Dr. Bibbo’s Presidential Address on Automation in Cytology: Were Her Predictions Right, Wrong, or Somewhere in the Middle?

David C. Wilbur
Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

Abstract
In 1983, Dr. Marluce Bibbo gave the Presidential Address at the Annual Meeting of the American Society of Cytology in Denver, CO, USA. The lecture was entitled “Analytic and Quantitative Cytology,” a field in which Dr. Bibbo was intimately involved. In the presentation, she included a summary of 30 years of work already accomplished, the present state of the art, and musings about issues encountered, potential resolutions, progress that needed to be made, and her perception of how the field needed to evolve in order to become ultimately successful as a clinical service. This commentary looks back 34 years, with observations about Dr. Bibbo’s predictions and how the field of cytology automation did actually evolve in the decades following her address. New challenges are identified and possible paths forward are discussed.

Keywords
Automation · Analytic cytology · Quantitative cytology

1983: Ronald Reagan was President of the USA, the last of the Delorian cars was produced (remember Back to the Future), Sally Ride became the first woman in space on space shuttle Challenger’s second flight, Korean Air Flight 007 was shot down by Russian fighters after straying into Soviet airspace, the USA invaded the Caribbean Island of Grenada to rescue US medical students, catastrophic earthquakes in Turkey and Japan killed thousands, Martin Luther King Day was declared a US national holiday, hundreds of US Marines were killed in bombings in Lebanon, and I was just finishing my third year of pathology residency.

In another historical event (from a cytologic viewpoint) in November of that year, Dr. Marluce Bibbo gave the Presidential Address at the American Society of Cytology (now Cytopathology) in Denver, CO, USA [1]. The topic of her address was “Analytic and Quantitative Cytology” which, in 1983, was a 30-year-old field which was both maturing and at the same time approaching a philosophical junction. Dr. Bibbo, through her work at the University of Chicago with her mentor and associate George Wied, was already a recognized leader in this arena, and her address detailed the relevant background as well as her musings on the future and the way forward.

For readers who might not be so familiar with Dr. Bibbo’s biography, a few “tidbits” are in order. Of course, all cytologists will know her from her textbook Comprehensive Cytopathology [2], now in its fourth edition, which is appropriately, and perhaps even affectionately, referred to by students merely as “Bibbo.” Dr. Bibbo is originally from Brazil where she completed medical school and her
pathology residency in Sao Paulo. Coming to the USA in 1969, she completed a Cytopathology Fellowship at the University of Chicago where she remained on the staff, rising to the rank of full professor. She moved to, and is currently still active at, the Thomas Jefferson University in Philadelphia, where she is the Warren G. Lang Professor of Pathology and Cell Biology. She was Editor-in-Chief of Acta Cytologica from 2004 until 2013 [3].

Automated analysis of cervical cytology specimens as a screening method for cervical neoplasia had been the “holy grail” of a large number of investigations with large amounts of cash spent for the previous 30 or so years. Flow cytometric and image-based instruments had been developed which could recognize very fine differences in cell populations [4, 5]. Anyone training the uninitiated first-time student of cytology will appreciate the subtleties inherent in this endeavor, even when working with a human brain! However, to date, these devices had not achieved usability in a practical sense – their cost was too high and their reliability suspect. Dr. Bibbo does, however, note that “considerable progress has been made” but laments on whether “fully automated cytologic diagnoses ... will ever be achieved” [1].

One might drop the subject right there – money down the rabbit hole – cut your losses – don’t throw good money after bad – all of these are potentially appropriate comments, given the status of the field as described in 1983. But the problems to be addressed still existed:

1. Cervical cancer is preventable by screening (and is therefore desirable to accomplish). True.
2. Screening for cervical cancer by human eyes through the microscope is “tedious and exhausting”, and is fraught with at least a 10% false negative rate for missed cancer or precancer. True.
3. If all eligible women were screened by the current standard of manual human microscopy, the workforce is inadequate. True.
4. Improved Pap test preparation could make the test more accurate (but Dr. Bibbo opines that doing so would ultimately not be cost-effective). See below.
5. The most common cause of a missed cervical cancer is no Pap test at all. True (and No. 3. above becomes a real concern).

The point is, despite the apparent dismal future for cytology automation, there is still a great need for it. So what does Dr. Bibbo set as the “state of the art” in analytic and quantitative cytology in 1983? And how should the field find fertile ground to overcome its inadequacies – she has a plan!

As was mentioned above, in 1983, the field was at a philosophic junction in the road. One of the paths represented what, in retrospect, was an error in practical thinking, i.e., that human observers do not have “planned” error rates. Early investigators were wedded to the principle that automated devices should not have an inherent (known) error rate; they should “be like humans” who never planned to make a mistake. Of course, we all know (but perhaps are not willing to admit [or quantify]) that humans do indeed have false-positive and false-negative rates and that these rates vary, depending on the process of thresholding. For example, some days I am more sensitive, perhaps because yesterday I found that I had missed a case via the retrospective review process. But inherent in this increased sensitivity is poorer specificity because today (in order to be more sensitive) I will have a tendency to overcall. If this sounds like the accuracy calculations we make for instrumentation in the clinical laboratory – well, it is indeed like that. Humans are not perfect and they have operating characteristics that vary just like machines do. However, leading up to 1983, cytologists were not willing to admit such, and hence they were also not willing to allow a device to perform autonomous clinical work with a “known” accuracy of less than 100%. It was thought that the use of such a device would be unethical when compared to the alternative of “no planned inaccuracies” in a human manual examination. The devices that had been developed had inherent (known because of extensive testing) disease detection sensitivities of 90–95%. In contrast, specificity, as Dr. Bibbo points out, was also a huge problem. The identification of what was truly normal was, in actuality, the most difficult task for instrumentation. Abnormal cells were actually much easier for the computers to identify, but if normal events were actually “detected” as abnormal (if the device were “set” to achieve maximal sensitivity), the number of “alarms” would easily overwhelm any automated system, essentially eliminating its practical use. When I entered the field in the late 1980s, this raging debate over the ethics of devices with known error rates being approved for use was stifling the field. As Dr. Bibbo notes, “It is unlikely, in our judgement, that any cytology smear reading device will be in clinical use in the near future” and that “FDA approval appears to be a long way off.” Fortunately, the junction in the road had another path that Dr. Bibbo alludes to in her presentation.

Strides in computer image analysis were advancing rapidly. New devices, such as laser scanning microscopes, were being developed that could acquire cellular images in high resolution from an entire Pap slide in under 1 min
[6]. Computing power was lagging behind, so these microscopes had to actually be slowed so that the information would not overload the buffers. In the University of Chicago Laboratory, Drs. Wied and Bibbo could now image entire slides, storing the coordinates for potentially abnormal cells, which could then be later examined manually by human observers [7]. With such systems, the power and utility of the alternate pathway was about to be recognized. Perhaps devices did not have to be perfect. Why not let the instruments do the heavy lifting of prescreening and reserve further evaluation for humans? Let cytologists do what they do best: assess high probability “alarm” cells as to whether they are normal or abnormal. The concept of “guided-screening” functionality was coming into its own. This epiphany of thought moved the field down that alternative pathway. This was the turning point for automation. It still took another decade, but freed from the need for perfection, realistic and practical clinical applications were identified, computing power advanced by orders of magnitude, and commercial ventures sought and received FDA approval.

The industry was helped greatly by parallel developments in image acquisition and analysis. The development companies found unlikely bedfellows in defense departments. If satellites could obtain high-resolution images and image analysis could find tanks and missiles amidst all the camouflage, then similar analyses could power other medical applications in radiology and pathology. In fact, the first several commercial ventures in the automated cytology industry that achieved FDA approval (namely the PAPNET [8] and NeoPath AutoPap [9] systems, both approved in 1995) came from development teams originally involved in satellite image analysis. Interestingly, neither of these devices was initially approved for primary screening tasks. This reflected the need for this emerging technology to be heavily scrutinized by regulators. Both devices were initially approved for quality control rescreening, arguably a “hard nut to crack” in the commercial clinical arena. However, in laboratories where QC automation was implemented, small, but quantifiable and significant improvements were noted, namely in the reduced numbers of false-negative Pap tests [10, 11]. Of course, the manufacturers had no intention of stopping with QC applications and they continued with FDA applications for primary screening – more robust economic business plans! This led to clinical trials which were subsequently successful for the AutoPap, first in 1998 as a slide profiler ranking slides based on probability of abnormality, and for which a percentage (25% in the USA, higher in the international market) could be archived as normal without any human review or further QC [12]; and secondly, in 2008 (renamed FocalPoint) as a true guided primary screening device, localizing potentially abnormal cells for further human scrutiny [13]. True to the discussion above, these devices were not perfect. The clinical trials of the AutoPap primary screener showed a 3% false-negative rate for HSIL [12]. There was also a fair degree of skepticism among users. Examples of device false-negatives could contain obviously abnormal cells that humans would “never miss.” However, the reality showed that the devices found more “subtle” abnormal cases that humans more frequently missed, and when the final scores were tallied, automated devices almost always found more abnormality than did human screeners, with the localization feature actually allowing users to more accurately classify the abnormal cells detected [14].

Dr. Bibbo also mentions the use of special preparations as a potential means to improve the automated screening process, but immediately dismisses this avenue as too expensive to be a likely pathway. As all readers will know, liquid-based cytology (LBC) preparations were introduced very shortly after her Presidential Address in 1983 [15]. I had the opportunity to visit the original Cytyc facility in the mid-1980s to review the new ThinPrep technology and the computer that was being designed from the ground up for screening the new LBC preparations [16]. As Dr. Bibbo mentions, DNA analysis was a potential method of abnormal cell detection, and the final version of the ThinPrep Imaging System (TIS) took advantage of this principle, using a new Feulgen-like stain which was stoichiometric for DNA (with an optical density proportional to DNA content), and could therefore identify cells having too much of it. This process took a significant amount of time, with the TIS not being finally approved until 2003, possibly not from a lack of technology, but because of another revelation that took place in the automated cytology industry. This revelation was a not-entirely-intended (but fortuitous) consequence of the investigation of the LBC method, as a “front end” to improve computer analytic performance. It was quickly found that it also improved the performance of human screeners. In addition, LBC had other advantages besides the improved cellular preservation and uniformity of appearance needed for automation. It randomized the sample, leading to improved abnormal cell “capture,” achieved significantly better adequacy rates, and observers found it a much more “pleasing” specimen to view. All of these benefits led to an improved Pap test, but as Dr. Bibbo noted, it would have been unlikely to improve the cost-benefit due to its much higher cost. The equation

Dr. Bibbo’s Presidential Address (1983)
changed when FDA approvals were piggy-backed onto the cytologic examination for other tests from the same collected sample, most notably chlamydia/gonorrhea and HPV tests. As a “one-stop” test for a variety of important diseases, the case for using LBC and the cost-benefit equation changed. Laboratories very rapidly changed from using conventionally prepared slides to LBC. Now the common use of these special preparations could be much more easily and efficiently mated with automation: the best of both worlds.

So, as it turns out, was Dr. Bibbo right, wrong, or somewhere in the middle with regard to her predictions about cytology automation? Well, to me as an observer, it was a bit of all three. Her most important correct conjecture referenced the need to change the opinion about how automation would be used in the field of cervical cancer. Her observation that the devices would never be perfect and that there was a need to change that goal was clearly prescient, and presaged exactly what subsequently happened in the commercialization of automated primary screening. She was, however, wrong that it would be a “long way off” when, in fact, it was only a few years before QC devices which were also fully capable of performing guided primary screening were approved. History will certainly forgive this “inaccuracy” as the state of computerization and the tremendous improvement in image processing and analysis coming from other industries would have certainly been difficult to predict at the time. Her second “right” prediction was that special preparations might lead the way into automation. This prediction came true with a bit of a “side track” of human LBC screening, as its advantages for that task alone were appreciated. She was “in the middle” when it came to predictions about the use of nontraditional methods to detect abnormality. She spoke about “malignancy-associated changes (MAC)” as a means to screen (even in normal cells) for patients at a high risk of harboring cervical neoplasia [17]. The presence of MAC was a real finding, but its use was never shown to be of commercial value. She does not mention immunologic markers as a means to detect disease. In 1983, of course, pathologists were just beginning to use immunohistochemical markers as a method to identify “lineage” specific features in cells (e.g., keratin for epithelial cells, common leukocyte antigen (CD45) for lymphoid cells, etc.). Dr. Bibbo had no way of knowing at the time that markers of “transformation” (e.g., p16) or of specific genetic mutations (e.g., p53) would be available in the not-too-distant future.

So what does the future hold? In some ways, I currently share some of the uncertainty that Dr. Bibbo must have felt while giving her address in 1983. It is clear that the field of cervical cancer screening is still awash in its evolution. Widespread preventive HPV vaccination implementation will change the dynamic of how screening programs must be administered. Lower disease prevalence in the population is already being achieved [18] and, as a result, cytology screening sensitivity will also decline: the 2 factors (prevalence and sensitivity) go hand in hand [19]. Even automated screening, for which habituation (human inattention) has no effect, will see some decline in true-positive case detection (more high-scoring “alarm” cells will be mimics rather than real disease). Therefore, some countries have already moved to HPV primary screening with no cytology, or cytology only as a secondary test in some HPV-positive cases [20]. HPV molecular testing remains sensitive (either viral DNA is detected or it is not, but the specificity for preneoplastic disease is decreased (more cases will be infection rather than neoplasia, particularly in the vaccinated population). These problems, most of them caused by the success of vaccination, will require a more focused method of detection for the rare “breakthrough” true-positive cases at risk for cervical cancer development. The detection of molecular “events” which are specific to preneoplastic and neoplastic transformed cells may be the appropriate screening procedure in the future. Already, tests such as the colocalized combination of p16 and ki67, detected by dual-stain immunohistochemistry, have shown high sensitivity and specificity for high-grade lesions and cancer [21]. The best use of automated screening in the future may be the detection of this or another marker (e.g., methylation status) in cellular samples. Of course, such positive results could also be analyzed in solution without the need for any morphologic examination. In addition, the detection of specific genetic events tied to neoplasia could also be achieved by rapid (and increasingly cheap) screening genome-sequencing procedures.

Also on the horizon is the advancing integration of artificial intelligence in pathology. Cytology automation formed the basis for the current expansion of image analysis functionality into histopathologic automated specimen examination. Already, computer vision and deep learning (neural networks) have shown excellent results (sometimes better than humans) in specific tasks, such as finding small metastases in lymph nodes or identifying features, heretofore not identified, that predict outcome in cancer better than traditional morphologic grading systems [22]. Future practice – sometimes referred to as “computational pathology” – will involve the merging of

DOI: 10.1159/000477374

Wilbur
many data streams [23]. For instance, cytology or histology features may be merged with radiology, laboratory, and/or clinical information in a complex matrix which might arrive at conclusions that are just not accessible to us with current methods.

In her closing paragraph, Dr. Bibbo opines: “We need not fear that automation will replace cytopathologists and cytotechnologists; instead, new developments will require acquisition of additional skills to keep up with modern systems …” This statement reflects the balance between comfort in our current positions and tension about events in the future, and is as true today as it was at the time of Dr. Bibbo’s lecture. In 1983, the fear was for the replacement of humans by a machine that would perform the equivalent of manual screening. Today, it is for our replacement by technology that will require no manual screening or review at all. Cytotechnologists clearly see this “writing on the wall” as they have already begun the planning of an expansion of their scope of practice as so-called “mid-level practitioners” trained for a variety of morphologic-based tasks hitherto performed by pathologists [24]. This dovetails nicely with the already expanding role of the pathologist into the performance and interpretation of the many new tests coming on-line in immunology and molecular genetics. Clearly, the pathologist workforce is not expanding rapidly enough to perform all the newly generated work as well as the old routine tasks [25, 26]. New midlevel practitioners should be ready to fill these gaps as they arise. Indeed, we do not have to fear for our jobs, we have only to be excited about all the new challenges that lie ahead. I am certain that Dr. Bibbo would agree!

**Disclosure Statement**

The author has no conflicts of interest to declare.
Presidential Address

Marlue Bibbo, M.D., Sc.D., F.I.A.C.

Everyone who has stood in this position where I am today, namely delivering the presidential address, has probably done so with some trepidation. I am no different.

I have attended this particular function of the society for years, and I have heard my predecessors addressing issues that we are confronted with in our field. Since I have been involved for many years in quantitative cytology, I have chosen to review the efforts made to achieve automation of cytology as well as some of the related developments that may revolutionize our field.

Attempts at automation of the cytologic diagnosis of uterine cancer began over 30 years ago. Papanicolaou worked in the early 1950s with Robert Mellors on the possible design of an automated quantitative system. In the mid-1950s the Airborne Instruments Laboratories in Mineola, New York, developed the Cytoanalyzer, which was based on the analysis of global descriptors of the cells. Since that time, both government and industry-supported projects have been undertaken in several high technology countries, such as Canada, France, Germany, Great Britain, Japan, the Netherlands and the United States. From this influx of research funds and industrial support has emerged a new scientific discipline, which one could best describe as analytical and quantitative cytology.

Highly complex and ingenious technologies have been developed. Flow cytometers that measure a few parameters in thousands of cells per second give the immunologist important analytical information. Image analysis systems give us the tools to recognize extremely fine differences between cell populations. What has not been accomplished so far is the development of an economical, practical, clinically acceptable and diagnostically reliable system to recognize and to classify malignant lesions in the female reproductive tract. However, considerable progress towards implementation of such a system has been made.
The question remains unanswered, however, as to whether fully automated cytologic diagnoses on the female reproductive tract can be achieved in the near future and whether clinicians will accept the reduced amount of information provided by an automated system while being required to provide specimens that need more work-intensive preparation techniques.

It is thus most informative at this point to pause and reflect on the original goals, the rationale of all these efforts and to examine where we stand today.

Proponents of the efforts to automate cytologic smear reading have put forward a number of arguments. One argument is the visual prescreening by cyto technologists is a tedious and exhausting activity, which may, during periods of diminished alertness, lead to failure to recognize positive cases. It is argued that field studies have shown that a worrisome number of such errors escape the quality control measures taken in most laboratories.

It has, furthermore, been argued that, in many countries, if every woman in the reproductive age group were to have the recommended number of cytologic smears taken, the existing cytologic laboratories could not possibly handle the workload. Finally, there was, at least initially, the idea that automation may bring a reduction in cost to the patient.

Many of these arguments are valid. Visual prescreening is tiring. A recent study of 750 cyto technologists from more than two dozen countries has shown that cyto technologists miss an average of 7% to 10% of all cancer cases, with some individuals missing as many as 30%, and this under conditions in which the participants knew they were being observed.

Laboratory quality assurance, i.e., review of prescreened slides by supervisory personnel and cytopathologist, has been found in other studies to miss an equal percentage, namely 10% of the presented carcinoma in situ cases.

Automation certainly would provide a non-tiring, highly discriminating evaluation, and one can state today that machine recognition of normal and abnormal cervical cells has been technically perfected to a high degree. Whether automation is necessary to handle the anticipated workload is, at least in the United States, debatable. As far as cost-efficiency is concerned, it appears at this time that the special handling and preparation requirements for an automated device make a cost saving unlikely.

One should be fully aware, however, that automation would affect only the actual cytologic evaluation of a clinical sample. Yet, extensive studies have clearly shown that the single most significant cause of erroneous negative readings of cytologic smears is variability in the taking of the clinical samples. If one had to make a recommendation as to what practical step would bring about a substantial increase in the reliability of cervical cancer detection, it would be either training of clinicians in the proper taking of the smear or training of nurses for that purpose. Patients with carcinoma in situ can be expected to have about 10% tumor cells in their ectocervical smear, i.e., even in a smear of only 30,000 cells there should be an ample supply of tumor cells for the lesion to be recognized by visual screening of the smear. Statistical sampling variables would, in only rare cases, lead to smears with a very small number of tumor cells, e.g., in only one of 100,000 women with carcinoma in situ. However, variability in the clinical sample taking may cause as many as 1% of women with carcinoma in situ to have smears in less than 10 tumor cells.

If one had to point to the single most significant cause for the failure to detect early cervical cancer, though, it would be the failure of the patient to have a cytologic smear taken. At the University of Chicago, the records for patients in various prescreening programs over a period of 21 years show that of 273,000 participants in the programs, 76.6% did not come back for a repeat smear in the following year and fully 56.2% never came back—and this in patients whose socioeconomic situation places them in a group with a higher-than-average risk of cervical cancer. Only 0.2% of the participants had a regular annual smear over an extended period of time.

Any discussion of the need for automation in cervical cancer detection must be seen in the context of these practical realities. To us, they suggest that the effectiveness of cervical cancer detection is still decisively dependent on the patient-physician relationship, on the follow-up for repeat examinations and on the clinical sample taking.

Assuming that patient compliance and proficient sample taking were achieved, what is the state of the art in automated smear reading?

The cytopathologist who visually examines smears does not set for himself an acceptable failure rate. The cytopathologist has the professional goal of diagnosing every malignant tumor and does not wish to label any normal case as cytologically "suspicious." The cytopathologist has the definite ethical goal of achieving the correct diagnosis in every case, whether or not this is humanly or technically possible. As cytopathologists, we admit that we may
make an error, but we do not plan this error in advance. We wish to avoid any and all errors, and if we make an error then we certainly do not plan to repeat the same.

Modern machine systems are not restricted to the assessment of one cell after another but also establish profiles and atypically indices for certain diagnostic groups. A machine, however, requires a threshold value of an acceptable rate of errors.

Any automated device works with an a priori established decision rule. The higher the sensitivity of tumor diagnosis is, the lower will be the specificity of the diagnostic decision. If we establish thresholds by means of which every tumor cell be recognized as such, then we have to accept that a large number of nonmalignant cells will be classified falsely as “alarms.” On the other hand, if no normal cell shall result in a false alarm, then we will miss a large percentage of tumor cells. The developer of any automated device thus has to set a threshold that seems “acceptable.”

Machine recognition of tumor cells from the ectocervix can be accomplished today with a threshold value of 90% to 95%. Even if the sample contained only 20 tumor cells, an automated device would raise 18 alarms. These figures assume, however, an evaluation of every cell on the smear in full detail. This cannot be accomplished at the present in a cost-effective manner but may be possible in a few years due to the rapid development of microprocessors and laser scanners. The Optical Sciences Center at the University of Arizona in Tucson has a laser scanner microscope in operation that can perform 128 million measurements per second. It has been slowed down to 96 million measurements per second because of the current inability of computers to accept and analyze the data at the full available scanning speed of the laser microscope. With such a laser microscope, one can screen the entire smear in high resolution in about one minute, with registration and evaluation of 1.6 billion image elements. Using the usual criterion of multiplication of the scan rate by the signal/noise ratio this laser scanner system performs better by a factor of 400 than a television-based scanning system. This also means that it is about 20 times faster than the human eye can perceive image information: we are therefore dealing here with the first microscope that is not meant to be controlled by a human eye but which serves as the “eye” of a computer system. With the currently possible data rates of the laser microscope, we can admit the full content of the available information in any given smear at very high resolution (better than 0.5 sq µm) in one single screening run. Systems based on such technologies will rapidly surpass all hybrid systems, the systems which work consecutively in coarse and fine resolution, such as we are using currently in our laboratory. How-through systems, which were developed to examine a few parameters in coarse resolution in many cells, will also become obsolete. Nowadays, we can examine many cells in high resolution while we are retaining the original Papanicolaou-stained cytologic smear for visual control and reexamination. However, the newest developments in image technology and computer science will not obviate the requirement for an excellent sample, as needed for visual examination, but in addition, will require samples specifically prepared for machine reading by a sample preprocessing device.

At this time, we ourselves work at the University of Chicago with a system that has been in operation there for over eight years and that suppresses electronically the normal epithelial cells, the leukocytes, erythrocytes and most contaminants. The remaining “unidentified alarms” are stored with their coordinates for subsequent reexamination by a high resolution system. Several other systems work on a similar principle.

If processing misses as many as 50% of all tumor cells present on the slide of an adequate clinical sample, this would still guarantee the diagnostic classification of a patient as having a malignancy.

Paradoxically, the problem for a machine classification of a smear is not the correct recognition of tumor cells. Even relatively primitive devices can do this with sufficient reliability; the problem for an automated device is the correct recognition of a normal cell as normal. Clinical sampling makes the detection of tumor cells a search for a possibly rare event since the vast majority of cells in the smear are normal cells. A 50% error rate for tumor cells will still cause a good number of alarms and flag the slide for visual review.

But, and that is the decisive point, in a sample of 50,000 cells, the automated device has 50,000 chances to make a mistake and erroneously call a normal cell a tumor cell. An error rate of 0.1% would raise 50 false-alarm signals, at which point human review of the slide would be mandatory. Unless the recognition of normal cells can be made reliable to better than one in 500,000, every single normal slide may be erroneously declared to contain tumor cells. To recognize every normal cell with such reliability is a very difficult task indeed, to say noth-
Figure 1:
Some intermediate cells from a patient with abnormal cytology are surrounded by yellow/red frames (indicated by arrows in this black-and-white illustration), which indicates the presence of atypia.

Figure 3:
Sample field for a Feulgen-stained tissue. A nuclear boundary tracking algorithm has delineated the boundaries of acceptable nuclei, the integrated optical densities of which are measured automatically pre-screened slides defeats the very purpose of automation in the first place.

Most research groups concede, at this point, that correct recognition of normal cells with the required level of reliability may not be attainable. Results from a number of experimental prescreening devices undergoing clinical testing confirm the above concerns: the devices tend to give rise to 15% to 30% false-positive diagnoses. Even more troubling, the

Figure 2:
Sample field for a Papanicolaou smear with boundaries of accepted nuclei delineated.

Figure 4:
DNA ploidy histogram for carcinoma in situ.
test devices seem to miss a substantial number of positive cases.

Approaches to overriding a certain number of “false alarms” range from simply setting a threshold to the number of false alarms to a search for clues for the presence of malignancy in the normal-appearing cells of the sample.

The search for diagnostic clues offered by the normal cells of the sample has proven unexpectedly fruitful, having led to the discovery of measurable marker features in the normal-appearing intermediate cells. So-called malignancy-associated changes were described many years ago by Nieburgs and others, but they were observed only visually and not by automated techniques. Even experienced cytopathologists might have great difficulties performing these assessments on intermediate cells, although Nieburgs and Rome examined slides under oil immersion for many years to determine if a patient should be placed in a high-risk group.

It seems evident that automated scanning for those clues would be superior to visual examination. Burger and his associates found these diagnostic features in intermediate cells using only monochromatic light.

Our own studies have demonstrated color-based features that can distinguish seemingly normal intermediate cells in patients with dysplasias, carcinomas in situ and invasive lesions from intermediate cells from truly healthy women (Figure 1). The computed values of these features indicate whether the patient has normal cytology, whether dysplasia is present or whether malignant disease is present.

With this finding, which still needs to be explored for consistency in a large patient sample, the prescreening search would no longer be a search for a rare event. Rather, a moderate size sample of 200 to 400 intermediate cells would be examined, and a statistical test would ascertain whether human visual review of the sample is needed or whether the sample should be dismissed as definitely normal.

In the computer-based cell diagnostic system developed at the University of Chicago, the TICAS project, computer graphics are used to draw a colored frame around each cell examined by the computer. The cytopathologist thus sees the cell as a high-resolution microscope image, and simultaneously he sees the cell as the computer does, through a statistical/analytical filter telling him by the hue of the color encoding this cell’s clues as to the diagnostic condition of the patient. The assessment of an entire sample, based on an analysis of these marker features, could be done even today in an economical fashion and with existing technology; yet, this research is only in its initial stages.

It is unlikely, in our judgment, that automated cytologic smear reading devices will be in clinical use in the near future. There may be some clinical testing of experimental systems, but commercially available, FDA-approved systems appear to be a long way off.

What then can we, as visual examiners of cytologic samples, realistically expect of all these new technological developments?

Certain technological advancements to assist the diagnostician are already operational. This is not only true for the above demonstration of subvisual cellular criteria. One can introduce today a rather inexpensive cell-detect module, which is easily programmable by means of mathematical filters to adjust for anatomic sites and/or other clinical data (such as age of patient, etc.). Such a system is in routine operation in Chicago and demonstrates to the visual observer only those cells or artifacts that have caused an “alarm” and that are then shown to the cytopathologist or cytotechnologist for final visual analysis. We have to admit that the system—at least in our hands—does not operate cost-effectively.

Automation has been shown to be of great value in the standardization and in cytologic and histologic differential diagnoses. Especially in the case of DNA measurements on nuclei in smears and tissues, automation is eminently successful (Figures 2 and 3). It has taken often hours and days to perform meaningful and accurate DNA measurements on cells and tissues to measure cell sizes, cell shapes, nuclear sizes and nuclear shapes then to register these data and evaluate them mathematically.

All this can now be done in a few minutes, with the findings immediately displayed (Figure 4) or printed for the patient’s chart. Only a few years ago, such measurements required special microscopes with scanning stages. These instruments were exorbitantly expensive and outside of the financial reach of most practicing cytopathologists. Today, such measurements are possible with any good microscope attached to a microcomputer, which can also be used for other laboratory functions, such as quality control, e.g., to determine whether the cellular sample and the histologic material are representative of the same lesion, as shown in Figure 5 for a case of carcinoma in situ. The microcomputer can also be used for the maintenance of patient files and the preparation of patient letters and other communications with a word-processing software addition.
If one thus considers these various important potentials of automation in cytology, one should admit that it would be a rather narrow view to see cytology automation only as a reduction in the workload of the cytotechnologist. The field of analytical and quantitative cytology, which was largely pioneered by Caspersson\textsuperscript{15–17} in Sweden, by Atkin\textsuperscript{1,2} in England and by Sandritter\textsuperscript{13,24} in Germany, provides the potential to recognize chronic, subthreshold environmental influences from which we require protection. The technologic developments currently used primarily in the field of diagnosis represent the basis for a general biotechnology, which may provide immuno-therapeutical products for future treatments of disease.

Image analysis techniques have made it possible to perceive information that has never before been utilized for diagnostic purposes; these techniques permit us to distinguish between cells that the human eye sees as absolutely identical, thus adding the power of large computers to our ability to differentiate and diagnose.

The rapid computer graphic capabilities available make it possible for the practicing cytopathologist to assimilate this novel information and to use it as a diagnostic aid (Figure 6).

It is quite possible that the real benefit of all the research into automation has always been in the improvement of diagnostic procedures and in the providing of novel and powerful diagnostic aids to the
skills to keep up with modern systems and quality controlled methods, of which our patients will be the final beneficiaries.

Thank you.

References
References