the management of cytology interpretations of atypical squamous cells of undetermined significance

The ASCUS-LSIL Triage Study (ALTS)* Group

OBJECTIVE: This study was undertaken to compare alternative strategies for the initial management of a cytologic diagnosis of atypical squamous cells of undetermined significance (ASCUS).
STUDY DESIGN: A total of 3488 women with a community-based ASCUS interpretation were randomly assigned to immediate colposcopy, triage that was based on enrollment HPV DNA testing and liquid-based cytology at a colposcopy referral threshold of high-grade squamous intraepithelial lesion (HSIL), or conservative management based on repeat cytology at a referral threshold of HSIL. All arms included 2 years of semiannual follow-up and colposcopy at exit. Loop electrosurgical excision procedure was offered to women with histologic diagnoses of cervical intraepithelial neoplasia (CIN) grade 2 or 3 at any visit or persistent CIN grade 1 at exit. The study end point was 2-year cumulative diagnosis of CIN grade 3.
RESULTS: The 2-year cumulative diagnosis of CIN grade 3 was 8% to 9% in all study arms. The immediate colposcopy strategy yielded 53.6% sensitivity for cumulative cases of CIN grade 3 diagnosed over 2 years. The human papillomavirus (HPV) triage strategy referred 55.6% of women and detected 72.3% of cumulative cases of CIN grade 3. A conservative management strategy of repeat cytology at the HSIL threshold referred 12.3% of women while detecting 54.6% of cumulative CIN grade 3. To compare triage tests, we re-estimated the performance of HPV and cytology in successfully referring women with underlying CIN grade 3 (ie, ignoring the insensitivity we discovered in colposcopically directed biopsies). A single enrollment HPV test identified 92.4% of the women diagnosed with CIN grade 3. Serial cytology, even at an ASCUS threshold, would have required two visits to achieve similar sensitivity (95.4%) and would have referred 67.1% to colposcopy.
CONCLUSION: HPV triage is at least as sensitive as immediate colposcopy for detecting CIN grade 3 and refers about half as many women to colposcopy. Follow-up that used repeat cytology is sensitive at an ASCUS referral threshold but requires two follow-up visits and ultimately more colposcopic examinations than HPV triage. (Am J Obstet Gynecol 2003;188:1383-92.)

Key words: Atypical squamous cells of undetermined significance, human papillomavirus, cervix, clinical management, randomized clinical trial, cytology, colposcopy

There are approximately 55 million Papanicolaou (Pap) tests performed in the United States yearly, of which about 5% are interpreted as "atypical squamous cells of undetermined significance" (ASCUS) and 2% are interpreted as "low-grade squamous intraepithelial lesion" (LSIL). Therefore, more than 3 million women per year are affected by an ASCUS or LSIL interpretation, with a total resultant clinical management cost that has been estimated to be several billion dollars. The clinical dilemma is the need to identify the small minority of women with underlying cervical intraepithelial neoplasia (CIN) grade 3 or cancer, which must be weighed against the very high prevalence of ASCUS and LSIL, arguing against aggressive management.

The ASCUS-LSIL Triage Study (ALTS) was designed to compare alternative strategies for the initial management of these common cytology interpretations found during cervical cancer screening. A decade ago, the discovery that infection with oncogenic types of human papillo-

From the ALTS Group.
Supported by the National Cancer Institute, National Institutes of Health Department of Health and Human Services contracts No. CN-55153, CN-55154, CN-55155, CN-55156, CN-55157, CN-55158, CN-55159, and CN-55105. Some of the equipment and supplies used in this study were donated or provided at reduced cost by Digene Corporation, Gaithersburg, Md; Cytyc Corporation, Boxborough, Mass; National Testing Laboratories, Fenton, Mo; Denvu, Tucson, Ariz; and TriPath Imaging, Inc, Burlington, NC.
Received for publication December 20, 2002; revised February 12, 2003; accepted March 25, 2003.
Reprint requests: Diane Solomon, MD, Breast and Gynecologic Cancer Research Group, Division of Cancer Prevention, National Cancer Institute, NIH, DHHS, Executive Plaza North, Room 2123, Rockville, MD 20852. E-mail: ds87v@nih.gov
**A complete list of ALTS investigators begins on page 1411.*
doi:10.1067/mob.2003.457

mavirus (HPV) causes virtually all cases of cervical cancer raised the possibility of incorporating knowledge of HPV natural history and HPV testing into cervical cancer screening programs that had previously relied on combinations of cytology and colposcopically guided biopsies. ALTS was designed after the conclusion of the 1991 Bethesda Workshop,[1] taking into account differing opinions regarding the optimal management of ASCUS and LSIL cytology interpretations. Different participants at the Workshop and a subsequent National Cancer Institute (NCI)–sponsored advisory meeting favored either immediate colposcopy, cytology follow-up, or triage of ASCUS and LSIL by using HPV DNA testing.[2] Immediate colposcopy was considered the reference standard of optimal sensitivity and safety. Cytology follow-up was considered more conservative and less invasive and permitted regression of self-limited HPV infections known to cause abnormal cytology interpretations. Triage that used HPV DNA testing was based on a corollary of the new knowledge that virtually all cases of cervical cancer contain HPV DNA. Specifically, all HPV-negative women were reasoned to be at extremely low risk of cervical cancer, permitting triage away from colposcopic referral. The experts did not agree on which of the three management options, if any, was superior. This clinical "equipoise" on a topic affecting millions of women annually justified a major randomized clinical trial.

This set of four manuscripts includes the major results from ALTS, apart from a cost-utility analysis that is still pending. The ALTS Group hoped that, by placing the main articles together, those interested could easily read a common introduction and methods, included in this first article focusing on the initial management of a cytology interpretation of ASCUS, followed by integrated results and conclusions related to LSIL[3] and postcolpos- copy management.[4,5]

Material and methods

Overview. ALTS was a randomized clinical trial comparing three strategies for women with ASCUS and LSIL separately: immediate colposcopy, HPV triage, and conservative management, based on a program of repeat cytology.[6] All women, regardless of randomization assignment and initial management during enrollment, were scheduled for follow-up with cytology at 6-month intervals for 2 years. We achieved very high retention rates and adherence to protocols, which permitted thorough analyses of the performance of clinical outcomes but might be difficult to reproduce in typical clinical practice. We also supported all the activities at the ALTS clinical centers with expert quality control (QC) groups, raising the issue of the most useful choice of study end points. As the surrogate for cancer risk, we chose a priori a scientific end point of histologic CIN grade 3 as diagnosed by the pathology QC group. To supplement this main analysis, we used a clinical end point of histologic CIN grade 2 or 3 as diagnosed by pathologists at the four clinical centers.

Recruitment. ALTS involved four clinical centers: University of Alabama at Birmingham (Birmingham, Ala), Magee-Womens Hospital of the University of Pittsburgh Medical Center Health System (Pittsburgh, Pa), the University of Oklahoma (Oklahoma City, Okla), and the University of Washington (Seattle, Wash). The study was approved by the NCI and local institutional review boards. All women had a community-read cytology result of ASCUS or LSIL as a prerequisite for study entry. These slides (virtually all were conventional Pap smears) were requested from the community laboratories and were sent to the Pathology QC group for rereview for ALTS research purposes only. The review diagnosis did not affect subject eligibility or management.

Randomization arms. A total of 5060 women enrolled in the study from January 1997 to December 1998: 3488 women with ASCUS and 1572 with LSIL cytology. Routine follow-up and exit visits concluded in January 2001. Demographic characteristics of the enrollees are described more completely elsewhere.[6]

After determining eligibility and obtaining written informed consent, participants were randomly assigned to one of the three management arms: immediate colposcopy (IC, referral to colposcopy regardless of enrollment test results), HPV triage (referral to colposcopy if the enrollment HPV result was positive or missing, or if the enrollment cytology was high-grade squamous intraepithelial lesion [HSIL]), and conservative management (CM, referral to colposcopy if cytology was HSIL).

All women in each arm underwent the same enrollment pelvic examination with collection of specimens as outlined below under examination procedures. Referral to colposcopy at enrollment was based on the randomization arm and enrollment test results. (This was the *only* management decision that differed among arms.) Subsequent follow-up was the same for all arms. An exit examination, with colposcopy scheduled for *all* women regardless of arm or prior procedures, was performed at 2 years, as described under follow-up and exit management below.

Examination procedures. At each patient visit, nurse–clinicians typically conducted the pelvic examination and collected two cervical specimens. The first cervical specimen was collected with a Papette broom (Wallach Surgical, Orange, Conn) and was rinsed directly into a PreservCyt vial (Cytyc Corporation, Boxborough, Mass). This specimen was used for both the preparation of ThinPrep cytology slides (Cytyc Corporation, Boxborough, Mass) and for HPV testing using the Hybrid Capture 2 (HC 2) high-risk probe set (Digene Corporation, Gaithersburg, Md). A second cervical specimen, collected with a Dacron swab, was obtained for investigational HPV DNA typing; these results were not used for patient man-

agement in the trial. After the collection of the cervical specimens, the cervix was rinsed twice with a 5% solution of acetic acid, and 2 Cervigrams (National Testing Laboratories Worldwide, Fenton, Mo) were taken.

Patient management at enrollment. Women randomly assigned to the IC arm proceeded immediately to colposcopy or were given an appointment to return for the procedure within 3 weeks if colposcopy could not be performed the same day. Women randomly assigned to the HPV triage arm were called back for colposcopy if the HPV test was positive or not performed (missing), or if there was an ALTS clinical center enrollment ThinPrep diagnosis of HSIL or a glandular abnormality (these interpretations as a group have been termed HSIL). A missing HPV test result was most commonly the result of having less than 4 mL of residual specimen in the PreservCyt vial after preparing the ThinPrep, an arbitrary minimum volume. Women in the HPV triage arm with no HPV test results were triaged to colposcopy because it was considered to be an impractical triage strategy to recall women for repeat collection of a specimen for the HPV test alone. In the CM arm, only women with a clinical center ThinPrep cytology diagnosis of HSIL were referred to colposcopy. Unsatisfactory cytology led to recall for repeat specimen collection unless the patient had already been referred for colposcopy based on randomization (IC arm) or HPV test result (HPV triage arm). Very rarely, clinicians referred patients to colposcopy on the basis of visualizing a lesion suspicious for cancer during the pelvic examination. Any safety net notification (see below) issued by a QC group also triggered colposcopy, the only time that QC results affected patient management.

Patient management at follow-up. Women were scheduled to return for follow-up visits at 6, 12, and 18 months from the date of enrollment regardless of arm and treatment received. Pelvic examinations and specimen collections were conducted as at enrollment. Women were referred (or referred again) to colposcopy for a clinical center cytology diagnosis of HSIL or a safety net trigger and were treated by loop electrosurgical excision procedure (LEEP) for histologic CIN grade 2 or 3 diagnosed by the clinical center. During follow-up, HPV results were masked in all arms.

Patient management at exit. Exit visits, scheduled for approximately 24 months from the date of enrollment, included colposcopy for all women. The purpose of the exit visit was 2-fold: to ensure patient safety and to provide complete ascertainment of disease end points before a woman exited the study. Therefore, all available clinical information was unmasked and provided to the clinician conducting the exit pelvic examination and colposcopy. This included all previous cytology and histopathology reported by the clinical center and the Pathology QC group, the most recent Cervigram photograph and report, as well as all previous HPV results.

At exit, colposcopy was performed in the same manner as at enrollment and follow-up. However, the threshold for treatment was lower at exit; in addition to treating women with CIN grade 2 or 3 on colposcopically directed biopsy, women with persistent low-grade lesions were offered LEEP. A woman was considered to have a persistent low-grade lesion if the colposcopically directed biopsy at exit showed CIN grade 1 *and* cytology results from at least one of the previous two visits showed LSIL or HPV+ ASCUS.

Laboratory processing and interpretation of cervical specimens. Liquid-based, ThinPrep cytology slides were prepared from PreservCyt vial specimens according to the manufacturer's standard protocol. Slides were screened at each clinical center by a cytotechnologist and evaluated by a cytopathologist trained to read ThinPreps according to routine practice. Cytology results were recorded on a standardized data collection form with the use of the 1991 Bethesda System with subcategorical distinctions between HSIL–CIN grade 2 and HSIL–CIN grade 3. After the clinical center evaluation, slides were sent to the Pathology QC group for rescreening and rereview.

After the preparation of the ThinPrep, a 4-mL aliquot of the residual PreservCyt specimen was taken for HPV testing (although only half of the aliquot was typically tested). We used HC 2, a molecular hybridization assay, to detect a group of cancer-associated HPV types, including types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68. The low-risk probe set (types 6, 11, 42, 43, and 44) was not used. A chemiluminescent tag in HC 2 emits light in proportion to the amount of target HPV DNA in the specimen. The Food and Drug Administration–approved threshold of 1 pg of HPV DNA per milliliter of test solution was used as the threshold for a positive result. An HPV QC group monitored the performance of the HPV assay by using mock specimen controls with each run for the first 2 years of the study, as well as by random retesting of a percentage of the clinical specimens throughout the entire study.

Colposcopy and treatment. The standard protocol for colposcopy included conventional visual assessment, application of 5% acetic acid and identification of the squamo-columnar junction and transformation zone. Biopsy specimens, obtained for colposcopically suspected CIN, were placed in separate labeled vials containing 10% buffered formalin. Endocervical curettage was performed according to clinician judgment in cases where the extent of the lesion, or squamocolumnar junction, was not visible. Colposcopic impression was recorded on a standardized form. A computer-assisted digital imaging system (Denvu, Tucson, Ariz) was used by the colposcopist to capture images of the cervix and record the biopsy sites selected. Histologic interpretation of biopsy specimens was conducted at each center by using CIN terminology. After the interpretation by the center, all histology slides were sent to the Pathology QC group for

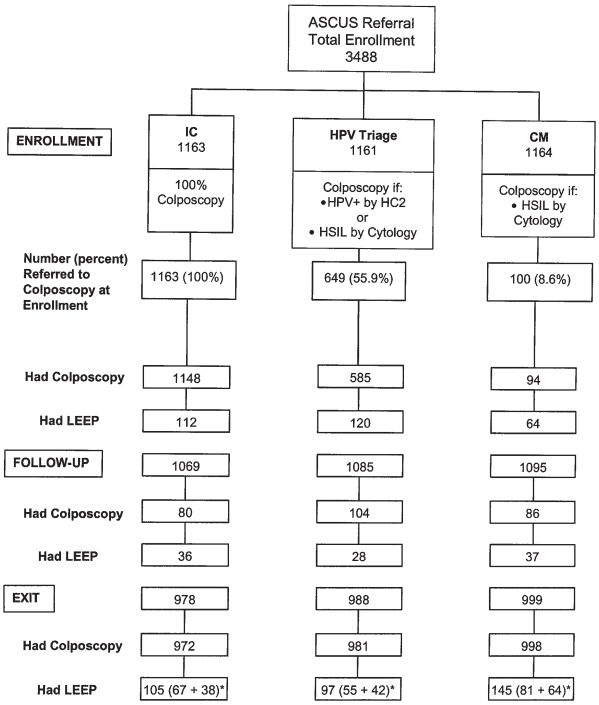

Figure. CONSORT diagram of participants in ALTS referred for ASCUS cytology. The number of women enrolled in each arm, the triage strategy for referral to colposcopy at enrollment, and the percent of women referred are shown at the top of the figure. The first row of numbers for the follow-up and exit periods reflect women remaining in the trial at that time. Subsequent rows indicate the number of women who had colposcopy and/or LEEP during the period. *Asterisk,* Exit LEEP numbers are subdivided in parentheses by the indication triggering the procedure: persistent low-grade disease vs presence or suspicion of high-grade disease, respectively.

re-evaluation; however, the management of the participant was based on the reading by the clinical center pathologist.

As per standard practice, all women with histologic CIN grade 2 or 3 diagnosed at the clinical centers were treated by LEEP. CIN grade 1 was not treated during the enrollment and follow-up periods. If the colposcopically directed biopsy result was discordant with an HSIL cytology or a high-grade colposcopic impression at the clinical center, repeat colposcopy was performed according to clinician judgment to clarify diagnostic discrepancies.

Pathology QC group. All referral slides, ThinPreps, and histology slides were sent to the Pathology QC group based at Johns Hopkins Hospital for rereview and final case definition. All cytology slides from enrollment and exit were rescreened by experienced study cytotechnologists and reviewed by the Pathology QC group. A portion of the cytology slides from follow-up were submitted for computer-assisted review (Neopath, TriPath Imaging, Burlington, NC). The pathology QC review protocol for histology slides included review by a QC pathologist masked to the original diagnosis. Any case with a diagnosis of CIN grade 2 or 3, by either pathology QC or the original clinical center, automatically went to a panel review composed of two to four QC pathologists unmasked to previous histology diagnoses. For all other cases, the first QC review diagnosis was compared with the clinical center diagnosis and, if concordant, that served as the final diagnosis. In the event of disagreement between the clinical center and the first QC reviewer, the case was sent to panel review. For all cases sent to panel, that review constituted the final diagnosis.

Safety notifications. In addition to providing expert interpretation for purposes of disease definition, the Pathology QC review was also designed to provide a "safety net" for study participants. For cytologic and histologic specimens, a pathology QC diagnosis of CIN grade 3 (that had been called less than CIN grade 2 at the center) triggered a safety notification sent by fax to the clinical centers. Cervigrams and digital colposcopic images also underwent external review for safety purposes. The threshold for safety notification for Cervicography and digital colposcopic images was "suspect cancer."

Statistical analyses. The primary study end point case definition was established a priori as a pathology QC histologic diagnosis of CIN grade 3, adenocarcinoma in situ (AIS), or cancer. Because there were so few cases of cancer (n = 2) or AIS (n = 1) among women with ASCUS, we refer to the scientific end point for simplicity as CIN grade 3. We also present a clinical end point of histologic CIN grade 2 or 3 as diagnosed at the clinical centers because women were treated on the basis of clinical center diagnoses at this threshold. For analyses related to time of diagnosis, we collapsed the periods into enrollment, follow-up, and exit. Additional procedures performed within 1 year of enrollment, as part of the continued diagnostic workup of a patient, are included in the enrollment period. A priori, the IC and HPV triage strategies were designed to detect CIN grade 3 at enrollment, based on the initial examination and colposcopic referral. However, the CM strategy relied on repeat cytology; therefore, detection of CIN grade 3 during either the enrollment or follow-up study periods was considered success.

The binomial distribution was used to compute exact CIs for proportions (eg, sensitivity). Pearson χ^2 tests for contingency tables were used to assess the associations between categorical variables (eg, cytology interpretations vs HPV test results). The McNemar test was used to assess the significance of differences in paired data, such as the comparison of the sensitivities of cytology and HPV testing in the same subjects. The χ^2 statistics for trend were calculated to test the significance of data with evident ordering (such as increasing severity of cytology interpretations related to HPV positivity). Life-table methods were used to account for censoring in analyses where such an

Table I. Clinical center enrollment liquid-based cytology diagnoses compared with HPV DNA test results*

	HPV DNA test result			
Cytology	*Negative (row %)*	*Missing (row %)†*	*Positive (row %)*	*Total (column %)‡*
Unsatisfactory or missing	9 (50.0%)	3 (16.7%)	6 (33.3%)	18 (0.5%)
Negative	935 (64.0%)	72 (5.0%)	453 (31.0%)	1460 (41.9%)
ASCUS	540 (47.8%)	38 (3.4%)	553 (48.9%)	1131 (32.4%)
LSIL	68 (10.7%)	37 (5.9%)	528 (83.4%)	633 (18.2%)
HSIL (CIN grade 2)	7 (3.4%)	9 (4.3%)	191 (92.3%)	207 (5.9%)
HSIL (CIN grade 3)	0 (0.0%)	3 (7.7%)	36 (92.3%)	39 (1.1%)
Total	1559 (44.7%)	162 (4.6%)	1767 (50.7%)§	3488 (100.0%)

*HC 2 includes probes for cancer-associated HPV types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68.
†4.6% of HC 2 results were missing because of <4 mL residual cytology specimen in Preservcyt collection vial after preparing the ThinPrep cytology slides.
‡$P_{trend} < .001$ for association between grade of cytologic abnormality and HPV DNA positivity.
§53.1% (1767/3326) of the HC 2 results that could be performed were positive.

adjustment was needed. All statistical tests were two sided and were considered significant at $P < .05$.

Results

ALTS enrolled 3488 eligible women with a community Pap smear interpretation of ASCUS, and randomly assigned 1163 to the IC arm, 1161 to HPV triage, and 1164 to the CM arm (Figure). Throughout the study, 93% of women had at least one follow-up visit, and 85% had an exit visit; retention did not differ by study arm. Virtually all women referred to colposcopy did attend, and virtually all referred for LEEP received treatment. Therefore, the details of the Figure reflect meaningful differences between study arms and periods, not biased participation rates.

Referral to colposcopy during the follow-up study period depended on HSIL cytology (or safety net findings) and the percentages were low (<10%) and not significantly different between study arms. Approximately one fifth of women had LEEP over the course of the study; this fraction did not vary by study arm. However, the timing of the LEEP differed among arms as a corollary of the time of detection of disease. About half of the LEEPs in IC and HPV were performed during the enrollment time period compared with one fourth in the CM arm. At exit, in addition to treating women with histologic CIN grade 2 or 3, women with persistent low-grade lesions (defined as CIN grade 1 on colposcopically directed biopsy *and* cytology results from at least one of the two immediately preceding visits showing LSIL or HPV+ ASCUS) were offered LEEP. Two hundred three women underwent LEEP triggered for persistent low-grade disease, representing one fourth of all LEEPs performed during the study.

Table I demonstrates the clinical center enrollment cytology interpretations compared with HPV testing results. Although all these women had been referred for ASCUS Pap smear interpretations in the community, the clinical center interpretations of the enrollment ThinPrep slides, taken on average 2.0 months later (median 1.7, range 0.3-6.0 months), were heterogeneous. Only 32.4% were again interpreted as ASCUS. There was a significant trend between the severity of the ThinPrep interpretation and HPV DNA positivity ($P_{trend} < .001$). Overall, 50.7% of women with ASCUS cytology were HPV DNA positive and another 4.6% had missing results because of a lack of sufficient residual material in the PreservCyt vial after cytology had been prepared, or for other technical reasons.

Table II. Clinical center histologic diagnosis from first colposcopically directed biopsy in IC arm

Clinical center histologic diagnosis	*No.**	*Percentage*
Missing valid biopsy result	26†	2.3%
No biopsy taken, normal colposcopic impression	273	23.8%
Negative	443	38.6%
CIN grade 1	300	26.1%
CIN grade 2	71‡	6.2%
CIN grade 3	35‡	3.0%
Total	1148	100.0%

*Of the 1163 women in the IC study arm, 15 women referred to colposcopy did not attend.
†Of the 1148 that attended, 19 had no biopsy specimen taken despite an abnormal colposcopic appearance, 1 had an unsatisfactory colposcopy without biopsy, and 6 had unsatisfactory biopsy specimens (n = 26 missing biopsy results).
‡The numbers of CIN grade 2 and CIN 3 are clinical center diagnoses of the initial colposcopically directed biopsy; these numbers cannot be directly compared with subsequent tables that use the pathology QC group diagnoses.

Women randomly assigned to the IC arm usually had colposcopy on the same day as enrollment (mean = 5.9, median = 0, mode = 0). Given universal colposcopy in the IC arm, the clinical center results reflect the distribution of disease detected at initial examination after ASCUS cytology (Table II). CIN grade 2 or 3 was diagnosed in

Table III. Cumulative histologic diagnoses of CIN grade 2 and CIN 3* by pathology QC group, stratified by study arm

	IC (column %)	*HPV triage (column %)*	*CM (column %)*	*P value†*	*Total (column %)*
CIN grade 2	92 (7.9%)	85 (7.3%)	55 (4.7%)	.005	232 (6.7%)
CIN grade 3	97 (8.3%)	101 (8.7%)	108 (9.3%)	.72	306 (8.8%)
CIN grade 2 or 3	189 (16.2%)	186 (16.0%)	163 (14.0%)	.26	538 (15.4%)
Total No. of women	1163 (100.0%)	1161 (100.0%)	1164 (100.0%)		3488 (100.0%)

*CIN grade 3 includes two cases of invasive cancer (1 each IC and CM arms) and one case of adenocarcinoma in situ (HPV arm).
†*P* values from χ^2 test for comparison between study arms. Direct comparisons of CIN grade 2 by study arm were statistically significant for CM vs either IC ($P = .002$) or HPV triage ($P = .01$).

Table IVA. Cumulative histologic diagnoses of CIN grade 3* by pathology QC group, stratified by study arm and period†

	IC	*HPV triage*	*CM*	*Total CIN grade 3*
Enrollment	**58 (59.8%)**	**76 (75.2%)**	**44 (40.7%)**	178 (58.2%)
Follow-up	14 (14.4%)	6 (5.9%)	**22 (20.4%)**	42 (13.7%)
Exit	25 (25.8%)‡	19 (18.8%)‡	42 (38.9%)‡	86 (28.1%)
Total	97 (100.0%)	101 (100.0%)	108 (100.0%)	306 (100.0%)

$P < .001$ from χ^2 test for overall comparison of study arm versus time of diagnosis of CIN grade 3. $P_{trend} = .06$ for IC versus HPV triage (0.04 for standard χ^2). $P_{trend} = .01$ for IC versus CM. $P_{trend} < .001$ for HPV triage versus CM.
*CIN grade 3 includes two cases of invasive cancer (1 each IC and CM arms) and 1 case of adenocarcinoma in situ (HPV arm).
†The figures in bold indicate the a priori–defined period for the strategy to successfully detect CIN grade 3 within the study arm (ie, enrollment for IC and HPV triage, and enrollment plus follow-up periods for CM). These numbers of CIN grade 3 include cases detected through safety net interventions; such cases are not counted as successes in the comparison of management strategy performance in Table V (CIN grade 3 detected through safety intervention: IC, n = 6; HPV, n = 3; CM, n = 7).

Table IVB. Cumulative histologic diagnoses of CIN grade 2 or 3 by clinical center pathologists, stratified by study arm and period

	IC	*HPV triage*	*CM*	*Total CIN grade 2 or 3*
Enrollment	119 (64.0%)	126 (70.4%)	60 (35.7%)	305 (57.2%)
Follow-up	24 (12.9%)	15 (8.4%)	32 (19.1%)	71 (13.3%)
Exit	43 (23.1%)	38 (21.2%)	76 (45.2%)	157 (29.5%)
Total	186 (100.0%)	179 (100.0%)	168 (100.0%)	533 (100.0%)

$P < .001$ from χ^2 test for overall comparison of study arm versus time of diagnosis of CIN grade 2 or 3. $P_{trend} = .34$ for IC versus HPV triage. $P_{trend} < .001$ for IC versus CM. $P_{trend} < .001$ for HPV triage versus CM.

9.2%, but no CIN of any grade was diagnosed in 62.4% either because the colposcopic impression was normal and no biopsy specimen was taken, or the biopsy specimen was negative.

Tables III through VI present complementary approaches to the analysis of the longitudinal data. We separately considered (1) the findings in the "study arm," (2) the performance of the "management strategy," and (3) the optimized "triage test performance."

Study arm findings. As the simplest, descriptive comparison of the study arms, Tables III and IVA present all disease endpoints diagnosed by the pathology QC group during the trial. Table III shows that the cumulative diagnoses of CIN grade 3, the a priori study end point, did not vary significantly by study arm (IC 8.3%, HPV 8.7%, CM 9.3%). However, the cumulative percentages of CIN grade 2 did vary by study arm ($P = .005$). There were significantly fewer diagnoses of CIN grade 2 in the CM arm (4.7%) than in either of the other two study arms (IC 7.9%, HPV, 7.3%, $P = .005$), thought to be a consequence of regression of missed prevalent CIN grade 2 in the CM arm (see Comment).

Table IVA shows the cases of CIN grade 3 in each study arm, stratified by period. Although the total percentage of CIN grade 3 diagnosed by the pathology QC group in each arm was equivalent, the timing of diagnosis was significantly heterogeneous ($P < .001$). Of the total CIN grade 3 cases in each arm, those in the HPV triage arm were diagnosed earliest, followed by the IC arm ($P_{trend} = .06$). In the CM arm, CIN grade 3 cases were diagnosed significantly later than in either of the other two study arms, and 38.9% of cases were not diagnosed until the exit period.

It is important to note that 11 cases of CIN grade 3 at exit were found only by offering LEEP to 203 women with persistent low-grade lesions (Figure). In terms of contri-

Table V. Performance of management strategies for detection of cumulative histologic diagnoses of CIN grade 3* by pathology QC group

	Study arm			
	IC	*HPV triage*	*CM*	*P value*
Sensitivity† for CIN grade 3 (%)	53.6% (43.2-63.8)	72.3% (62.5-80.7)	54.6% (44.8-64.2)	.01
Referral to colposcopy (%)	100% (99.7-100)	55.6% (52.6-58.4)	12.3% (10.5-14.3)	<.001

$P = .01$ from χ^2 test for sensitivity of IC versus HPV triage, $P = 1.0$ for IC versus CM, $P = .01$ for HPV triage versus CM.
*CIN grade 3 includes two cases of invasive cancer (1 each IC and CM arms) and one case of adenocarcinoma in situ (HPV arm).
†The management strategy performance calculations consider as "successes" only those cases of CIN grade 3 detected by the clinical application of the management strategy at the centers within the a priori–defined period for that strategy (ie, enrollment period for IC and HPV triage, and enrollment plus follow-up periods for CM) (see bold areas of Table IV). Cases of CIN grade 3 missed by the strategy but detected by safety net interventions and cases detected after the defined period for that strategy are not included in the numerator for calculating sensitivity.

Table VI. Estimated* triage test performance for detection of cumulative histologic diagnoses of CIN grade 3† by pathology QC group

	Sensitivity for CIN grade 3 (%) (CI)	*Referral (%) (CI)*
Enrollment HPV DNA test	92.4% (88.7-95.2)	53.1% (51.4-54.8)
HSIL cytology threshold‡		
1	35.5% (30.0-41.3)	7.1% (6.2-8.0)
2	48.3% (38.8-57.7)	10.2% (8.5-12.0)
3	60.2% (50.8-69.6)	11.7% (9.8-13.6)
LSIL cytology threshold‡		
1	59.3% (53.4-65.0)	25.1% (23.6-26.6)
2	74.1% (65.8-82.3)	31.7% (29.0-34.4)
3	82.0% (74.7-89.4)	37.2% (34.4-40.1)
ASCUS cytology threshold‡		
1	83.4% (78.7-87.5)	58.1% (56.4-59.8)
2	95.4% (91.4-99.3)	67.1% (64.4-69.8)
3	97.2% (94.1-100)	72.7% (70.1-75.4)

*For these estimates, missing test results, missed visits, and the timing of visits were ignored, to focus on the performance of the tests according to how many were completed.
†CIN grade 3 includes two cases of invasive cancer (1 each IC and CM arms) and 1 case of adenocarcinoma in situ (HPV arm).
‡Each cytology threshold reflects the finding of a cytologic abnormality greater than or equal to the cut point when cytology is performed one, two, or three times at approximately 6-month intervals. The enrollment HPV test was compared with the first cytology using data from all study arms to maximize statistical power. Because of extensive censoring in the IC and HPV arms, only data from the CM arm were used to estimate the performance of two or three repeat cytology examinations.

bution to total numbers of CIN grade 3, these 11 cases represent 12.8% of the cases diagnosed at exit and 3.6% of the total number of cases of CIN grade 3 in the study.

Table IVB addresses the clinical endpoint of CIN grade 2 or 3 diagnosed by the clinical center pathologists. Although the numbers of endpoints are greater, the percent distribution of time of diagnosis mirrors the findings that were based on the scientific end point.

Management strategy performance. In Table V, the management strategy performance calculations consider as "successes" only those cases of CIN grade 3 detected by the clinical application of the management strategy at the centers within the a priori–defined period for that strategy (ie, enrollment period for IC and HPV triage, and enrollment plus follow-up periods for CM—see bolded figures of Table IVA). Cases of CIN grade 3 missed by the strategy but detected by QC safety net intervention, and cases detected after the defined period for that strategy, are not included in the numerator for calculating sensitivity.

Table V compares the alternative management strategies on the basis of the sensitivity for the detection of CIN grade 3 and the percentage of women requiring colposcopy under that strategy. In IC, only 53.6% of cumulative cases of CIN grade 3 diagnosed over the 2-year study period were detected during the enrollment period. This management strategy required colposcopy in 100% of women, significantly more than the other two study arms. In HPV triage, 72.3% of cumulative cases of CIN grade 3 were detected during enrollment, a sensitivity that was significantly greater than IC ($P = .01$) or CM ($P = .01$). The sensitivity for detection of CIN grade 3 did not differ significantly between the IC and CM strategies. Of note, the HPV triage strategy theoretically depended on either

HPV-positive results or HSIL cytology at enrollment, but none of the cases of CIN grade 3 were referred to colposcopy on the basis of cytology alone. The percentage of women referred to colposcopy at enrollment by HPV triage was only about half of the universal referral in IC but significantly greater than the 12.3% referral (enrollment plus follow-up) for the CM strategy.

Triage test performance. While Table V shows the actual performance of the three alternative management strategies in clinical settings subject to the limitations of colposcopically directed biopsy and loss to follow-up, Table VI gives estimates of the theoretical, optimal test performance for HPV testing and cytology at three thresholds of colposcopic referral. For these estimates, we ignored the imperfect sensitivity of colposcopically directed biopsy and excluded missing test values, to evaluate (1) what percentage of cases of CIN grade 3 would have been referred on the basis of a positive triage test and threshold (% sensitivity) and (2) how many referrals would have resulted by using each triage test and threshold (% referral). Of the women originally referred to ALTS with ASCUS cytology who were ultimately found to have CIN grade 3, enrollment HPV testing would have properly triaged 92.4% (CI = 88.7-95.2) while referring 53.1% of women overall (CI = 51.5-54.9) (exclusion of missing tests accounts for difference with Table I). Examination of the sensitivities and referral percentages for various thresholds of repeat cytology, determined from the CM arm, demonstrates that repeating cytology twice at the ASCUS threshold would provide comparable sensitivity for detection of CIN grade 3. However, such a program of repeat cytology would refer significantly more women (67.1%) than a single HPV test (53.1%).

Comment

For women with ASCUS cytology interpretations, the ALTS data demonstrate that HPV triage is at least as sensitive as immediate colposcopy in the detection of underlying CIN grade 3, while nearly halving the number of women referred for colposcopy. Repeat cytology with colposcopic referral at an ASCUS threshold is also sensitive in detecting CIN grade 3 but requires repeated visits and leads to significantly more colposcopic examinations than does a single HPV test. Although the final cost-utility analyses of ALTS data are not yet complete, HPV testing is obviously an excellent strategy for the initial management of ASCUS, particularly when liquid-based cytology permits "reflex" HPV testing of a single cytology specimen. This conclusion was adopted by a recent American Society for Colposcopy and Cervical Pathology sponsored consensus conference on the management of women with abnormal cervical cytology.[7]

In our opinion, the ALTS data successfully address the controversy that has existed in the United States regarding the best clinical management of women with ASCUS cervical cytology[2,8-11] since the Bethesda System was introduced in 1988.[12] In the recently introduced 2001 Bethesda System, the ASCUS classification is revised to exclude the subcategory of "favor reactive," but retains the concept of "undetermined significance" (ie, the lack of sufficient morphologic features to permit definitive interpretation).[13] Unless the new "atypical squamous cells" category proves to be much more specific and restricted than anticipated, the results of ALTS based on the ASCUS terminology will still be fully applicable.

Three major management strategies have been used in the United States for management of women with ASCUS cervical cytology: immediate colposcopy, triage based on HPV DNA testing, and repeat cytology at 4- to 6-month intervals.[2] Immediate colposcopy has been assumed to be the safest option but with the disadvantages of high cost and potential overtreatment. Because the sensitivity of a single repeat conventional cervical cytology is relatively low, a program of repeat cytology has been proposed. Several large studies have evaluated the performance of HPV DNA testing to guide management in the ASCUS population, with sensitivity for detection of CIN grade 2 or 3 by using HC 2 reported as 83% to 100%.[14-19]

ALTS evaluated these three alternative strategies in a prospective, randomized fashion. Retention and compliance with recommended interventions in the trial were excellent, did not differ by arm, and therefore did not influence the outcomes of the study.

By their nature, ASCUS interpretations are not highly reproducible.[20] Only 32.4% of women with community ASCUS cytology had a repeat ASCUS interpretation on the enrollment ThinPrep as read by the clinical center pathologists. The association of severity of the clinical center enrollment ThinPrep interpretation and HPV positivity mirrors the broader observation that approximately half of women with ASCUS cytology from the community were positive for HPV, whereas more than 80% of women with LSIL cytology were HPV positive.

The overall percentage of CIN grade 2 or 3 in the ASCUS population as diagnosed by the pathology QC group was 15.4%, similar to that reported in other studies.[9,14-16,18,21] The cumulative rate of detection of CIN grade 3 did not vary by study arm. However, the cumulative percentage of CIN grade 2 alone did vary by study arm with the CM arm having significantly fewer CIN grade 2 diagnoses than either of the other study arms. These data strongly suggest a spontaneous regression of some cases of missed prevalent CIN grade 2 in the CM arm, and point to the advantage of using the a priori scientific end point of CIN grade 3 as a more stringent surrogate of cancer risk.

By trial design, each arm represented an alternative management strategy. These can best be evaluated by comparing the number of CIN grade 3 cases detected by the strategy without crediting CIN grade 3 detection by the

multiple safety interventions. We judged the success of the immediate colposcopy and HPV triage protocols by detection of CIN grade 3 during enrollment, while we evaluated conservative management based on CIN grade 3 detection during both enrollment and follow-up. The HPV triage strategy detected a higher percentage of CIN grade 3 than IC and CM. The seemingly greater sensitivity of HPV triage compared with IC may reflect that the colposcopist was generally aware of the positive HPV test and enrollment ThinPrep cytology results at the time of colposcopy (enrollment test results determined triage to colposcopy in this arm) and, thus, may have been more diligent in the colposcopic evaluation. We have no data to address this hypothesis directly. More than a third of the cases of CIN grade 3 in the CM strategy (which relied on HSIL referral to colposcopy) were not diagnosed until the exit visit that included universal colposcopy and liberal offering of LEEP treatment for persistent low-grade lesions. This suggests that repeat cytology at an HSIL threshold, although referring few women and similar in sensitivity to a single colposcopy, is not optimally sensitive for the timely detection of CIN grade 3.

The imperfect sensitivity of the initial colposcopy in all arms after colposcopic referral is of concern. For CIN grade 3 diagnosed during follow-up or exit, it is impossible to accurately separate missed prevalent from newly incident cases. Therefore, it is possible that some of the CIN grade 3 developed after enrollment and was detected appropriately at follow-up visits. Our review of complete records for cases of CIN grade 3 diagnosed after enrollment suggests, however, that many cases represented missed prevalent disease falling below the triggers for safety net notifications.

Because colposcopy and directed biopsy are not completely sensitive, we compared the theoretical optimal performance for HPV testing and cytology at three thresholds of referral on the basis of (1) the percentage of women ultimately found to have CIN grade 3 that would have been be referred on the basis of a positive triage test and (2) the percentage of the entire population that would be referred at that triage threshold (Table VI). A single enrollment HPV test would have appropriately triaged 92% of the women who were ultimately found to have CIN grade 3, while referring 53% of the total ASCUS population. Only 1.4% of the women who were HPV negative at enrollment were ultimately found to have CIN grade 3 over 2 years. The sole strategy of cytology follow-up that was equally sensitive (95%) in detecting women with CIN grade 3 would depend on at least two repeat cytology tests at an ASCUS threshold, referring 67% of women. In comparison to HPV triage, the additional number of referrals would be very large in absolute terms.

HPV triage is at least as sensitive as immediate colposcopy for detecting CIN grade 3 among women with ASCUS. A program of repeat cytology is also sensitive if an ASCUS threshold is maintained and loss to follow-up is minimal. The immediate colposcopy strategy is certainly the least specific, referring 100% of women to colposcopy. The ALTS longitudinal data suggest that HPV triage is the most effective strategy for management of women with ASCUS, as already suggested by a cost-utility analysis that is based on a model that closely approximated published ALTS data.[22]

REFERENCES

1. Broder S. Rapid communication. JAMA 1992;267:1892.
2. Kurman RJ, Henson DE, Herbst AL, Noller KL, Schiffman ME. Interim guidelines for management of abnormal cervical cytology: the 1992 National Cancer Institute Workshop. JAMA 1994;271:1866-9.
3. ASCUS-LSIL Triage Study (ALTS) Group. A randomized trial on the management of LSIL cytology interpretations. Am J Obstet Gynecol 2003;188:1393-400.
4. Cox JT, Schiffman M, Solomon D, ASCUS LSIL Triage Study Group. Prospective follow-up suggest similar risk of subsequent cervical intraepithelial neoplasia grade 2 or 3 among women with cervical intraepithelial neoplasia grade 1 or negative colposcopy and directed biopsy. Am J Obstet Gynecol 2002;188:1406-12.
5. Guido R, Schiffman M, Solomon D, Burke L, ASCUS LSIL Triage Study Group. Postcolposcopy management strategies for patients referred with low-grade squamous intraepithelial lesions or human papillomavirus DNA–positive atypical squamous cells of undetermined significance: a two-year prospective study. Am J Obstet Gynecol 2003;188:1401-5.
6. Schiffman M, Adrianza ME, ALTS Group. The ASCUS-LSIL Triage Study (ALTS): design, methods, and characteristics of trial participants. Acta Cytol 2000;44:726-42.
7. Wright TC Jr, Cox JT, Massad LS, Twiggs LB, Wilkinson EJ, 2001 ASCCP-Sponsored Consensus Conference. 2001 Consensus guidelines for the management of women with cervical cytological abnormalities. JAMA 2002;287:2120-9.
8. Frable WJ. ASCUS! ASCUS! Down the rabbit hole [editorial]. Cancer Cytopathol 1999;87:319-21.
9. Jones DE, Creasman WT, Combroski RA, Lentz SS, Waeltz JL. Evaluation of the atypical Pap smear. Am J Obstet Gynecol 1987;157:544-9.
10. Koss LG. The Papanicolaou test for cervical cancer detection: a triumph and a tragedy. JAMA 1989;261:737-43.
11. Bosch FX, Manos MM, Munoz N, Sherman M, Jansen AM, Peto J, et al. Prevalence of human papillomavirus in cervical cancer: a worldwide perspective: International Biological Study on Cervical Cancer (IBSCC) Study Group. J Natl Cancer Inst 1995;87:796-802.
12. National Cancer Institute Workshop. The 1988 Bethesda System for reporting cervical/vaginal cytologic diagnoses. JAMA 1989;262:931-4.
13. Solomon D, Davey D, Kurman R, Moriarty A, O'Connor D, Prey M, et al. The 2001 Bethesda System terminology for reporting results of cervical cytology. JAMA 2002;287:2114-9.
14. Wright TC Jr, Lorincz AT, Ferris DG, Rickart RM, Ferenczy A, Mielzynska I, et al. Reflex HPV DNA testing in women with abnormal Papanicolaou smears. Am J Obstet Gynecol 1998;178:962-6.
15. Manos MM, Kinney WK, Hurley LB, Sherman ME, Shieh-Ngai J, Kurman RJ, et al. Identifying women with cervical neoplasia: using human papillomavirus DNA testing for equivocal Papanicolaou results. JAMA 1999;281:1605-10.
16. Bergeron C, Jeannel D, Poveda J, Cassonnet P, Orth G. Human papillomavirus testing in women with mild cytologic atypia. Obstet Gynecol 2000;95:821-7.
17. Lin CT, Tseng CJ, Lai CH, Hsueh S, Huang HJ, Law KS. High-risk HPV DNA detection by Hybrid Capture II: an adjunctive test

for mildly abnormal cytologic smears in women τ or + 50 years of age. J Reprod Med 2000;45:345-50.
18. Shlay JC, Dunn T, Byers T, Baron AE, Douglas JM Jr. Prediction of cervical intraepithelial neoplasia grade 2-3 using risk assessment and human papillomavirus testing in women with atypia on Papanicolaou smears. Obstet Gynecol 2000;96:410-6.
19. Solomon D, Schiffman M, Tarone R, ALTS Study group. Comparison of three management strategies for patients with atypical squamous cells of undetermined significance: baseline results from a randomized trial. J Natl Cancer Inst 2001;93:293-9.
20. Stoler MH, Schiffman M, ALTS group. Interobserver reproducibility of cervical cytologic and histologic interpretations: realistic estimates for the ASCUS-LSIL Triage Study. JAMA 2001;285:1500-5.
21. Eskridge C, Begneaud WP, Landwehr C. Cervicography combined with repeat Papanicolaou test as triage for low-grade cytologic abnormalities. Obstet Gynecol 1998;92:351-5.
22. Kim JJ, Wright TC, Goldie SJ. Cost-effectiveness of alternative triage strategies for atypical squamous cells of undetermined significance. JAMA 2002;287:2382-90.