

AD_____

AWARD NUMBER: W81XWH-04-C-0083

TITLE: Internet-Based Cervical Cytology Screening Program

PRINCIPAL INVESTIGATOR: David C. Wilbur, M.D.
Barbara. A. Crothers, D.O.
John H. Eichhorn, M.D.
Min S. Ro, M.D.
Jeffrey A. Gelfand, M.D.

CONTRACTING ORGANIZATION: Massachusetts General Hospital
Boston, Massachusetts 02114-2554

REPORT DATE: April 2006

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE (DD-MM-YYYY) April 2006		2. REPORT TYPE Annual		3. DATES COVERED (From - To) 15 Mar 05 – 14 Mar 06	
Internet-Based Cervical Cytology Screening Program				5a. CONTRACT NUMBER W81XWH-04-C-0083	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) David C. Wilbur, M.D., Barbara. A. Crothers, D.O., John H. Eichhorn, M.D. Min S. Ro, M.D., Jeffrey A. Gelfand, M.D. E-mail: dwilbur@partners.org				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Massachusetts General Hospital Boston, Massachusetts 02114-2554				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
13. SUPPLEMENTARY NOTES					
14. ABSTRACT This project explores the combination of computerized automated primary screening of cervical cytology specimens in remote sites with interpretation of device-selected images transmitted via the Internet. The project is in 3 phases: 1) hardware/software and interface development and end user training with 200 case pilot trial; 2) a 500 case prospective pilot study with hardware/software adjustment, with the development of clinically applicable specimen triage and management guidelines; and 3) a 5000 case prospective clinical trial of the completed system, with report and development of a training and operation manual. During this report period, phase 1 data was analyzed and reported, hardware/software/ interface issues were resolved, and training materials were developed for phase 2. Internal IRB approvals were obtained for phase 2 (MGH/WRAMC), however, Army Office of Research Protections has not yet approved the phase 2 protocols (see full report), hence phase 2 patient accrual has not been initiated. Phase 3 preactivities have therefore also ceased.					
15. Subject Terms (keywords previously assigned to proposal abstract or terms which apply to this award) telecytology, cytopathology, telemedicine, cancer screening, health care information systems					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 26	19a. NAME OF RESPONSIBLE PERSON USAMRMC
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U			19b. TELEPHONE NUMBER (include area code)

Table of Contents

Cover.....	1
SF 298.....	2
Introduction.....	4
Body.....	5
Key Research Accomplishments.....	10
Reportable Outcomes.....	11
Conclusions.....	12
References.....	13
Appendices.....	14

Introduction

Cervical cancer is theoretically completely preventable by effective screening using cervical cytology methods (the Pap test). The process of preparing and interpreting Pap tests remains one of the last high-volume manual processes in the clinical laboratory. Recent technological advances in specimen preparation and computerized primary screening make automated approaches to cervical cancer screening possible. In addition, advances in information technology have facilitated the Internet transmission and archival storage of digital images and other clinical information. The combination of automated preparation and screening of cervical cytology specimens, with Internet transmission of selected images, and remote interpretation and reporting of results has not been previously attempted.

This project develops a highly automated cervical cytology screening system, a software interface capable of transmitting and presenting images to remote reading stations, with facility for immediate results reporting back to the specimen source. Clinical studies utilizing this developed system will be performed to test accuracy and functionality against the current on-site manual screening process. Primary development of the system has been accomplished at the Massachusetts General Hospital (MGH) site and reading stations have been installed at MGH and at Walter Reed Army Medical Center (WRAMC). A phase 1 pilot study has been completed, data has been analyzed and reported. A phase 2 study of 500 prospectively obtained, consented patients is currently awaiting US Army IRB oversight approval. A phase 3 clinical trial is planned but as per US Army oversight IRB advice, the protocol submission for this portion of the study will not come under review until phase 2 is completed and data analyzed. Phase 3 clinical sites are the 121st Army Hospital in Seoul, Korea and MGH. The US Air Force Academy was dropped from Phase 3 as originally planned because their cytopathology laboratory was closed since the last annual report.

Body

The following is a summation of the work completed to the present time based on the project's accepted Statement of Work. Details follow below in an expanded version of the Statement of Work:

Statement of Work

Task 1: Complete hardware, software and network development required for testing of the internet-based cervical cytology screening system

- a) Modify the FocalPoint device to accept, process and analyze ThinPrep specimens - **completed**

- b) Adapt FocalPoint hardware for internet transmission of digital images from ThinPrep and SurePath specimens - **completed**

- c) Adapt commercial software (Wellogic) to permit rapid and secure transmission of digital images to remote review stations - **completed**

- d) Procure and install remote microscopy stations (2) - **completed**

- e) Adapt commercial software/hardware (Wellogic) to allow secure, automated reporting of cervical cancer screening results - **completed**

- f) Adapt commercial software (Wellogic) to integrate screening results reporting with medical decision support system - **Phase 1 completed, pending Phase 2 modifications.**

- g) perform initial testing of integrated hardware/software/network - **completed**

Task 2: Develop morphology and terminology for digital images and perform pilot clinical trial

- a) Develop a set of learning cases with known diagnostic outcome - **pilot set of 200 cases completed (100 SurePath, 100 ThinPrep).**
- b) Develop morphologic criteria for accuracy of interpretation - **Phase 1 completed, pending Phase 2 modifications.**
- c) Develop reporting terminology appropriate for case management - **Phase 1 completed, pending Phase 2 modifications.**
- d) Develop medical decision support algorithms - **Phase 1 completed, pending Phase 2 modifications.**
- e) Perform pilot trial using a set of 500 unknown specimens to identify preliminary system performance characteristics - **This is a Phase 2 task. Local IRB approval for Phase 2 was received for MGH on 8/26/2004 (reapproved on 8/16/2005), and tentative approval following second level review for WRAMC on 4/26/2005. US Army oversight IRB (Office of Research Protections) protocol review was submitted on 5/16/2005. Request for revisions from ORP was received on 1/3/2006, and a resubmission was returned on 2/20/2006. At the time of this Annual Report, no response has yet been received from ORP and Phase 2 patient accrual, as such, cannot begin.**
- f) Modify procedures/equipment based on pilot trial results - **Phase 1 completed, pending Phase 2 modifications.**
- g) Develop training methods/materials for clinical practice - **Phase 1 completed, pending Phase 2 modifications.**

Task 3: Complete large, prospective clinical trial of the performance of the internet-based system compared to conventional on-site screening.

- a) Develop and receive approval for clinical trial protocol and consent forms - **As per ORP, Phase 3 protocols cannot be submitted for IRB review until Phase 2 is completed and data analyzed.**
- b) Install equipment at selected sites - **future**
- c) Train clinical personnel participating at selected sites - **future**
- d) Conduct the clinical trial - **future**
- e) Perform trial data analysis - **future**
- f) Prepare report of trial with implementation recommendation - **future**

Expanded Discussion

A) Phase 1 of the project has been completed. This Phase included:

- 1) development of hardware, software, and interfaces between computerized scanning device and Internet-linked servers and reading stations.
- 2) development of a 200 case test set of slides with known reference diagnosis (100 SurePath and 100 ThinPrep slides)
- 3) analysis of the test set on the prototype system with interpretation by 6 individuals (3 cytotechnologists, 3 pathologists)
- 4) data analysis
- 5) reporting of the data in 3 abstracts submitted to the US-Canadian Academy of Pathology Annual meeting (February 2006) - see attached
- 6) development of training materials to guide and improve performance
- 7) submission of revisions/improvements to software

Comments: Phase 1 showed a successful first feasibility trial of this system. 191 cases were included in the analysis (SP-101, TP-90; 99-NILM, 4-ASC-US, 3-ASC-H, 4-AGC, 63-LSIL, 18-HSIL). ≥ 3 reviewers agreed on the correct general categorization for unsatisfactory/normal in 87%, and for abnormal in 83%. For specific Bethesda interpretation, ≥ 3 reviewers agreed on the correct categorizations as follows: ASC-US - 75%, ASC-H - 100%, AGC - 25%, LSIL - 83%, HSIL - 94%. These results indicate that correct triage of abnormal cases could be performed at a sensitivity very comparable to the manual screening standard. In addition it was noted during the data analysis/training phase, that a substantial number of the "missed" cases had to do with experience of the observers in identifying clues present in the review station images or with institutional "biases," meaning differences in interpretations that could be traced to practice setting differences between MGH and WRAMC.

B) Initiation of Phase 2 of the project has been significantly delayed

- 1) Local IRB approvals were granted for Phase 2 at MGH and WRAMC

- 2) US Army "oversight" IRB (Office of Research Protections - ORP) required review submitted - 5/17/2005.
- 3) ORP requested revisions - 1/3/2006
- 4) Revisions submitted to ORP - 2/20/2006
- 5) No response from ORP at the time of this Annual Report
- 6) Phase 2 cannot be initiated at either the MGH or WRAMC sites without this approval

Comments: This IRB "oversight" process has significantly delayed the project. At the time of ORP submission, the timeline for completion of Phase 2 showed a final date in the Fall of 2005, with Phase 3 initiation before the end of 2005. At present, Phase 2 should be completed within 4 months of an OPR approval of the protocol. The investigators have been frustrated by this long review process of previously locally IRB-approved protocols. Momentum for the project has been lost as well a loss of the functionality of already completed training activities. These training activities will all subsequently have to be repeated. This delay seems unjustified given that virtually all other research protocols require local IRB approval only.

C) Phase 3 changes since the last Annual Report

- 1) US Air Force Academy clinical trial site was eliminated due to a closure of the cytopathology laboratory there - Dr. Emily Miller is no longer involved in the project.
- 2) USAFA has been replaced by MGH for Phase 3 - a military site could not be obtained as USAFA was initially selected due to their use of SurePath, which is not available at any other potential military trial site.
- 3) An assurance has been granted for the 121st Military Hospital by the Tripler jurisdiction that will allow the research of Phase 3 to be performed there. This required a reallocation of \$10,000 of contract funds to cover the cost of this assurance - a request for reallocation of contract funds is underway.
- 4) Dr. Wilbur visited 121st in Seoul to update the investigators there in August of 2005 (this was as part of another trip in Asia).

Key Research Accomplishments

- 1) Modification of FocalPoint System to accommodate ThinPrep Slides
- 2) Development of Wellogic database, reading station, and report generation software package
- 3) Interfacing of the FocalPoint and Wellogic systems
- 4) IRB submissions
 - Phase 1 at MGH and WRAMC approved
 - Phases 2 at MGH approved
 - Phase 2 at WRAMC tentative approval, awaiting ORP review
 - Phase 2 at ORP pending
 - Phase 3 at 121st WRAMC, and MGH not yet submitted
- 5) Learning test slide set development completed
- 6) Equipment installation at MGH and WRAMC sites completed
- 7) Development of morphology criteria, terminology, and clinical algorithms completed for Phase 1
- 8) Pilot test running of 200 slide test set with the system completed
- 9) Data analysis of the 200 slide set run on the system completed
- 10) Three abstracts presented at the USCAP Annual meeting (Atlanta, Feb 2006)

Reportable Outcomes

- 1) First publication accepted in Cancer Cytopathology - see attachments

- 2) Phase 1 data analysis completed - 3 abstracts presented at USCAP Annual meeting - Feb 2006 - see attachments.

Conclusions

- 1) System development is progressing satisfactorily.

- 2) Data analysis of Phase 1 shows substantial feasibility of the System for appropriate triage of cervical cytology samples.

- 2) IRB (ORP) issues are significantly delaying progress.

- 3) Morphology training sets have been developed for use in ongoing training activities.

- 4) System testing with consented patient cases is slated to begin within one month after ORP approves the Phase 2 protocol.

- 5) System installation at phase 3 clinical sites is postponed as per the requirements of ORP.

References

1) Cancer Cytopathology Publication

Eichhorn JH, Gelfand JA, Brauns TA, Crothers B, Wilbur DC. A Novel Automated Screening and Interpretation Process for Cervical Cytology Using Internet Transmission of Low Resolution Images: A Pilot Study. Cancer (Cancer Cytopathol) 2005;105:199-206.

2) Abstracts presented at USCAP Annual Meeting 2006

Nanji SS, Eichhorn JH, Crothers BA, Wilbur DC. Determining the effective cervical cytology triage point using an automated Internet-based telecytology system with hierarchical image review. Mod Pathol 2006;19;67A.

Nagle JA, Eichhorn JH, Crothers BA, Wilbur DC. Determination of Specimen Adequacy and the Presence of Transformation Zone Component Using an Automated Internet-Based Telecytology System for Cervical Cytology. Mod Pathol 2006;19;67A

Eichhorn JH, Buckner L, Buckner S, Beech DP, Crothers BA, Wilbur DC. Internet-Based Gynecologic Telecytology with Remote Automated Image Selection: Results of a First-Phase Development Trial. Mod Pathol 2006;19;56A.

3) Case data report form for interpretation and criteria development

attached

4) Wellogic software computer review/reporting station screens

attached

HOME | PATIENTS | MESSAGES | REPORTS | LIBRARY | RULES | SETTINGS | HELP

W012, A - 28 yrs, F
 PATIENT ID: AW 530093641888 DOB: 01/01/1978

Beech, David | Workspace: CLINICAL | LOGOUT

CHART DETAIL

View by Event: TYPE

Consultation

Cervical Cancer Consult...

Cervical Cancer Consultation Image(s)

Please Choose... ESCALATE COMPLETE

Details | 0 Message(s) OPEN FOR REVIEW:

Show: FOUR IMAGES

HOME | PATIENTS | MESSAGES | REPORTS | LIBRARY | RULES | SETTINGS | HELP

W015, A - 30 yrs, F
 PATIENT ID: AW 530094137278 DOB: 01/01/1976

Beech, David | Workspace: CLINICAL | LOGOUT

CHART DETAIL

View by Event: TYPE

Consultation

Cervical Cancer Consult...

Cervical Cancer Consultation Image(s)

Please Choose... ESCALATE COMPLETE

Details | 0 Message(s) OPEN FOR REVIEW:

Provider: Wilbur
 Bar Code #: 99-0001554-1
 Quintile: 2
 Adequate Squamous Cells: Detected
 Endocervical: Detected
 Adequacy assessment: <50%

Collected: 3/14/05
 Received: 3/30/05
 Specimen: Thin-Prep

Show: ONE IMAGE

Title: Determination of Specimen Adequacy and the Presence of Transformation Zone Component Using an Automated Internet-Based Telecytology System for Cervical Cytology

J A Nagle¹, J H Eichhorn¹, B A Crothers² and D C Wilbur¹. ¹Department of Pathology, Massachusetts General Hospital, Boston, MA, United States and ²Department of Pathology, Walter Reed Army Medical Center, Washington, DC, United States.

Background: In order to implement a newly designed cervical screening system that allows review of only a portion of the slide surface it is necessary to determine if adequacy assessments can be reliably made. We tested our ability to do so using a computerized location guided screening process in which images for review were device-captured and transmitted to review stations via the Internet.

Design: Liquid-based cervical cytology slides (ThinPrep (TP)(Cytoc) and SurePath (SP)(Tripath)) were scanned using an automated screening device (FocalPoint (Tripath)). Up to 30 low resolution black and white images were acquired from each case and transmitted via the Internet to custom reading stations. Two independent reviewers determined if each case was satisfactory or unsatisfactory based on squamous cellularity assessments. The presence of transformation zone components (TZC) (endocervical or metaplastic cells) was also noted as well as the first FOV (arranged hierarchically) in which they appeared. Results were compiled and compared between reviewers, and against device generated adequacy assessments.

Result: In total 189 slides (90 TP, 99 SP) were analyzed. In 3 slides with reference interpretations of unsatisfactory, 2 were labeled as such, and 1 was labeled as satisfactory by both reviewers. Of 378 TZC assessments (2/slide), 316(84%) indicated that TZC was present within the first 10 FOV's, 29(8%) in the second 10 FOV's, 11(3%) in the third 10 FOV's, and 22(6%) showed no TZC present. Reviewers agreed on exact FOV number containing TZC in 44% of cases. In 21% of cases, TZC was noted by both reviewers in the first FOV examined. The screening device indicated lack of endocervical component in 19 cases - in 11(58%) of these cases both reviewers noted the presence of TZC (4 with metaplastic cells only).

Conclusion: The presence of TZC was consistently identified by reviewers primarily within the first, first 10, or 20 FOV's of the review process (44, 84, 92%, respectively). Reviewers agreed on the exact FOV of first evidence of TZC in nearly ½ of cases. Accurate identification of unsatisfactory cases (2/3, 67%) may require additional training and experience. Device assessments of endocervical component were judged to be incorrect in greater than 1/3 of cases. The data suggest that this system of limited FOV review may be acceptable for assessments of specimen adequacy.

Title: Determining the effective cervical cytology triage point using an automated Internet-based telecytology system with hierarchical image review

S S Nanji¹, J H Eichhorn¹, B A Crothers² and D C Wilbur¹. ¹Department of Cytopathology, Massachusetts General Hospital, Boston, MA, United States and ²Department of Pathology, Walter Reed Army Medical Center, Washington, DC, United States.

Background: Despite the effectiveness of cytological screening, cervical cancer remains one of the most common cancers among women worldwide, primarily due to lack of access to screening programs. Utilization of automated location-guided screening with Internet transmission and remote interpretation may be an effective solution to optimize productivity. The present goal of the study is to determine the number of images that need to be reviewed in order to provide effective specimen triage.

Design: Liquid-based cervical cytology slides (ThinPrep (Cytec) and SurePath (TriPath)) were scanned using an automated screening device (FocalPoint, TriPath). Up to thirty low-resolution (black and white) fields of view (FOV) from each slide were transmitted to reading stations. The images were hierarchically arranged - the first image had highest probability of abnormality and so forth. FOVs were interpreted independently by 7 cytologists (3 cytotechnologists and 4 physicians) and general (NILM/abnormal) and specific (Bethesda 2001) interpretations rendered. For each abnormal interpretation, the first FOV containing abnormality was recorded, as this would represent the "triage point." Data was compiled to show the percentage of triage points in the 1st 10, 2nd 10 and 3rd 10 FOV's.

Result: 92 slides (92x7=644 interpretations) with general reference interpretations of abnormal (5 AGC, 3 ASC-H, 4 ASC-US, 62 LSIL, 18 HSIL) were analyzed (95 NILM cases were scanned and reviewed but are irrelevant in this analysis). Of the 644 interpretations, 75% were identified as abnormal and of these, the triage point was present in the first 10 FOV's in 93% of cases (ASC-US (100%), ASC-H (82%), HSIL (93%), LSIL (92%), AGC (94%)). In the first 20 FOV's the triage point was noted in 100% of ASC-US, ASC-H, HSIL and AGC cases and in 99% of LSIL cases. The first FOV was the triage point in 50% of the abnormal cases.

Conclusion: The system was found to be very robust in the identification of abnormality in the first FOV (50%), the 10 highest (93%) and the 20 highest (99%) FOV's. These findings suggest that using this system, examination of 20 images may be sufficient as a screening (triage) tool. Implementation of this process has the potential to improve efficiency and accessibility of screening to underserved populations via highly productive and centralized interpretation.

Title: Internet-Based Gynecologic Telecytology with Remote Automated Image Selection: Results of a First-Phase Development Trial

Authors: ¹JH Eichhorn, ²L Buckner, ³S Buckner, ¹DP Beech, ²BA Crothers, ¹DC Wilbur. ¹Department of Cytopathology, Massachusetts General Hospital, Boston, MA, United States, ²Department of Pathology, Walter Reed Army Medical Center, Washington, DC, United States, and ³Department of Pathology, National Naval Medical Center, Bethesda, MD, United States.

Background: The combination of liquid-based slide preparation, automated screening with field of view (FOV) selection, image capture, internet transmission, and remote interpretation of selected FOV's is a means to allow centralization of gynecologic cytology expertise and improve productivity.

Design: A retrospective set of gynecologic cytology slides (ThinPrep (TP)(Cytoc) and SurePath (SP)(TriPath)) with reference interpretations was run on an automated screening device (FocalPoint (TriPath)) that selects FOV's based on a hierarchical probability of abnormality being present. An interface between the device and a remote server with custom image review software (Consult (Wellogic)) was developed. FOV's were reviewed by 6 cytologists (3 CT, 3 MD) and general categorizations (unsatisfactory/normal, abnormal) and specific interpretations (Bethesda 2001) were rendered. No training prior to FOV review was performed. For this phase, results were tabulated based on correct general categorizations versus the reference interpretation for each case and reported based on how many reviewers achieved the correct interpretation.

Results: 191 cases were included in the analysis (SP-101, TP-90; 99-NILM, 4-ASC-US, 3-ASC-H, 4-AGC, 63-LSIL, 18-HSIL). ≥ 3 reviewers agreed on the correct general categorization for unsatisfactory/normal in 87%, and for abnormal in 83%. For specific Bethesda interpretation, ≥ 3 reviewers agreed on the correct categorizations were as follows: ASC-US - 75%, ASC-H - 100%, AGC - 25%, LSIL - 83%, HSIL - 94%.

Conclusion: Based on comparisons to prior studies of manual screening of slides, the results of this study, in which no prior training was given to reviewers, suggest that this procedure has comparable sensitivity and specificity and would be an effective initial triage to further evaluation such as full manual screening, HPV testing, or colposcopy.

A Novel Automated Screening and Interpretation Process for Cervical Cytology Using the Internet Transmission of Low-Resolution Images

A Feasibility Study

John H. Eichhorn, M.D.^{1,2}
 Timothy A. Brauns, M.B.A.³
 Jeffrey A. Gelfand, M.D.^{3,4,5}
 Barbara A. Crothers, D.O.⁶
 David C. Wilbur, M.D.^{1,2}

¹ Cytopathology Unit, Massachusetts General Hospital, Boston, Massachusetts.

² Department of Pathology, Harvard Medical School, Boston, Massachusetts.

³ Center for Integration of Medicine and Innovative Technology, Partners Health Care, Boston, Massachusetts.

⁴ Infectious Disease Unit, Massachusetts General Hospital, Boston, Massachusetts.

⁵ Department of Medicine, Harvard Medical School, Boston, Massachusetts.

⁶ Department of Pathology, Walter Reed Army Medical Center, Uniformed Services School of Medicine, Washington, DC.

Presented in part at an annual meeting of the United States and Canadian Academy of Pathology, Washington, D.C., March 22–28, 2003.

The preliminary work was supported in part by the Edmund C. Lynch, Jr., Cancer Fund. The ongoing portions of this project are supported by the U.S. Army Medical Research and Materiel Command under contract no. W81XWH-04-C-0083.

Address for reprints: David C. Wilbur, M.D., Cytopathology Unit, Warren 120, Massachusetts General Hospital, 55 Blossom Street, Boston, MA 02114; Fax: (617) 724-6564; E-mail, dwilbur@partners.org

The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Department of the Army or the Department of Defense.

Dr. Wilbur is a member of the speakers' bureau of TriPath Imaging, Inc.

Received December 8, 2004; revision received March 8, 2005; accepted March 16, 2005.

BACKGROUND. Transmission over the Internet of low-resolution images acquired by automated screening of cervical cytology specimens has the potential to provide remote interpretation and, hence, centralization of a cytology workforce.

METHODS. Liquid-based cervical cytology slides were scanned using the FocalPoint® System. Ten black-and-white images that had the greatest probability of containing abnormality were acquired from each of 32 reference slides (16 negative samples, 3 samples of atypical squamous cells of uncertain significance, 5 samples of low-grade squamous intraepithelial lesions [LSIL], 5 samples of high-grade squamous intraepithelial lesions [HSIL], 1 adenocarcinoma in situ sample, and 2 carcinoma samples) and were transmitted as e-mail attachments in JPEG format to remote reading stations. The slides were interpreted independently by two pathologists and were assigned to either of two groups: 1) suspicious for \geq HSIL or 2) \leq LSIL. The interpretations were compared with the reference diagnoses. The specimens were then randomized, and the image sets were redistributed to the pathologists for another round of interpretation and scoring.

RESULTS. The initial and subsequent trials yielded similar results. Pooling the interpretations of the two pathologists, the concordance rate between reference and assigned diagnostic groups for each of the two trials was 84%, the false-positive rate was 8.3%, and the false-negative rate was 37.5%. Review of the discrepant slides revealed subtle cellular changes that may be utilized to reduce errors and, with training, to optimize sensitivity and specificity.

CONCLUSIONS. This procedure showed promise for allowing remote interpretation of device-selected images. The procedure may represent an effective way to centralize cervical cytology services and to allow the provision of services to previously unscreened populations that lack an effective cytology infrastructure. *Cancer (Cancer Cytopathol)* 2005;105:199–206. © 2005 American Cancer Society.

KEYWORDS: cervical cytology, automation, Internet, telepathology, telecytology, FocalPoint.

The FocalPoint® screening system for cervical cytology (TriPath Imaging, Inc., Burlington, NC) is approved by the U.S. Food and Drug Administration (FDA) for use in the primary screening of conventional and SurePath® (TriPath) liquid-based slides. Studies have shown that this automated system efficiently and reliably identifies more abnormal specimens compared with manual screening alone and that it can be used in the triage of slides to “no manual review” or “complete manual review.”^{1–3} Inherent to the operation of the FocalPoint device is its ability to identify fields of view (FOV) on each slide that contain individual cells or cell groupings that have the greatest probability of being abnormal (Fig. 1).⁴ In doing so, the



FIGURE 1. This image of the Focal-Point® GS reviewing station shows black-and-white image of the field of view captured in the lower left part of the instrument (box).

device captures low-resolution JPEG images designed to allow accurate localization to observers during subsequent manual microscopic review. This image-capture capability also makes possible the rapid transmission of such images over the Internet for interpretation at remote sites. Although digital transmission and interpretation of cytology images has been investigated by others,⁵⁻⁷ the combination of automated screening and Internet-based telecytology remote interpretation has not been explored. Because it bypasses on-site manual review and the large-scale transportation of perishable glass slides, such a system could lead to substantial centralization and optimization of cytology screening resources. The use of these complimentary technologies also may allow the introduction of cervical carcinoma screening programs to countries where none currently exist.

To implement such a program, it will be necessary to show that accurate and reliable interpretations can be made using FOV review alone on a computer monitor, without the benefit of a full manual microscopy screening. It has been shown that microscopic FOV review with triage to full manual screening when abnormality initially is identified has the potential to be very accurate⁴; however, the operating characteristics of the review of only low-resolution images, by its very nature, may have a lower sensitivity for the detection of disease. In addition, lack of the benefit of full screen

or “through the microscope” review may be expected to provide less information on specimen adequacy, reactive changes, organisms, and subtle low-level abnormalities, such as atypical squamous cells. For an intended application of high-grade specimen triage, however, it may be sufficient to identify those lesions that require prompt intervention. For the current study, we used this approach with a triage point at the “suspectious for high-grade squamous intraepithelial lesion (HSIL) or a more serious process” cut-off level to assess the feasibility of remote interpretation of digital images obtained by automated screening and transmitted over the Internet.

MATERIALS AND METHODS

The study population consisted of a set of 32 reference SurePath® liquid-based cervical cytology slides from the files of the Massachusetts General Hospital with known diagnoses confirmed by 1 of the study pathologists (D.C.W.) approximately 2 months prior to the study initiation. The interpretations were made using the criteria of the 2001 Bethesda system.⁸ The set included 2 samples of carcinoma, 1 adenocarcinoma in situ (AIS), 5 samples of HSIL, 5 samples of low-grade squamous intraepithelial lesion (LSIL), 3 samples of atypical squamous cells of undetermined significance (ASCUS), and 16 samples that were negative for intraepithelial lesion or malignancy (NILM). The

TABLE 1
Diagnoses of 32 Reference Specimens

Diagnosis	No. of specimens
Carcinoma	2
Squamous	1
NOS	1
AIS	1
HSIL	5
LSIL	5
ASCUS	3
NILM	16
Candida species	3
Atrophy	2
Trichomonas vaginalis	1
Bacterial vaginosis	1
NOS	9

NOS: not otherwise specified; AIS: adenocarcinoma in situ; HSIL: high-grade squamous intraepithelial lesion; LSIL: low-grade squamous intraepithelial lesion; ASCUS: atypical squamous cells of undetermined significance; NILM: negative for intraepithelial lesion or malignancy.

slides that were diagnosed as HSIL, AIS, and carcinoma were confirmed on biopsy. Among the NILM slides, four showed *Candida* species or *Trichomonas vaginalis* organisms, two were atrophic, and one showed bacterial vaginosis (Table 1).

All patient-identifier information was delinked from the specimens, and the slides were scanned at a remote site (TriPath, Redmond, WA) on the FocalPoint® GS System, and a set of the 10 highest scoring FOV images were captured in the JPEG compressed format (image size, 12–16 KB each). The FocalPoint GS System captures only black-and-white images, and each image corresponds closely to a microscopic $\times 200$ magnification when viewed on a monitor. Sets of 10 JPEG images per slide were bundled as e-mail attachments and transmitted over the Internet to reading stations at the Massachusetts General Hospital. Each “reading station” was a Partners network computer (Microsoft Windows 2000 Professional; Microsoft, Redmond, WA), and the images were opened using Microsoft Outlook 2000 and Internet Explorer software and were reviewed on standard, 17-inch desk top monitors. An example of an e-mail received at a review site is shown in Figure 2.

Each of the 32 study slides was interpreted independently by 2 board-certified cytopathologists with a combined 30 years of postgraduate experience (J.H.E. and D.C.W.) who were masked to the reference diagnoses. During the review, notations were made for learning purposes regarding why a given diagnosis was favored. Slides were diagnosed as specifically as possible and then were assigned in a binary triage process to either of two categories: 1) NILM or “low-grade” (normal findings, reactive changes, ASCUS, or

LSIL) or 2) suspicious for HSIL or more a serious process (atypical squamous cells, cannot exclude HSIL, HSIL, carcinoma, or suspicious for any of these diagnoses). A binary classification was used, because one objective of the study was to determine whether the combination of automated screening and remote transmission of low-resolution images could be utilized to triage patients with a cut-off point of biologic significance for true precancerous high-grade lesions versus negative results and those with low-grade or possibly transient human papillomavirus (HPV) infections. After both pathologists had recorded their categorizations of each slide, the “test” interpretations were compared with the reference diagnoses. With the knowledge of this comparison between “test” and “reference” categorizations, each of the pathologists had an opportunity again to view the bundled digital images for any given case or to review any notations that they had recorded at the time of the initial interpretation. Glass slides were not studied together with the black-and white images. After a 2-month hiatus, the 32 slides were “shuffled” (randomized), and their bundled images were retransmitted to the 2 reviewers for a second round of interpretations to assess whether their ability to categorize the lesions was improved by their knowledge of the types of errors that had been made (or their threshold bias) in the first round. The results of the first and second rounds were compared for each of the two pathologists. From the pooled interpretations of the two pathologists, rates of concordance ($(\text{true-positive} + \text{true-negative})/\text{total}$), false positivity ($\text{false-positive}/\text{total negative}$), false negativity ($\text{false-negative}/\text{total positive}$), sensitivity ($\text{true-positive}/[\text{true-positive} + \text{false-negative}]$), and specificity ($\text{true-negative}/[\text{true-negative} + \text{false-positive}]$) were calculated for both trials and were compared.

RESULTS

Figures 3–6 show representative FOV images from selected slides that were interpreted as NILM, LSIL, HSIL and carcinoma, respectively. Compared with standard light microscopy, the images displayed on the computer monitor showed poorer resolution for a given degree of magnification, and their plane of focus could not be manipulated. These characteristics resulted in poorer definition of the chromatin texture in small nuclei and the possible lack of detection or appreciation of objects of very small dimensions, such as some microorganisms. From the standpoint of the binary categorization used in the current study, these limitations were most problematic when considering the distinction between benign squamous metaplasia and the cells of HSIL (most notably, those of moderate

From: [Identifying information removed]
Sent: Sunday, July 28, 2002 5:28 PM
To: Wilbur, David
Subject: slide 28



[Identifying information removed]

<<98-0272781-0_1_36034_10614.jpg>> <<98-0272781-0_2_34063_8643.jpg>>
<<98-0272781-0_3_25052_15683.jpg>> <<98-0272781-0_4_32936_9769.jpg>>
<<98-0272781-0_5_25615_13712.jpg>> <<98-0272781-0_6_35471_12304.jpg>>
<<98-0272781-0_7_34344_13430.jpg>> <<98-0272781-0_8_35752_17091.jpg>>
<<98-0272781-0_9_27868_16528.jpg>> <<98-0272781-0_10_30402_12585.jpg>>

FIGURE 2. Example of an E-mail received at a review site, with 10 image icons.

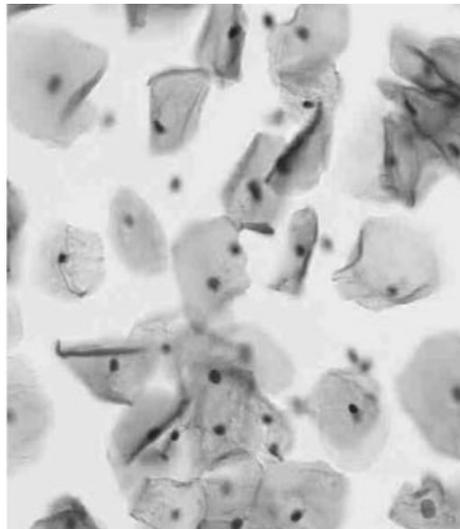
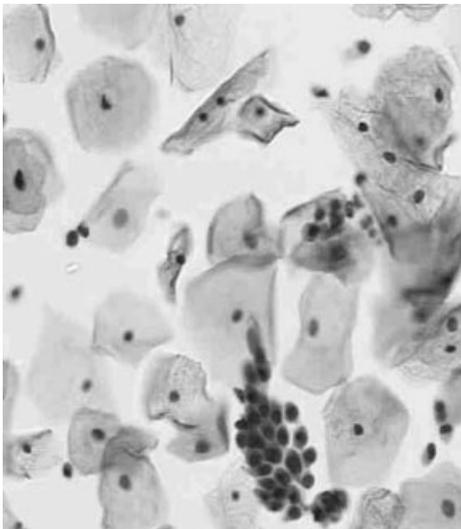


FIGURE 3. These field-of-view images show a specimen that was negative for intraepithelial lesion or malignancy.

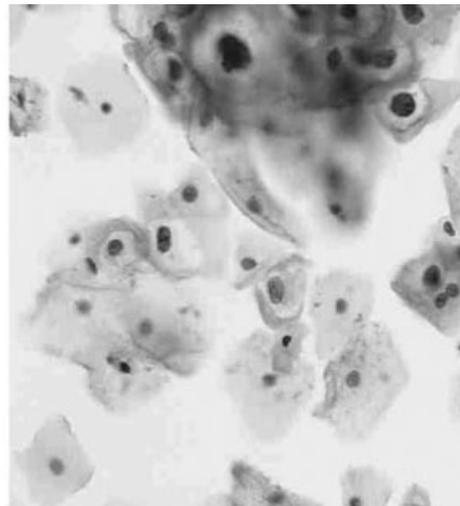
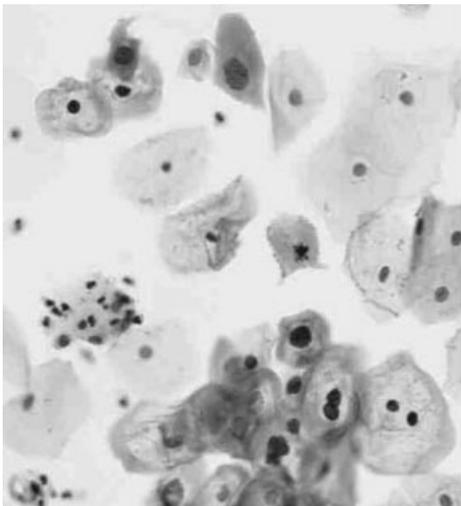


FIGURE 4. Field-of-view images from a specimen of low-grade squamous intraepithelial lesion. Note the aggregates of cells with large, irregularly shaped hyperchromatic nuclei and perinuclear halos.

FIGURE 5. Field-of-view images from a specimen of high-grade squamous intraepithelial lesion. Note the individual, small cells with increased nucleus-to-cytoplasm ratio and irregular hyperchromatic nuclei.

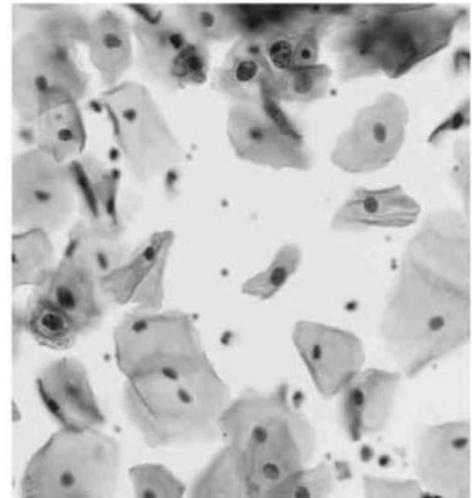
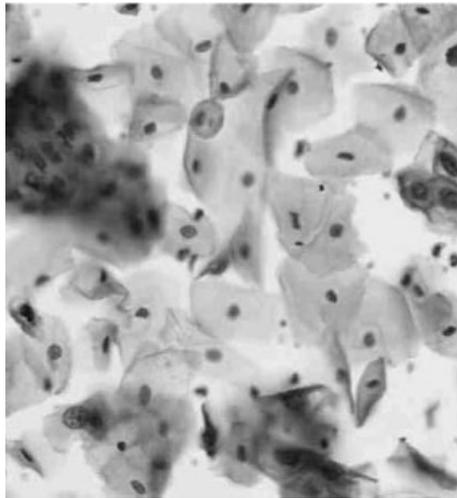
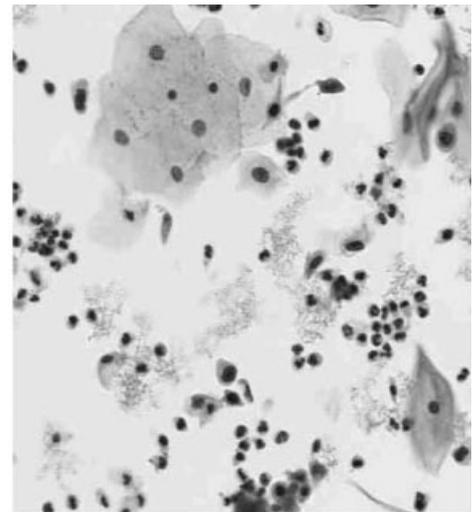
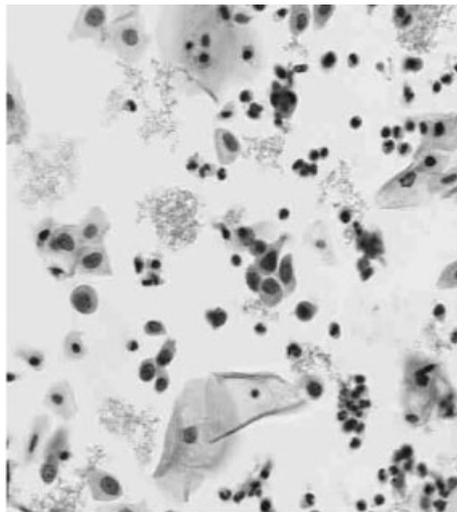


FIGURE 6. Field-of-view images from a specimen of carcinoma; note that tumor diathesis, which is a useful clue.



squamous dysplasia), particularly when just a few single cells were available in the FOVs for review. Low-image resolution and the lack of a full color spectrum also made the evaluation of subtle cytoplasmic qualities more challenging, particularly in the smaller cells of a given population.

We found that the most useful clues in the identification of dysplasia and its categorization as low-grade or high-grade were the nucleus-to-cytoplasm ratio, the nuclear shape, the degree of hyperchromasia, the presence or absence of HPV-associated cytopathic perinuclear halos, and the evaluation of aggregates of potentially abnormal cells for their cellular relation (three-dimensionality, cellular distribution and overlap, orientation, and the degree of variation in size and shape). These features were not affected (or were affected only to a limited degree) by the decreased resolution, fixed plane of focus, and loss of color. In addition, an increased emphasis on back-

ground features (brisk inflammation or tumor diathesis) could be used in some slides to compensate for the loss of fine detail. On average, each pathologist spent < 5 minutes per slide to review the images and render an interpretation. No attempt was made to do a formal time analysis in this pilot study.

The reference diagnostic groups and those assigned by each of the 2 pathologists in the initial and subsequent trials are compared in Table 2. Both pathologists had rates of concordance between test and reference categories that exceeded 81% in each of the 2 trials. One pathologist tended to under diagnose specimens that were suspicious for high-grade lesions in the first trial, with improved sensitivity in the second trial. The other pathologist tended to over diagnose specimens that were suspicious for high-grade lesions in the first trial and showed improved specificity in the second trial.

Table 3 combines the results of the two pathologists

TABLE 2
Comparison of Test and Reference Diagnostic Groups:
First and Second Trials of Each Pathologist

Category	Reference diagnostic category ^a	
	Neg, ASCUS, or LSIL	≥ HSIL
Pathologist 1 interpretation		
Negative or LSIL	24 (22)	6 (2)
At least suspicious for HSIL	0 (2)	2 (6)
Pathologist 2 interpretation		
Negative or LSIL	20 (22)	0 (4)
At least suspicious for HSIL	4 (2)	8 (4)

Neg: negative for dysplastic atypia; ASCUS: atypical squamous cells of uncertain significance; LSIL: low-grade squamous intraepithelial lesion; HSIL: high-grade squamous intraepithelial lesion.

^a Expressed as the number of specimens in the initial (subsequent) reviews.

TABLE 3
Comparison of Test and Reference Diagnostic Groups:
Composite Results of Both Pathologists

Category	Reference diagnostic category ^a	
	Neg, ASCUS, or LSIL	≥ HSIL
Assigned diagnostic category, both pathologists		
Negative or LSIL	44 (44)	6 (6)
At least suspicious for HSIL	4 (4)	10 (10)

Neg: negative for dysplastic atypia; ASCUS: atypical squamous cells of uncertain significance; LSIL: low-grade squamous intraepithelial lesion; HSIL: high-grade squamous intraepithelial lesion.

^a Expressed as the number of specimens in the initial (subsequent) reviews.

in each of the two trials, respectively, and compares the reference diagnostic groups to the interpretations of the pathologists. From the latter tabulation, the concordance rate between the test and reference interpretations was 54 of 64 interpretations (84%), the false-positive rate was 4 of 48 interpretations (8.3%), and the false-negative rate was 6 of 16 interpretations (37.5%) for both trials. The rates of sensitivity and specificity were 63% and 92%, respectively, for both trials.

DISCUSSION

Worldwide, it has been estimated that 471,000 women per year are diagnosed with cervical carcinoma, and 233,400 women per year die from the disease.⁹ Overall, cervical carcinoma remains a leading cause of death of young women in countries that lack any screening program.¹⁰ Conversely, the incidence and mortality figures have decreased substantially in countries that have instituted population-based screening programs.¹¹ In the U.S., for instance, it has been estimated that the incidence of cervical carcinoma has decreased by 50% in 50 years¹² and that the mortality

has decreased by 70% since widespread screening was instituted in the 1950s.¹³ The incidence and mortality rates for cervical carcinoma in the U.S. recently were reported as 13,000 and 4000 women per year, respectively¹⁴; however, again, a majority of these deaths occur among women in subpopulations that largely are under-screened.¹⁵ Women in many developed countries benefit not only from periodic sampling of cells from the cervix and an adequate pool of trained specialists (cytotechnologists and cytopathologists) to screen and interpret these specimens, but also from new technologies, such as liquid-based specimen preparations and computerized automated screening. From the perspective of healthcare delivery systems in regions of the world that lack any cytology infrastructure, however, the most difficult and costly hurdle to the implementation of screening is the establishment of a trained cytology workforce. Even in the U.S., deficiencies in this workforce are anticipated,¹⁶ particularly in certain settings, such as the military health care system.

It is the hypothesis of the current study that Internet transmission of images derived from automated screening devices for interpretation at remote locations has the potential to provide an effective screening program for countries that lack a trained cytology infrastructure and to provide a more efficient, economic, and centralized way of triaging patient samples. The results of this pilot study provide preliminary feasibility support for this hypothesis. The system of image selection, digital storage, electronic transmission, and remote interpretation worked effectively. Diagnoses were rendered on small, finite sets of images that had been selected and encoded by an automated screening device, sent over the Internet from a remote location, and interpreted at a centralized site. After independent masked reviews by two pathologists in two separate testing events, case discrimination at the "suspicious for ≥ HSIL" level of triage showed a false-negative rate of 37.5%, a false-positive rate of 8.3%, and an overall concordance rate of 84% for each of 2 trials. Stated differently, the specificity was 92%, and the sensitivity was 63%. Although the relatively high false-negative rate compares unfavorably with the current standard of < 10% for ≥ HSIL, the authors observed that experience and learning obtained between the first and second phases of this relatively small trial had the potential to improve the sensitivity in ongoing use. In addition, if the desired objective is the identification of as many patients with high-grade lesions as possible in populations with limited screening, then this process offers an advantage over no screening at all. Ultimately, the sensitivity of the screening process may be augmented when coupled

with developing molecular assays, such as those for HPV-associated antigens and proliferation markers, which also could be performed and interpreted remotely over the same web-based network.¹⁷

However, the current data suggest that interpretative accuracy with this system can be improved through a reassessment of the diagnostic criteria and training in the use of them. It may be conjectured that the interpretative criteria that are most useful in this new modality are different from those that are relied on in manual microscopy. Image magnification could not be increased, and the lack of fine focusing imposed limits on the visualization of fine chromatin detail and group morphology. In this context, a more subtle assessment of background material and cell-cell relations received greater emphasis. Both pathologists altered their performance in the second trial after a review of their mistakes in the first, but they did so in opposite measures: "Pathologist 1" improved his specificity from the first trial to the second, and "Pathologist 2" improved his sensitivity. Moreover, the utilization of alternative triage cut-off points not only may alter the sensitivity and specificity rates but also may be tailored to different public health care objectives.

The system we describe differs from manual microscopic screening in at least five respects that may decrease its comparative performance: 1) a finite set of images (FOVs) is selected by the automated screening device, 2) the image resolution is lower than that of microscopy, 3) the plane of focus is fixed, 4) the magnification is predetermined, and 5) the images are black and white. A factor to be analyzed in future studies is the optimum number of FOVs that need to be examined. Will increasing the number of FOVs presented improve the accuracy? The size and resolution of the presented FOVs also can be altered. Will changing of these parameters improve the accuracy? Finally, the ability to image in color and the creation of scanned "virtual" FOVs with the ability to focus through planes need to be investigated as means of improving the overall accuracy without compromising the novel issues of Internet transmission, remote interpretation, and centralization capability.

This pilot study had obvious limitations. The study population was a small teaching set with known diagnoses, which is not equivalent to a population of patients examined prospectively. Only specimens that were prepared using the SurePath method were studied, primarily because the FocalPoint screening equipment is calibrated and FDA-approved only for their use; accordingly, the results obtained from conventionally smeared or ThinPrep[®] liquid-based slides (Cytoc Corporation, Boxborough, MA) may differ. Finally,

the pathologists, by necessity, were modifying their habitual criteria as the study progressed. Although their performance improved after a review of their mistakes in the first trial, it is not known whether this was due to a modification of their criteria or diagnostic thresholds, subconscious (or conscious) compensation for their biases in the initial test, or memory of individual difficult slides or images. Future planned prospective trials will address these issues.

The screening system presented was conceived as a potential solution to a need for population-based cervical carcinoma screening in the many countries of the world that currently lack an infrastructure for cytology screening. Such a system would serve an immediate need in developing countries that have a large "at-risk" population of unscreened women, a commitment to improving public health, some medical care capabilities, and some telecommunications infrastructure, but an absence or limited supply of a trained cytology workforce. Potentially, the number of women residing in such countries greatly exceeds that of the developed countries that have active screening programs in place. The images that were used in the current study required very modest amounts of memory (< 20 KB) and, thus, would not require significant bandwidth for transmission. At comparatively low cost, automated screening devices could be located in these regions, and personnel trained only in their operation and maintenance and in the preparation of liquid-based specimens could tend them. Images derived from such specimens then would be transmitted and interpreted elsewhere; and the diagnoses, triage groups, or both would be returned electronically to a party responsible for the care of the patient.

Another potential application would be for far-flung organizational systems, such as the military cytology service. Such services strive to apply U.S. "standard of care" for military personnel throughout the world, often with a highly mobile and difficult-to-track group of patients and caregivers. A cruder triage of patients into diagnostic groups that do or do not require timely medical intervention, however, may suffice for many of the costly deployment decisions that could arise. Automated screening equipment could be located at existing military medical facilities, sites that nevertheless are too remote for a fixed pool of trained cytotechnologists. The necessity of making rapid deployment decisions and the declining availability of cytology human resources could be used to argue for the use of such a system with a centralized interpretive service and a linked electronic reporting system that automatically would follow the patient and health care provider to any point on the globe.

The scheme described above represents a novel

melding of two complementary technologies—the Internet and computerized, automated screening. This pilot study demonstrated the feasibility of this methodology with the possibility of improvements from further training and refinement of hardware, incorporation of ancillary markers, and investigation of clinically relevant endpoints. At its current level of performance, it appears to offer an advantage over no screening for geographic areas that lack a cytology infrastructure. A much larger and well designed clinical trial will be required, however, before its application to public health problems can be considered. Finally, it has the potential to enhance efficiency and centralization of cytology services greatly for organizations that maintain large and complex cytology infrastructures, such as the U.S. military services or large laboratory organizations.

REFERENCES

1. Richart RM, Nelson AC, Wood T. Clinical test results for automated quality control and primary screening of cervical smears using the NeoPath AutoPap 300 System [abstract]. *Acta Cytol.* 1994;38:805.
2. Stevens MW, Milne AJ, James KA, Brancheau D, Ellison D, Kuan L. Effectiveness of automated cervical cytology re-screening using the AutoPap 300 QC System. *Diagn Cytopathol.* 1997;16:505–512.
3. Wilbur DC, Prey MU, Miller WM, Pawlick GF, Colgan TJ. The AutoPap system for primary screening in cervical cytology. Comparing the results of a prospective, intended-use study with routine manual practice. *Acta Cytol.* 1998;42:214–220.
4. Wilbur DC, Parker EM, Foti JA. Location guided screening of liquid-based cervical cytology specimens: a potential improvement in accuracy and productivity is demonstrated in a preclinical feasibility trial. *Am J Clin Pathol.* 2002;118:399–407.
5. Raab SS, Zaleski MS, Thomas PA, Niemann TH, Isacson C, Jensen CS. Telecytology: diagnostic accuracy in cervical-vaginal smears. *Am J Clin Pathol.* 1996;105:599–603.
6. O'Brien MJ, Takahashi M, Brugal G, et al. Digital imagery/telecytology. International Academy of Cytology Task Force summary. Diagnostic cytology towards the 21st century: an international expert conference and tutorial. *Acta Cytol.* 1998;42:148–164.
7. Kaplan KJ, Burgess JR, Sandberg GD, Myers CP, Bigott TR, Greenspan RB. Use of robotic telepathology for frozen-section diagnosis: a retrospective trial of a telepathology system for intraoperative consultation. *Mod Pathol.* 2002;15:1197–1204.
8. Solomon D, Nayar R. The Bethesda system for reporting cervical cytology. Definitions, criteria and explanatory notes. 2nd edition. New York: Springer-Verlag, 2004.
9. Parkin DM, Bray F, Ferlay J, Pisani P. Estimating the world cancer burden: Globocan 2000. *Int J Cancer.* 2001;94:153–156.
10. Ferlay J, Parkin DM, Pisani P. Globocan 1: cancer incidence and mortality worldwide. IARC CancerBase 3. Lyon: IARC, 1998.
11. Sankaranarayanan R, Budukh AM, Rajkumar R. Effective screening programmes for cervical cancer in low- and middle-income developing countries. *Bull WHO.* 2001;79:954–962.
12. Nanda K, McCrory DC, Myers ER, et al. Accuracy of the Papanicolaou test in screening for and follow-up of cervical cytologic abnormalities: a systematic review. *Ann Intern Med.* 2000;132:810–819.
13. Boring CC, Squires TS, Tong T, Montgomery S. Cancer statistics, 1994. *CA Cancer J Clin.* 1994;44:7–26.
14. Jemal A, Thomas A, Murray T, Thun M. Cancer statistics, 2002. *CA Cancer J Clin.* 2002;52:23–47.
15. Schneider V, Henry MR, Jimenez-Ayala M, Turnbull LS, Wright TC, International Consensus Conference on the Fight Against Cervical Cancer IACTFCIUSA. Cervical cancer screening, screening errors and reporting. *Acta Cytol.* 2001; 45:493–498.
16. Allen KA. ASCT cytotechnologist salary and vacancy report. *ASCT J Cytotechnol.* 2002;5:5–7.
17. Sampson SJ, Parker MR, Hessling JJ, et al. Data comparison of three molecular markers (MCM7, p16, and integrin-B6) for the detection of cervical dysplasia in SurePath™ cytology specimens [abstract]. *Acta Cytol.* 2004;48:658.