Milbank Memorial Fund Quarterly/*Health and Society, Vol. 56, No. 4, 1978*

The Annual Pap Test: A Dubious Policy Success

ANNE-MARIE FOLTZ AND JENNIFER L. KELSEY

Graduate School of Public Administration, New York University; and Department of Epidemiology and Public Health, Yale School of Medicine

SCREENING FOR DISEASE among apparently healthy people has long been an important part of public health work. Certain types of screening have become national policy either through Congressional mandate or through informal diffusion. In recent years, the utility of screening tests has been increasingly debated by public health personnel (Fogarty International Center, 1976; Lave and Lave, 1977; Shapiro, 1977). Among the most common screening tests whose widespread use is being questioned is the Papanicolaou smear test, or "Pap" test, for cancer of the cervix (the neck of the uterus). During the past 30 years, this test has been promoted for annual use for all women over 20 years of age. So well established is the Pap test that in 1973 nearly 50% of the women over 17 years of age reported having had one during the previous year, and 75% reported having one at least once in their lives (U.S. Department of Health, Education and Welfare, 1975). Despite widespread use, and despite endorsement by the American Cancer Society and by the National Cancer Act Amendments of 1974, benefits of the Pap test remain unclear.

0026-3745-78-5604-0426-37/$01.00/0

This paper is an examination of the rationale and development of policies for the use of the Pap test. As Shapiro (1975:83) pointed out:

> . . . unfortunately, no provision was made years ago for a rigorous test of the efficacy of the Pap smear through randomized clinical trials and now because of its widespread acceptance as a diagnostic tool, it is no longer possible to do so.

Therefore, one has to pass judgment on the basis of available data, inconclusive though these data may be.

We will first review the scientific basis for establishing screening policies, and then assess the extent to which the Pap test meets these criteria for screening tests. We will then examine the question of why the annual Pap test has been adopted as national policy in the United States despite serious questions about its usefulness. In so doing, we will contrast the policies of the United States with those of the United Kingdom and Canada.

Criteria for Screening

When the National Conference on Preventive Aspects of Chronic Disease in 1951 (Commission on Chronic Illness, 1957) first established a definition for screening, it also set a precedent that public screening programs could be undertaken where the health benefits were directed to the individual rather than to society as a whole. This permitted the application of screening tests to diseases such as cancer, whereas previously they had been applied only to infectious diseases (Wilson, 1968; Hart, 1975; Rosen, 1975). Following this pattern, a few years later the Commission on Chronic Illness (1957) defined screening as:

> the presumptive identification of unrecognized disease or defect by the application of tests, examinations, or other procedures which can be applied rapidly to sort out apparently well persons who probably have a disease from those who probably do not. A screening test is not intended to be diagnostic. Persons with positive or suspicious findings must be referred to their physicians for diagnosis and necessary treatment.

Since 1968, the World Health Organization (Wilson and Jungner, 1968), McKeown (1968), Cochrane and Holland (1971),

Whitby (1974), Frankenburg and Camp (1975), and Cole and Morrison (1978) have summarized criteria that screening tests should fulfill before they are applied to populations. For purposes of discussion, we have combined these criteria into five groups, as listed below. Although failure of a screening test to fulfill any single criterion does not in itself completely negate the value of the screening test, most of the criteria should be met if a mass screening program is to be effective.[1]

1. *Importance of the Disease.* The disease should be an important health problem and have a high prevalence in the community.

2. *Characteristics of the Screening Test.* The test should be simple to administer, accurate, reliable, and acceptable to the population.

3. *State of Knowledge of the Natural History of the Disease.* The disease should have a recognizable latent or early presymptomatic stage, and its natural progression from latent to declared disease should be well understood.

4. *Efficacy of Treatment.* Diagnosis and treatment should be available for patients with recognized disease and should be acceptable to them. Consensus should exist on what is appropriate efficacious treatment.

5. *Justifiability of Screening Costs.* The costs of case-finding through screening must be politically and socially acceptable. This includes the cost of the test itself, the costs of diagnosis and treatment, and the personal and social costs associated with suggesting there is disease where none exists (false positive) and suggesting absence of disease where in fact it does exist (false negative).

---

[1]For example, phenylketonuria (PKU) is of low prevalence, but the test itself is accurate and simple, the natural history of the disease known, and the disease amenable to treatment, all of which help to justify the costs and make it an acceptable screening procedure. Nevertheless, of late, it too has come under criticism.

Screening for Cancer of the Cervix

1. Is cervical cancer an important health problem with a high prevalence?

Cancer of the cervix is not a major cause of death among women in Western countries. In the United States, as Fig. 1 shows, cancer of the cervix trails far behind heart disease and stroke as a cause of death, and it also trails behind several other major cancer sites in women. In 1976, 5525 deaths in the United States were attributed to cancer of the cervix, showing a steady decline from 7108 deaths in 1968 (U.S. Department of Health, Education, and Welfare, 1978b).

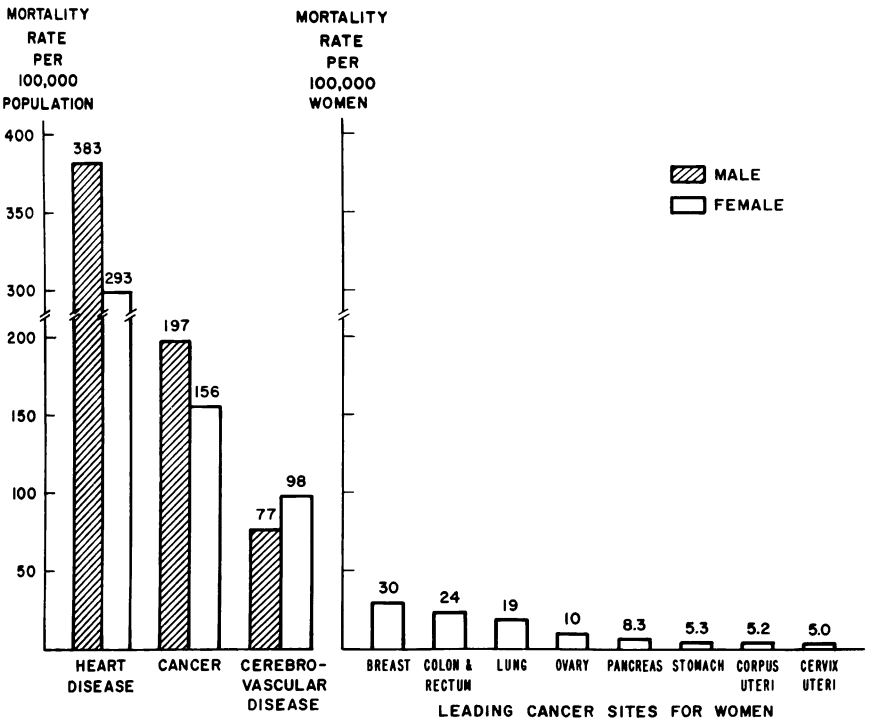

FIG. 1. Mortality rates from three leading causes and leading cancer sites for women in the United States, 1976. *Source:* U.S. Department of Health, Education, and Welfare. 1978b.

The average annual incidence rate for invasive cervical cancer from the Cancer Surveillance, Epidemiology and End Results (SEER) program (U.S. Department of Health, Education, and Welfare, 1978a) for the years 1973 through 1976 was 12.6 per 100,000. Applying this to the estimated number of women in the United States over this same time period (U.S. Department of Commerce, 1978) yields an estimate of the annual number of new cases of invasive cervical cancer of about 14,000.[2] From these mortality and incidence rates, it is evident that cancer of the cervix is a health problem of only moderate significance in the United States.

If a disease is of low prevalence in a community, screening brings in a low yield for a high effort. Moreover, the lower the prevalence, the less likely it is that a positive test will correctly identify a woman who really has cancer. This means that many women with positive test results but without disease will be referred unnecessarily for further diagnostic tests and treatment, with concomitant costs and worry.[3] The actual prevalence rates of cervical cancer, carcinoma in situ, and dysplasia are not known, since these rates can only be determined by surgery, which obviously cannot be undertaken on a sample of the general population. However, results of the Pap test indicate that cancer of the cervix is a disease of low prevalence in the general population. Stern and Neely (1963) found prevalence rates at initial screening of 5.4 per 1000 for dysplasia, 5.1 per 1000 for carcinoma in situ, 1.5 per 1000 for Stage 1 invasive carcinoma, and 1.0 per 1000 for Stage 2–4 carcinoma. The yield is even lower among those rescreened: among women with previously negative tests who had been screened from 1 to 7 years before, the respective rates were 2.4, 0.07, 0.09, and zero.

---

[2]It is realized that the SEER reporting areas are not completely representative of the United States, but they do provide a better cross section than any other source of incidence data.

[3]If the prevalence of a disease is 5 per 1000, the false-positive rate 5%, and the false-negative rate 20%, only about 8% (4 of 54) of those screened as positive will be true positives. If the false-positive rate is 20%, which is probably a more realistic figure, the false-negative rate still 20%, and the prevalence 2 per 1000, only about 0.8% (1.6 of 201.2) of those screened positive will be truly positive. In other words, 99.2% of those screened as positive will not be found to be truly positive by subsequent definitive testing. Thus, a costly burden is laid on the diagnostic facilities for an extremely low yield, while a psychological burden is placed on the women involved.

Among Connecticut women undergoing the Pap test between 1970 and 1974, the prevalence of invasive cancer ranged from 0.2 to 0.3 per 1000 women screened; for carcinoma in situ the prevalence rate ranged from 0.9 to 1.2 per 1000 women screened (American Cancer Society, Connecticut Division, 1976). Brindle, Wakefield, and Yule (1976) noted that almost 29% of the women screened in a program in northwestern England presented evidence of some gynecological disorder, and that among this 29%, about 2% had Pap test findings that required further investigation. Thus, although the yield among women with symptoms may not be low, among symptomless women in the general population the test is likely to yield few new cases for a large effort, especially if the women have had a negative test in previous years.

These differences in prevalence rates for cervical cancer result in part from variation in the types of women having Pap tests (Alvan R. Feinstein, 1978). These women come from at least four different groups represented in varying proportions in screening programs: 1) women not under current medical surveillance who would otherwise not have visited a gynecologist; 2) women under medical surveillance, visiting a doctor for some reason other than suspicious symptoms of cervical cancer, and who have the Pap test performed by that doctor during the course of that examination; 3) women for whom the Pap test is undertaken because they have signs or symptoms of cervical pathology; and 4) women previously found to be negative who are being reexamined. Each of these populations has particular prevalence rates, but very few studies take these different groups into account.

In addition, because the high frequency of hysterectomy has reduced the number at risk for cervical cancer, it is difficult to estimate accurately mortality, incidence, and prevalence rates for the disease. Nearly one-fifth of the women in the age range of 40–49, for instance, have had a hysterectomy (National Center for Health Statistics, 1966; Stern, Misczynski, Greenland et al., 1977).

2. Is the Pap test simple to administer, accurate, reliable, and acceptable to the population?

The Pap smear test consists of an analysis of Papanicolaou-stained cells taken from the uterine cervix (neck of the uterus) by scraping. The test may be taken in a doctor's office, clinic, or hospital. The

procedure is quick, simple, and may cause some discomfort. Its safety has never been in question, and it seems to be readily accepted by women.

The primary purpose of the Pap test is to detect cancer or lesions that may be pre-cancerous. Traditionally, the results of the Pap smear are reported in five classes: I = normal; II = atypical; III = suspicious (dysplasia); IV = carcinoma in situ; and V = invasive carcinoma. Some laboratories use as many as seven classifications, however, and the names of the classes may vary. Classes III through V are considered "positive" by most physicians and require follow-up with the more definitive diagnostic procedure, the cervical biopsy.

The accuracy and reliability of the Pap test have not been established despite its use for over 30 years. The accuracy of the test has been stated to be between 95% and 100% (Dickinson, 1972c; American Cancer Society, 1975; Talebian, Shayan, Krumholz et al., 1977). This statement is misleading, however. With any condition of low prevalence, this statistic can hide a high rate of false negatives.[4]

Direct estimates of false-negative rates are not available, since this would require knowledge of the proportion of women known to have cervical dysplasia, carcinoma in situ, and invasive carcinoma who have negative Pap tests. Instead, false-negative rates have been estimated indirectly by comparing results from a first Pap test to findings in second and subsequent tests or by comparing first and second readings of the same smear. False-negative rates for a single smear estimated in these ways have varied considerably, but are for the most part about 20% (Yule, 1972) to 30% (Garrett, 1964).

A cervical scrape produces lower false-negative rates than vaginal aspiration (Shulman, Leyton, and Reed, 1974). Rates from cervical scrapes alone, from many different studies, range from 2.4%

[4]In this instance, the authors of these articles presumably use the term "accuracy" to mean the proportion correctly classified by the test out of all those screened. This is dependent not only on the false-positive and false-negative rates, but also on the prevalence of the disease. If very few people have the condition, most people will be correctly classified as negative. For instance, assume that 1000 persons are screened, of whom 5 are true positives; with a false-negative rate of 20% (1 of 5) and a false positive rate of 5% (50 of 995), the "accuracy" may be calculated to be 95% (949 of 1000). Even if the false-negative rate were 100% and the false-positive rate were 5%, 945 of 1000 persons would still be correctly classified. Thus, the statistics of 95% accuracy is misleading since it can hide a very high false-negative rate for a condition of low prevalence.

to 26% (Husain, 1976). Coppelson and Brown (1974), using various sources of data, estimated false-negative rates of 40% for dysplasia, 20% to 45% for carcinoma in situ, and 24% for invasive cancer. They concluded: "These studies show the false negative rate to be so high that it cannot be ignored in the design of strategies for cancer prevention."

This lack of repeatability in Pap tests has several causes. One is that the abnormal cells may be present in one sample and not in another. Sedlis, Walters, Balin et al. (1974) found that if two samples of cellular material were taken from a woman at the same time at least one-third of the time the abnormal cells seen in one sample would not be present in the second sample for carcinoma in situ, and at least half the time for mild or moderate dysplasia. Noting the reported false-negative rates of 25%, Sedlis et al. (1974: 296) stated: "Our results tend to support a view that in real practice, as opposed to highly selective studies, the prevalence of false negative results is probably much higher than generally believed." The authors concluded that taking two samples instead of one could increase by as much as 50% the chances of getting positive results. However, this would, of course, increase the false-positive rate as well.

A second reason for the lack of repeatability is that different cytologists reading the same slides report different results (Seybolt and Johnson, 1971; Lambourne and Lederer, 1973; Evans, Shelley, Cleary et al., 1974). Kern and Zivolich (1977) had two cytologists, four cytotechnologists, and four cytotechnology students examine 112 slides from 60 subjects. They found a high correlation among readers for dysplasia, but "significant disagreement" with 38% of the carcinomas in situ and 44% of the invasive carcinomas.

Data on false-positive rates are not available since one would need as a denominator a group of women known by a definitive method (surgery) to be free from cervical dysplasia, carcinoma in situ, or invasive carcinoma. To our knowledge, no estimates have been made of the false-positive rate either by direct means (such as assembling a group of women who have had hysterectomies for reasons unrelated to findings on the Pap test) or by indirect means. As noted earlier, even low false-positive rates in a disease of low prevalence result in high numbers of disease-free women being misidentified, thus creating high follow-up costs and unnecessary anxiety.

The problems of accurate assessment of the disease extend not only to the cervical screening but also to the diagnostic tool of biopsy. Inter-observer variation among pathologists has been frequently documented (Siegler, 1956; Kirkland, 1963; Cocker, Fox, and Langley, 1968; Lambourne and Lederer, 1973). For instance, in comparing histological diagnoses made by the pathologists submitting the specimens and by a panel of expert pathologists, Brudnell, Cox, and Taylor (1973) revealed disagreement in 32% of the 728 specimens examined. Moreover, the submitting pathologists were more likely to rank the specimens as being of greater severity (carcinoma in situ) while the panel of expert pathologists tended to rank them as dysplasia or other bland epithelial abnormalities. Brudnell et al. estimated a possible 150 unjustified hysterectomies out of their sample of 728 specimens. Hulme and Eisenberg (1968) found that pathologists were well aware of the variability, with one pathologist quoted as saying: "One man's CIS [carcinoma in situ] is another man's dysplasia."

Furthermore, there is also debate as to whether carcinoma in situ is really cancer or merely a frequent precursor. Dr. Lewis Robbins (1977) described some of the reasons for the confusion:

> The Pap smear isn't diagnosing cancer, it's diagnosing a precursor. Why do they call it cancer then? Because nobody would pay attention to it when they would call it a dysplasia . . . if you call it carcinoma-in-situ then they will examine it, do something with it.
>
> I remember a battle royal at Roswell Park in 1946. . . . The pathologist was saying that the Pap smear is no good. But [one physician in the group] said carcinoma-in-situ is not cancer but we have to call it cancer. The pathologist said we can't call it cancer if it doesn't metastasize or if it hasn't already metastasized. Well, they did.

If carcinoma in situ is a precursor of invasive cancer, then this would be demonstrated in the natural history. However, as Dunn (1977) pointed out in an interview:

> We are still arguing whether in situ is a pre-cancer or not. You know, pathologists got to the point they hated to say carcinoma in situ, because carcinoma has meant one thing to a surgeon—you take it out. The pathologist had a surgical specimen. He wasn't sure whether it was a cancer if it were an in-situ lesion.

In summary, the test is acceptable to most women and simple to administer, but its accuracy and reliability have been insufficiently documented. The available evidence is not encouraging.

3. Does cervical cancer have a recognizable latent or early presymptomatic state, and is its natural progression from latent to declared disease well understood?

It is assumed that dysplasia proceeds to carcinoma in situ, which in turn proceeds to invasive cancer. However, the frequency with which each of these progressions occurs is unknown. Moreover, diagnostic biopsies applied to the suspicious lesion may themselves alter the natural course of the disease.

Some suggestive data on the relationship between carcinoma in situ and invasive cancer have been derived from establishing correlations of socioeconomic status between groups with the two conditions (Wakefield, Yule, Smith et al., 1973). Other evidence for association of carcinoma in situ and invasive cancer was gathered by Cramer (1974): the average age of women with carcinoma in situ is lower than that of women with invasive cancer; retrospective studies have revealed some overlooked cases of carcinoma in situ in invasive cancer; carcinoma in situ is often seen at the margins of frankly invasive lesions; and serial sectioning of carcinoma in situ specimens has revealed some stromal invasion on occasion.

The Canadian Task Force (1976), in summarizing all previous studies, concluded that the natural history of cervical cancer passed through three stages: dysplasia, carcinoma in situ, and invasive cancer, with the process taking perhaps 35 years. This is inferred from the evidence that the peak incidence of each of these stages occurs at progressively higher ages.

This indirect evidence is, however, insufficient. Some direct longitudinal studies have also been carried out. In Copenhagen, of 127 patients with diagnosed carcinoma in situ who were not treated, 33% had developed invasive cancer during a 9-year period (Peterson, 1956). Two other studies by Spriggs (1971) and Kinlen and Spriggs (1978) traced women in England with positive Pap tests (defined as invasive cancer or carcinoma in situ) who had refused biopsies. In the latter study, of 60 women so traced, one-third had negative tests an average of 5.2 years later, although all the women whose positive test regressed were under the age of 40 at the time of the first test. Kinlen and Spriggs concluded (1978:464):

> If these cases are representative, then one-third of the biopsies (usually cones) done in Britain because of positive cervical cytology are performed for lesions which are insignificant or would have disappeared if left alone.

These studies, because of their small samples and possibly self-selected populations are not totally reliable, and for ethical reasons, they are hard to replicate. Nevertheless, they indicate that regression can occur and that progression is not inevitable.

Nor does dysplasia always lead to carcinoma in situ. Stern and Neely (1963) found that dysplasia could appear rapidly and also regress, though it tended to appear again. In a study of women having shown dysplasia on a Pap test 1 to 7 years previously, 48% remained dysplastic, 40% regressed to normalcy, 11% progressed to carcinoma in situ, and 1% progressed to invasive cancer. Hall and Walton (1968) reported higher progression rates, 29% for those with severe dysplasia over a period of 1 to 14 years. Here again, it is likely that women with dysplasia are at a considerably higher risk for carcinoma in situ, but the progression is by no means certain.

Even if one assumes that the tumors do progress some of the time, the question still remains as to how quickly they grow. Great variation has been noted in the rate of growth of cervical tumors (Pederson, Høeg, and Kolstadt, 1971). A screening program offers the opportunity to identify slow-growing tumors with a long pre-invasive phase rather than fast-growing ones that quickly cause symptoms. With other forms of cancer, such as breast and lung, these fast-growing tumors unfortunately have been the ones associated with the worst prognosis (Charlson and Feinstein, 1974; Wells and Feinstein, 1977), and there is no reason to believe that this would not apply to cancer of the cervix as well. Unless tests are carried out every few months, these fast-growing tumors will be missed in screening programs.

In summary, women with cervical dysplasia are at higher risk for carcinoma in situ, and women with carcinoma in situ are at higher risk for invasive carcinoma. These progressions do not always occur, however, and certainly not at a constant rate. Much remains to be learned about the natural history of cervical cancer.

4. Are diagnosis and treatment available and acceptable to the public? Is treatment for cervical cancer effective?

Recommended treatment procedures have changed over time. By 1977, the usual procedure following a "positive" finding from a Pap test was a biopsy. In recent years, this biopsy has been recommended

to be carried out using colposcopy (visualization and magnification of the cervix 15×). This is an office procedure and is more accurate than other types of biopsy. Other office procedures for excision of lesions include cryotherapy and electrocautery (Homesley, 1977). Before extensive use of colposcopy, and in the many areas where physicians trained in colposcopy are not available, the usual procedure has been to carry out a conization (removal of part of the cervix), which requires hospitalization. The extent to which colposcopy is available in the United States is not well known. Although most expert pathologists agree that it is not good practice to move directly from positive Pap tests to cone biopsies or to cryosurgery, or from punch biopsies to hysterectomies, the actual frequencies with which such procedures take place are unknown.

For cases of invasive carcinoma confirmed on biopsy, the usual procedure is total hysterectomy. Cases of carcinoma in situ may be treated by conization or cryosurgery if the woman is interested in future child bearing, but more likely by a total hysterectomy. In recent years, physicians in major medical centers have recommended not doing hysterectomies for cases of mild or moderate dysplasia but maintaining follow-up through colposcopic evaluation, biopsies, and further Pap tests. Since the patterns of diagnosis and treatment following positive Pap tests have not been surveyed in this country, it is difficult to know whether these recommendations represent actual practice or a sought-after ideal. Both biopsies and hysterectomies seem to have been well accepted by the medical profession and the public. For the most part, facilities have been available throughout the country.

The method of treatment of certain types of lesions has been debated. For example, Hulme and Eisenberg (1968) found that Connecticut hospitals varied in their treatment of micro-invasive carcinoma that is histologically on a spectrum between carcinoma in situ and invasive carcinoma. Two-fifths of the hospitals treated it as carcinoma in situ, an equal proportion treated it as invasive cancer, and the rest had no consistent policy. The issue in micro-invasive carcinoma (which goes by many different names), of whether to perform a radical or simple hysterectomy, is not easily resolved (Burghardt and Holzer, 1977; Benson and Norris, 1977).

Whether present treatment methods are effective has not been satisfactorily shown. The American Cancer Society, in its pamphlets to the public (1975), emphasized that, when pre-cancerous con-

ditions and carcinoma in situ were detected sufficiently early, treatment was almost 100% successful. Several studies, however, have cited recurrences of carcinoma in situ or invasive cancer following treatment by either conization or hysterectomy (McIndoe and Green, 1969; Kolstad, 1970; Creasman and Rutledge, 1972). Kolstad and Klem (1976) estimated a 2% recurrence rate over a 10-year period following either hysterectomy or conization for carcinoma in situ. For most women, the disease recurred within 3 years, and the form of treatment seemed irrelevant.

The ultimate goal of mass screening for cervical cancer is to reduce mortality from this disease. Since adequate trials of the efficacy of the Pap test were not established when screening programs were first implemented, and since controlled randomized trials are now considered to be unethical, one has to conclude what one can from trends in mortality rates in areas where screening programs have been in effect, compared to areas without formal screening programs. This comparison is complicated by differences that may be due to other confounding variables such as socioeconomic status, by inadequacies in the reporting of mortality statistics (Feinstein, 1968), and by the knowledge that mortality rates for uterine cancer were decreasing before the Pap test was widely used.[5]

Analyses of trends in mortality rates for cervical cancer in areas where the Pap test is widely used have produced varying results, both when such areas are studied by themselves (Dickinson et al., 1972a and 1972b) or when such areas are compared to localities with lower Pap screening rates (Christopherson, Parker, Mendez et al., 1970; Ahluwalia and Doll, 1968; Kinlen and Doll, 1973; Cramer, 1974; Grünfeld, Horwitz, and Lysgaard-Hansen, 1975; MacGregor and Teper, 1978). However, it seems that in areas where the Pap test has been extensively used for a long enough period of time, there is a small but real decrease in mortality rates attributable to the Pap test. Even then, Pap screening programs have not been responsible for a substantial reduction in mortality.

Perhaps the most useful, although by no means conclusive, study was reported from Canada. Miller, Lindsay, and Hill (1976) studied mortality trends at the county and census division levels in

[5]Until 1949, mortality rates reported for uterine cancer combined the two parts of the uterus (the corpus and the cervix). Cervical, rather than corpus, cancer almost certainly accounted for most of the decrease in mortality rates before 1949.

Canadian provinces according to the intensity of screening activity; they also took into account census-derived indices of socioeconomic status that were known to be related to the incidence of cervical cancer. They found that, although much of the variation in mortality rates for cervical cancer from one area to another could be attributed to factors other than the intensity of screening, and although there was little evidence that screening brought about any decrease in mortality rates from 1950–52 to 1960–62, screening programs did contribute significantly to the overall reduction in mortality between 1960–62 and 1970–72.

None of these investigators took into account how much of the decrease in mortality from cervical cancer could be attributed to the increasing frequency of hysterectomies: this would, of course, mean that fewer uteri are at risk for developing cervical cancer. Estimates of the effect of hysterectomies on the decline of cervical cancer mortality rates vary from 10% (Lyon and Gardner, 1977) to 25% (Stern, Misczynski, Greenland et al., 1977).

Finally, none of the studies showing the effect of screening on cervical cancer mortality has been differentiated by recommended screening intervals for individual women. In Aberdeen, for example, women are recalled for screening every 5 years (MacGregor, 1976), while in British Columbia (Ahluwalia and Doll, 1968; Kinlen and Doll, 1973) and Louisville (Christopherson et al., 1970), the intervals are closer, but not specified. Yet all three programs report mortality rate decreases. Also, it is not always clear whether these screening programs include mostly women who are under medical surveillance anyway, or whether they attract high risk women who would otherwise not be screened.

5. Is the cost of case-finding through screening politically and socially acceptable?

A necessary condition for any effective program is that individual and social benefits from the screening procedures outweigh the costs such programs generate. Different ways of assessing costs and benefits have been tried with varying intensity and interest on both sides of the Atlantic and, at least in Britain and Canada, they have generated much debate. The issue is not so much whether the Pap test should be used, but who should be screened and with what frequency.

United Kingdom. In 1968, Knox in Britain had raised the issues of the cost of screening and suggested that the cost of preventing a clinical carcinoma would be £1000, about the same as the cost of treating it. In 1972, a *Lancet* editorial stated that large investments in screening at best would be likely to reduce mortality over a long period of time by only about 50%. (These latter findings were based on Knox's work.)

The next year, the *British Medical Journal* (1973) followed with an even stronger statement: "It is not a question of proving that screening has *no* value . . . but of deciding whether it has sufficient value to justify the risks and effort it entails" (italics in the original). This editorial prompted two defensive responses upholding the effectiveness of screening in reducing mortality: Christopherson (1974) cited studies in Louisville, Kentucky; and MacGregor (1974) cited data from Aberdeen, Scotland showing that the mortality rate in that area had been falling more rapidly than in the rest of the U.K.

Meanwhile, Brudnell, Cox, and Taylor (1973) had raised issues about unjustified hysterectomies because of improperly read histological specimens, and Richards (1974) had indirectly raised the issue of the personal costs of hysterectomy. Richards found that hysterectomy patients reported that an average of 11.9 months had elapsed before they felt completely convalesced, while a comparison group of women who underwent other forms of major surgery reported an average of 3 months. Certainly, this was an area in which little research had previously been done and demonstrated a concern not only for monetary costs but for personal costs as well.

The concern for costs and a conservative approach remained strong in Britain. Holland (1974) concluded that the *Lancet* series on screening "leaves the impression that total screening is hard to justify for any condition, save PKU." Randall (1974), in the same series, had stated that cervical cancer screening did satisfy most of the criteria for a selective screening program for high risk populations, but the wrong people were being screened.

These concerns were reflected in suggestions that doctors examining women under 35 years old be less zealous in taking repeat smears if there were no symptoms (Brindle, Wakefield, and Yule, 1976). The same year Knox (1976), using a more complex model than he had in 1968, concluded that a series of 10 screening tests for women between the ages of 35 and 80 years would be optimal and capable of preventing about 77% of all deaths from cervical cancer.

These calculations were made with the acknowledgment that much factual information, such as error rates and natural history of the disease, was still lacking and could noticeably alter conclusions. Nevertheless, in Britain, one could detect a clear trend toward less frequent use of screening because of its questionable effectiveness, its monetary costs, and its personal and social costs through high hysterectomy rates.

United States. In the United States, the debate on costs was unfolding differently, with the question of social costs or benefits receiving less attention. A program analysis by the Department of Health, Education, and Welfare as early as 1966 had suggested that one would save $1000 per case by treating carcinoma in situ early rather than invasive cancer when it appeared later (U.S. Department of Health, Education, and Welfare, 1966). Based on persons appearing for screening and treatment at the Mayo Clinic, Dickinson (1972c) found that money could not be saved by screening activity.

Another approach (Schneider and Twiggs, 1972) had been to estimate the annual costs of screening and not screening a theoretical population of 100,000 women. The authors concluded (1972: 857) that the only practical way to limit costs was to reduce the number of smears taken to one every 3 years in order to "relieve the overburdened health worker of the tyranny of the tedious 'Annual Pap and Pelvic'."

These last two estimates of costs and benefits were based on false-negative rates under 10%. In fact, the rates have been much higher. Estimating a 40% false negative rate, Coppelson and Brown (1974) concluded that the cost of delivering screening and follow-up services to 80% of the eligible women in America would amount to $1 billion, or 1% of the annual U.S. health expenditures.

In actual practice, monetary costs in the United States had been higher than one might expect, although most of these costs have not been formally reported. During 1974, 484,773 Connecticut women (American Cancer Society, Connecticut Division, 1976) had Pap tests at a cost of about $2.6 million for laboratory work alone.[6] This is about the same as the state expended annually for all its public maternal and child health services. The cost per case of cancer

[6]If we add to this the cost to the woman of $25 to visit the gynecologist, the total cost is $6 million.

detected, including invasive cancer, carcinoma in situ, and adenocarcinoma, was $3322, and this did not include the costs of the diagnostic biopsies.

The New England Journal of Medicine (1976) in an editorial raised the issue of whether the wrong women were being screened, since it was the middle-class low-risk women who were coming in for Pap tests, while the low-income high-risk women were not being seen in clinics. Six months later the same journal reported a round table discussion of the personal costs of the hysterectomies that usually follow a diagnosis of carcinoma in situ and sometimes even dysplasia. In this discussion, Cole (1976) noted that the discounted cost of providing women over 45 years with prophylactic hysterectomies to prevent future cancer would be about $9800 per year of life saved. The disadvantages would be the mortality rate of hysterectomies (estimated at 0.06% to 0.2%), and possible adverse metabolic and endocrine effects. Cole concluded that the benefits of hysterectomy for the purpose of cancer prophylaxis were insufficient to justify costs. Notman (1976) pointed out that a high rate of severe mental depression seemed to follow hysterectomies, and these effects were of sufficient importance that they should be taken into account and studied further.

Despite the comments on costs, despite the concern for excessive hysterectomies, and despite suggestions that screening should be carried out less frequently, many groups in the United States continued to support the annual Pap test. The recommendations of the Workshop on Uterine-Cervical Cancer in 1973 (Koss and Phillips, 1974) had called for cervical smears for all nonvirginal women (the designation by sexual activity rather than by age was a welcome recognition of at least one etiologic agent for the disease), and suggested that a cervical smear be done on all hospital in- and outpatients unless the patient had participated in a cytologic survey the previous year. The only concession on the question of cost came when the Workshop noted that "routine annual cytologic examinations cannot be justified economically in all patients." But having said that, the Workshop reneged: "This Panel cannot at this time make a firm recommendation about the frequency of cytologic examinations beyond the currently accepted annual interval." In essence, annual examinations were too costly but the panel refused to consider other options and continued, at least tacitly, to approve the annual exam.

As for social and personal costs, the Workshop recognized them in regard to dysplasia: "These lesions should be managed in the fashion most consistent with the well-being of the patient and, if appropriate, the preservation of her reproductive function." It is difficult to see how a physician would translate this general and ambiguous policy into practice.

One can get some inkling of the physician's expected response to this controversy from the *Journal of Family Practice,* an American journal devoted to disseminating information to family physicians. A 1975 review of the cervical screening literature (none more recent than 1969) pointed out that previous studies "suggest that cancer of the cervix is a slowly progressive disease requiring 5–10 years to progress from dysplasia to invasive carcinoma" (Frame and Carlson, 1975). The authors did not indicate that the progression did not always occur, nor that duration of the progression was not well established and probably the duration was longer than they indicated. One wonders how such statements influenced physicians in their decisions to perform or recommend total hysterectomies. However, the authors did diverge from the traditional annual Pap test philosophy by recommending the test be used only every other year.

The American Cancer Society had continued to stress the annual Pap test in its lay and professional publications. Moreover, it confused the public by implying that the Pap test was aimed at the detection of uterine cancer as a whole rather than just cervical cancer. The Pap test, as a rule, is designed to detect cancer only of the uterine cervix, the neck of the uterus. Moreover, uterine cancers other than those of the cervix at present have higher incidence and mortality rates. The following statements were made:

> If every woman had the test every year, most uterine cancer could be discovered in time for cure (American Cancer Society, 1968).
>
> A Pap test regularly, once each year, is the way to help protect yourself (American Cancer Society, 1973).
>
> The choice of a screening interval depends upon many complex factors, but is generally chosen to be one year (Nelson, Averette, and Richart, 1975).
>
> Each year, on some regular date, have a thorough health checkup . . . including a Pap test (American Cancer Society, 1976).

The American Cancer Society also noted that: "Total abdominal hysterectomy is, without question, the most commonly used form of definitive treatment" (Nelson et al., 1975). If analysts were raising questions about costs, frequency of screening, and excessive hysterectomies, they were not being heard in the public forum.

Canada. Canada, like the United States, had had a long running cervical cancer screening program. It had also supported epidemiological research. Unlike their counterparts in the United States, however, Canadian researchers had begun to question the value of all types of screening programs, perhaps in response to their universal provincial health insurance programs and high health costs (Sackett, 1975). Spitzer and Brown (1975) questioned whether a program of biannual health exams for all persons in the Province of Ontario would be worth the costs, considering how little was known about the benefits. He was more optimistic about the benefits of the Pap test.

However, the Pap test was already under scrutiny by a task force appointed by the Canadian government to produce a report on one of several programs and health care activities "whose effectiveness was in doubt" (Canadian Task Force, 1976: 981). Known as the "Walton Report" and published in June, 1976, it carefully reviewed the scientific evidence on the effectiveness of the Pap test, its availability, accuracy, and costs. It recommended for women at low risk Pap tests 1 year apart, then tests at 3-year intervals (if the earlier tests were negative) until the age of 35 years, and then Pap tests at 5-year intervals until the age of 60. High risk women, who were identified as those of low income, with early onset of sexual activity and multiple sexual partners, were recommended for annual screenings, but it was recognized that these women were the hardest to reach.[7]

Reactions to the Walton Report. The reaction to the report was swift in Canada and Great Britain, and lethargic in the United States. In Canada, the report had been published in the *Canadian Medical Association Journal* so that it received immediate and wide

[7]Whether users of oral contraceptives should be included among high risk women is subject to debate. At present, the evidence is unclear as to whether users of oral contraceptives are at high risk independently of other risk factors (World Health Organization, 1978).

coverage in the professional and the lay press. Although it acknowledged the scientific merits of the report, The Society of Obstetricians and Gynecologists was the group most unhappy with the reduction in frequency of screening (Schmidt, 1977: 972).

> The periodic examination is the mainstay of our present health care delivery system and . . . annual cytologic screening of the cervix is a reasonable safeguard against cancer.

Nevertheless, because of the vast publicity, Canadian women were informed of the need for less frequent Pap tests.

In Great Britain, reaction came in the form of editorials in both the *Lancet* (1976) and *British Medical Journal* (1976) within a few months of publication of the Report. Both journals summarized and commended the findings. Since annual Pap tests had never been a part of British policy, no changes were required, although the findings did help support those who had been concerned with the high costs of screening.

In the United States, the Walton Report was not so much disputed as ignored. An editorial in November in the *Journal of the American Medical Association* (Danilevicius, 1976) suggested that at least in the United States, based on the Louisville studies and some New York studies, the screening programs had potential for decreasing morbidity and mortality, and that "the natural history of carcinoma of the cervix . . . as well as the anatomic availability [*sic*] of the cervix for repeated examination are circumstances particularly suited to a screening program." Danilevicius concluded (1976:2099):

> Using consultations with gynecologists and availing themselves of the cytologic expertise of pathologists, the majority of primary care physicians could include all their female patients in successful, continuous cervical cancer screening and prevention programs.

It was as if the Walton Report had triggered a response of recommending increased rather than decreased screening activity.

At the end of September, 1976, the Cancer Control Division of the National Cancer Institute sponsored a conference on the "state of the art" of cervical cancer screening. The issue of costs was raised only in terms of reimbursement for Pap tests and follow-up by third party payers. Overall program costs were not addressed. The issue of frequency of screening was also raised. At the meeting, it seemed as

if a consensus had emerged that annual screening should be continued. However, even though draft reports circulated informally, the conference had still not issued a final report by mid-1978, and it seemed unlikely that one would emerge. This indecisiveness may itself have been an indication of the impact of the Walton Report.

Since 1976, there has been a tendency among certain health professionals to move away from clear statements of recommended annual tests. With the annual health examination being questioned as well, Breslow and Somers (1977) suggested a lifetime health-monitoring program by which, for example, women 40 to 59 years of age would have a Pap test at 2- or 3-year intervals rather than annually. Even the American Cancer Society (1977: 17) rephrased its recommendations to suggest that: "Uterine cancer could be reduced dramatically as a cause of death if every woman had a Pap test with her regular health checkup . . . " Again, the term "uterine cancer" was used rather than cervical cancer, leading to confusion as to what the Pap test actually detects.

For the most part, however, Americans view the benefits of annual screening for all women as accepted dogma. The firm convictions of most health professionals are reflected in statements such as the following:

> Deaths from carcinoma of the cervix could be dramatically reduced if every adult woman had a Pap test annually. (Homesley, 1977)

These convictions have been communicated, as well, by the lay press:

> Young women should have an annual pelvic exam which includes a Pap smear. (*Glamour,* 1978)

> Getting rid of the concept that most people need a complete annual checkup would probably reduce our national health budget by several million dollars. (Let me add one caveat here. Most doctors still believe that women over 30 should have regular—probably annual—Pap smears, a test for cancer of the cervix, particularly if they are on the Pill.) (Nolen, 1978)

> Pelvic cancer cannot be seen or felt by a woman, therefore it is terribly important to have a Pap test at least once a year, starting at first menstruation or at least at first sexual activity. (Ask Beth, 1978)

In the United States, the political and social costs of screening are apparently acceptable. The annual program is not only accepted by women who pay the costs, but also by politicians in Congress who sponsor the program for those who cannot pay for it themselves. However, the costs in the United States have for the most part been hidden and ill defined. In contrast, Britain and Canada have found the political and social costs of an annual program unacceptable.

Persistence of Screening Policy

Why does annual screening persist as a policy in the United States if it is of dubious value?

What accounts for the differing points of view among British, Canadian, and American policy makers, and also among researchers themselves about their findings? One factor, already noted, is the state of the art, which is not as far advanced as those who work with the techniques or their results would like.

The second factor is the political environment for screening programs. To understand why the policy of an annual Pap test for all women over 20 years persists, we must examine the context for health policy-making in the United States. Screening enters the political context as soon as someone must pay for services not previously provided. For cytological screening in the United States, the individual woman pays for her visit and laboratory test, but laboratories, hospitals, screening programs, and cytologists in training have all been supported by public funds. Thus, public and private decisions are being made simultaneously. In Britain, and more recently, in Canada, virtually all funds for screening come out of the public purse one way or another. When funding for a screening program is mainly public, it is easier to assess competing priorities and determinations of need.

A program of dubious value such as the annual Pap test requires a fertile environment and an interested group of supporters in order to flourish. In the United States, the annual Pap test has had both. The fertile environment has been provided by American ideology, and the supporters have been a group most notably associated with the American Cancer Society. The contrasts between Britain and the United States help to illustrate how these factors have affected screening policy.

First, American ideology has always favored novelty and entrepreneurship. More than 150 years ago, de Toqueville (1955: 443) wrote that in America the idea of novelty is "indissolubly connected with the idea of amelioration." Americans often proceed with programs without assessing costs and benefits. Wilson (1963) noted that, whereas the British have emphasized the need for validating methods, Americans have tended to carry out screening before knowledge was great.

The number of new technologies implemented without appropriate clinical trials continues to grow in the United States and to be of some concern (Iglehart, 1977). Some, such as computerized tomography (CT scan), have been introduced regardless of the high costs of their use. Although Great Britain was the first country to pioneer these machines, they are not yet accepted there as standard equipment and, because of their costs, will probably not be for some time, if ever. In contrast, in the United States, hospitals and clinics throughout the country are vying with one another to purchase these expensive machines. Supporters say that new, important, and undreamed of uses have been discovered since the CT scanners were put into place. This may be a case where invention becomes the mother of necessity.

Other technologies may not only be costly, but potentially harmful as well. Coronary artery bypass surgery is so well established that about 60,000 operations are carried out yearly without the technique ever having been subjected to controlled trials to test its efficacy (Preston, 1977). The American rate of these operations is more than 10 times that of Great Britain or other western European countries.

American reactions to new technologies tend to be: Do what can be done, not do what needs to be done. The British tend to be conservative in their screening decisions: Do no harm. Americans think differently on social costs, as Hampshire (1976) noted:

> There is a notorious difference between Britain and the United States ... a deep difference, both of conscious feeling and unconscious strategy. The political argument between the parties in Britain typically circles around the question, How can we minimize the worst possible outcome being realized? In the United States the question is apt to be, How can we maximize the chances of the best possible outcome?

Second, the American political scene has traditionally been viewed by its analysts as an arena where competing interest groups with varying strengths, resources, and interests determine what policy is made (Truman, 1958). This viewpoint may be helpful in understanding the American commitment to mass screening. One distinguishing feature of the American health establishment is that it has been backed by a powerful lobby, the American Cancer Society (ACS), which has mobilized to promote cancer research, screening, and treatment. Its lobbying effort resulted in the establishment by federal law of the National Cancer Institute (NCI) in 1937. During the 1950s and 1960s, lobbying by the ACS and others was so successful that Congress during this period continuously raised appropriations of the National Institutes of Health (NIH), of which the NCI was a part, beyond the requests made by its own director (Strickland, 1972).

The ACS also played a key role in promoting the Pap test as the major and perhaps only cancer control device available (Breslow, 1977, I: 220). It sponsored the first interdisciplinary conference in 1948 and was (ibid., I: 225): "instrumental in providing the support necessary for establishing numerous cytology screening programs and cancer detection centers throughout the United States." In 1957, the ACS launched its "Uterine Cancer Year" which, with great publicity, was designed to promote the use of the Pap test for annual screening.

Meanwhile, the NCI had formed a cancer control branch which worked very closely with the ACS and supported screening projects at first to assess their feasibility and then, with the Louisville project, to attempt to demonstrate their impact on mortality (ibid., I: 225–232).

American Cancer Society activities culminated in the National Cancer Act of 1971. The "War against Cancer" was viewed with the same fervor as a moon shot or the development of the atomic bomb (Rettig, 1977), and the rhetoric was not always grounded in fact. The National Panel of Consultants on the Conquest of Cancer (1971: 42), a blue ribbon panel of scientific experts and distinguished laymen, reported to Congress:

> On a world-wide basis, cervix cancer is the most common form of cancer in women. Among all women in the United States, it is second only to cancer of the breast and, in low income groups, it is even more prevalent than breast cancer.

Figures elsewhere in the report as well as other sources, however, indicated that this statement was untrue. Yet, it has been widely cited. The Panel also maintained that studies had demonstrated incidence and mortality from cervical cancer could be greatly reduced if every woman were tested annually (ibid.: 42), even though the best data at the time could not confirm this.

When the National Cancer Act had been passed in 1971, a subinterest group had wanted funds for the cancer control division that had supported Pap test screening programs and that had been phased out by NCI and NIH nearly 2 years earlier. This cancer control group had a staunch champion in Representative Paul Rogers (1972: 495), the Chairman of the House Subcommittee on Health and the Environment, who said about the cervical cancer control program: "This screening and testing was and is the most effective tool we have in fighting cancer." Partially in response to the efforts of a "working Group 8," which was formed to advise NCI on cancer control, in 1974 Congress amended the National Cancer Act to include (Section 409a): "programs to provide appropriate trials of programs of routine exfoliative cytology tests conducted for the diagnosis of uterine cancer."[8] Authorizations of $55 million for fiscal 1975 rose to $89 million by fiscal 1977.

The ultimate and worthy goal of the ACS is to eliminate cancer. Meanwhile, it must keep the disease highly visible in order to raise funds for its many campaigns. Thus, its policy to encourage mass screening programs also supports the organization's viability and visibility. Through the Society's efforts, screening for cervical cancer becomes national policy without any member of the public speaking for it and without the whole-hearted endorsement of epidemiologists. One should not be surprised that the ACS, with its $100 million budget, is concerned primarily with eliminating cancer, but the American public is hard pressed to counterbalance the ACS's powerful appeal with its own weaker public interests such as social, personal, and economic costs.

---

[8]Exfoliative cytology had to be specifically cited because the 1971 Act had been interpreted to mean that NCI's activities in cancer control were to be limited to new techniques. Hence, the Pap test, as an old technique, was excluded. The 1978 Cancer Act amendments, which removed this earlier ambiguity about new and old tests, therefore dropped the specific reference to exfoliative cytology or any other procedure (Rogers, 1978).

Given these political and ideological constraints, a change in policy seems unlikely even if new technological information were uncovered in the near future. The only change could come if health professionals, politicians, and the public itself were to think in terms of nationwide social and economic costs, and about competing priorities, and in terms of minimizing monetary and personal costs as well as maximizing benefits.

Conclusions

As currently constituted, screening programs for cancer of the cervix do not meet the criteria for mass screening annually of all sexually active women. In fact, both the Walton report and Knox's (1976) studies indicated the desirability of discouraging annual tests for women with several negative Pap tests and those who are at low risk for cervical cancer. Thus, the extensive advertising for annual tests should be stopped. The frequency with which women in different risk groups should be tested needs to be reconsidered, in light of the Canadian and British experience. Screening programs would also be more effective if studies were carried out concurrently to improve the readings of the tests and to minimize the number of unnecessary biopsies and hysterectomies.

In the United States the issue remains: Who should have a Pap test and how often? This is no longer just an issue between a woman and her doctor. Much of the use of the Pap test is now mandated by state and federal law. In many parts of the country, hospitals are required to carry out a Pap test on any woman admitted who has not had one during the previous 3 years. Federal regulations require federally-funded family planning clinics to take annual Pap tests on women requesting any type of birth control device, despite a very low yield of positive results.

Theoretically, one should not make policy without facts. Yet the present annual Pap test policy was made without reference even to those facts that were readily available. A new policy should now be instituted; enough information can be collected to make it feasible to do this within a year. Haste is appropriate, since no new policy is also policy—bad policy.

The most rapid and satisfactory mechanism for changing policy in this case would be a national task force charged with determining

the appropriate standards and guidelines for using the Pap test for women of different ages and different risk groups. To operate effectively, such a body would have to be immune from many of the political pressures discussed earlier. This immunity would be best assured if the task force were located in a scientifically accepted and highly prestigious organization not already committed to the promotion of the Pap test program. The National Academy of Sciences comes to mind as one particularly appropriate organization. A second choice would be the National Institutes of Health. (Although the NIH have lodged within them the responsibility of cancer control programs, they have managed to maintain a fair degree of detachment at higher levels.)

The task force should be composed of experts in the fields of epidemiology, oncology, biostatistics, pathology, gynecology, economics, and policy analysis. If such understanding is not otherwise possessed within its specialist membership, the task force should include additional experienced persons sensitive to the special social and mental, as well as physical, problems women encounter in their roles as consumers of medical care. The task force should address itself both to immediate policy needs and to more long-term research requirements. Within 1 year of its formation, the task force should be able to use the basic findings of research undertaken to date to provide guidelines for physicians, clinics, and women as to the frequency with which the Pap test should be administered for different population risk groups during the next 4 or 5 years. The guidelines the task force will recommend after its year of deliberation are by no means clear-cut. As a *Lancet* editorial (1978: 1030) recently pointed out:

> Not enough is known for firm conclusions to be drawn about the best time for screening to begin and the best time for subsequent smears to be taken.

The editorial continued that a review was also needed of screening policy in the United Kingdom, particularly in the light of the fact that the British do not recommend screening women under 35 years of age even though women under this age appear to have an increasing incidence of cervical cancer.

Among the risk populations to be considered should be, first, all women with symptoms, then women of low income, women with early onset of sexual activity, and women with multiple sexual partners. Consideration should be given as to whether women taking

oral contraceptives and women whose mothers took diethylstilbestrol (DES) during pregnancy should be included in high risk groups. The policy guidelines should be set in terms of overall epidemiological assessment of benefits and risks and costs to populations, not in terms of individual cases. Cost-effectiveness studies, such as those undertaken by Knox (1976), should be reviewed. If insufficient, others should be commissioned.

It is most important that the task force consider the Pap test in terms of its effectiveness in combatting cervical cancer, not in terms of its possible value as a means of luring women to gynecologists' offices once a year so that other examinations can be carried out. If reasons exist why all women should be seen annually by a gynecologist, or anyone else, then these reasons should be demonstrated independently through epidemiological evidence.

Meanwhile, the task force should have undertaken the equally important job of reviewing the status of research in the field of cervical cancer and should set in motion the research that would establish the following information: what is the reliability of the Pap test under normal as opposed to special research conditions? What are the true costs of a screening program, not only in terms of screening costs, but also in terms of diagnosis and follow-up treatment for all women in the United States? What diagnoses and treatments actually follow from positive Pap smears in the United States? This latter study will necessitate surveys of what constitutes a positive Pap smear in a sample of various rural and urban, university and non-university medical care centers in the United States, as well as surveys of what diagnostic and treatment modality usually follows from such a positive smear—whether it is colposcopy, punch biopsies without colposcopy, conization, or whatever. Data on the cost of each procedure should also be collected.

The treatment and risks resulting from these diagnostic procedures should also be reviewed. Currently, there is a notable lack of knowledge about what diagnostic and treatment modalities are used around the United States. Particular attention should be paid to the proportion of carcinomas in situ for which hysterectomies are recommended. It has been suggested that recently physicians have tended to recommend fewer hysterectomies for women still of childbearing age; however, there is little documentary evidence on this point. In view of the evidence that carcinoma in situ sometimes regresses, hysterectomy may not always be the treatment of choice, since a hysterectomy has its own mental and physical

risks. Treatment by hysterectomy should be more carefully monitored in the future, particularly if it remains such a frequently performed operation in the United States.

It is possible, even likely, that after all these studies have been collected, the task force may want to revise its recommendations on the Pap test. Such a possibility should in no way discourage the group from using present knowledge to improve inadequate policies.

The task force is suggested as a particular solution for the case of the Pap test, a program that has been in force for 30 years without ever having been subjected to clinical trials. Ideally, new technologies should be regularly reviewed through regularly established services such as the National Center for Health Technologies in the Department of Health, Education, and Welfare, newly mandated by Congress in 1978. Until such review is a matter of routine procedure, ad hoc bodies, such as the task force recommended here, are the most adequate review mechanisms.

The successful diffusion of the annual Pap test is an excellent example of the ease with which a new technology or device can be rapidly disseminated to a population within a short period of time when given a great deal of publicity. Its widespread acceptance may, in part, have been dependent on the fact that it was directed to women who, as a group, seem to be more willing than men to undergo medical examinations. This misguided success may serve as a source of cheer for those who despair of ever getting any technologies diffused. It might better serve as ample warning that all medical devices and techniques need to be tested for safety and efficacy before they are diffused.

For nearly 20 years, American women have been bombarded with information issued by public service advertising that they should be screened annually for cancer of the cervix. One cannot be surprised that those American women who could afford it have gone dutifully to their gynecologists for annual tests. They have no interest in dying from cancer of the cervix, and the annual Pap test promises a cure, if cancer is caught early enough.

When the information is not correct, however, the federal government must reassess its policies and protect its citizens from inappropriate use of doubtful procedures. There is now ample evidence that it should not be national policy to screen all women for cervical cancer every year. The time for a policy change is long overdue.

References

Ahluwalia, H. S., and Doll, R. 1968. Mortality from Cancer of the Cervix Uteri in British Columbia and Other Parts of Canada. *British Journal of Preventive and Social Medicine* 22: 161–164.

American Cancer Society. 1968. *Cancer of the Uterus.* New York: American Cancer Society.

———. 1973. *Stay Healthy: Learn About Uterine Cancer.* New York: American Cancer Society.

———. 1975. *1976 Cancer Facts and Figures.* New York: American Cancer Society.

———. 1976. *Cancer Facts for Women.* New York: American Cancer Society.

———. 1977. *1978 Cancer Facts and Figures.* New York: American Cancer Society.

American Cancer Society, Connecticut Division. 1976. *Cytology Services in Connecticut, 1970–1974.* Woodbridge, Conn.: American Cancer Society.

Ask Beth. 1978. See Doctor. *New Haven Register* (April 28): 25.

Benson, W. L., and Norris, H. J. 1977. A Critical Review of the Frequency of Lymph Node Metastasis and Death from Microinvasive Carcinoma of the Cervix. *Obstetrics and Gynecology* 49 (May): 632–638.

Breslow, L. 1977. A History of Cancer Control in the U.S. with Emphasis on the Period 1946–1971. 4 vols. Los Angeles, Calif.: University of California at Los Angeles School of Public Health.

———, and Somers, A. R. 1977. The Lifetime Health-Monitoring Program. *The New England Journal of Medicine* 296: 601–608.

Brindle, G., Wakefield, J. and Yule, R. 1976. Cervical Smears: Are the Right Women Being Examined? *British Medical Journal* 1: 1196–1197.

British Medical Journal. 1973. Editorial. Uncertainties of Cervical Cytology. 4: 501–502.

———. 1976. Editorial. Commentary on Walton Report. 2: 659–660.

Brudnell, M., Cox, B. S. and Taylor, C. W. 1973. The Management of Dysplasia, Carcinoma in Situ and Microcarcinoma of the Cervix. *Journal of Obstetrics and Gynaecology of the British Commonwealth* 80(8): 673–679.

Burghardt, E., and Holzer, E. 1977. Diagnosis and Treatment of Microinvasive Carcinoma of the Cervix Uteri. *Obstetrics and Gynecology* 49 (6): 641–653.

Canadian Task Force. 1976. Cervical Cancer Screening. (Walton Report). *Canadian Medical Association Journal* 114 (June 5): 1003–1033.

Charlson, M. E., and Feinstein, A. R. 1974. Auxometric Dimensions: A New Method for Using Rate of Growth in Prognostic Staging of Breast Cancer. *Journal of the American Medical Association* 228: 180–185.

Christopherson, W. M. 1974. Letter. Mass Screening for Cervical Cancer. *British Medical Journal* 1: 453.

———, Parker, J. E., Mendez, W. M. et al. 1970. Cervix Cancer Death Rates and Mass Cytologic Screening. *Cancer* 26 (4): 808–811.

———, Lundin, F. E., Jr., Mendez, W. M. et al. 1976. Cervical Cancer Control: A Study of Morbidity and Mortality Trends Over a Twenty-One-Year Period. *Cancer* 38 (3): 1357–1366.

Cochrane, A. L., and Holland, W. W. 1971. Validation of Screening Procedures. *British Medical Bulletin* 27 (1): 3–8.

Cocker, J., Fox, H., and Langley, F. A. 1968. Consistency in the Histological Diagnosis of Epithelial Abnormalities of the Cervix Uteri. *Journal of Clinical Pathology,* 21: 67–70.

Cole, P. 1976. Elective Hysterectomy: Pro and Con. *The New England Journal of Medicine* 295 (5): 264–266.

———, and Morrison, A. S. 1978. Basic Issues in Cancer Screening. Paper presented at the International Union Against Cancer Workshop on Screening. Toronto, Canada, April 24–27, 1978.

Commission on Chronic Illness. 1957. *Chronic Illness in the United States.* Vol. 1. Cambridge, Mass: Harvard University Press.

Coppelson, L. W., and Brown, B. 1974. Estimates of the Screening Error Rate From the Observed Detection Rates in Repeated Cervical Cytology. *American Journal of Obstetrics and Gynecology* 119: 953–958.

Cramer, D. W. 1974. The Role of Cervical Cytology in the Declining Morbidity and Mortality of Cervical Cancer. *Cancer* 34 (6): 2018–2027.

Creasman, W. T., and Rutledge, F. 1972. Carcinoma *in-situ* of the Cervix: An Analysis of 861 Patients. *Obstetrics and Gynecology* 39: 373–380.

Danilevicius, Z. 1976. The Walton Report. Editorial. *Journal of the American Medical Association* 236: 2098–2099.

de Toqueville, A. 1955. *Democracy in America.* Vol. 1. New York: Vintage Books.

Dickinson, L. M., Mussey, E., Soule, E. H., and Kurland, L. T. 1972a. Evaluation of the Effectiveness of Cytologic Screening for Cervical Cancer. I. Incidence and Mortality Trends in Relation to Screening. *Mayo Clinic Proceedings* 47 (8): 534–544.

———, Mussey, E., and Kurland, L. T. 1972b. Evaluation of the Effectiveness of Cytologic Screening for Cervical Cancer. II. Survival Parameters Before and After Inception of Screening. *Mayo Clinic Proceedings* 47 (8): 545–549.

———. 1972c. Evaluation of the Effectiveness of Cytologic Screening for Cervical Cancer. III. Cost-Benefit Analysis. *Mayo Clinic Proceedings* 47 (8): 550–555.

Dunn, J. 1977. Interview with Devra Breslow. In Breslow, L. *A History of Cancer Control in the U.S. with Emphasis on the Period 1946–1971.* Vol. 4, Appendix 6, p. 35. Los Angeles, Calif.: University of California at Los Angeles School of Public Health.

Evans, D. M. D., Shelley, G., Cleary, B. et al. 1974. Observer Variation and Quality Control of Cytodiagnosis. *Journal of Clinical Pathology* 27 (12): 945–950.

Feinstein, A. R. 1968. The Identification of Rates of Disease. *Annals of Internal Medicine* 69: 1037–1061.

Feinstein, A. R. 1978. Personal Communication. New Haven, Conn.: Yale School of Medicine.

Frame, P. S., and Carlson, S. J. 1975. A Critical Review of Periodic Health Screening Using Specific Screening Criteria. Part 3: Selected Diseases of the Genitourinary System. *Journal of Family Practice* 2: 189–194.

Frankenburg, W. K., and Camp, B. W., eds. 1975. *Pediatric Screening Tests.* Springfield, Ill.: Charles C Thomas.

Fogarty International Center (National Institutes of Health) and American College of Preventive Medicine. 1976. *Preventive Medicine USA: Theory, Practice and Application of Prevention in Personal Health Services.* New York: PRODIST.

Garrett, W. J. 1964. Symptomless Carcinoma of the Cervix. *Journal of Obstetrics and Gynaecology of the British Commonwealth* 71: 517–528.

Glamour. 1978. Beauty and Health Report. 76 (April): 168.

Grünfeld, K., Horwitz, O., and Lysgaard-Hansen, B. 1975. Evaluation of Mortality Data for Cervical Cancer with Special Reference to Mass Screening Programs, Denmark, 1961–1971. *American Journal of Epidemiology* 101 (4): 265–275.

Hall, J. E., and Walton, L. 1968. Dysplasia of the Cervix. *American Journal of Obstetrics and Gynecology* 100: 662–671.

Hampshire, S. 1976. Thinking About Social Costs. *The New York Times* (March 24): 39.

Hart, C. R., ed. 1975. *Screening in General Practice.* Edinburgh: Churchill Livingstone.

Holland, W. W. 1974. Taking Stock. *Lancet* 1: 1494–1497.

Homesley, H. D. 1977. Evaluation of the Abnormal Pap Smear. *American Family Physician* 16: 190–194.

Hulme, G. W., and Eisenberg, H. S. 1968. Carcinoma in-situ of the Cervix in Connecticut. *American Journal of Obstetrics and Gynecology* 102: 415–425.

Husain, O. A. N. 1976. Quality Control in Cytological Screening for Cervical Cancer. *Tumori* 62: 303–314.

Iglehart, J. 1977. The Cost and Regulation of Medical Technology: Future Policy Directions. *Milbank Memorial Fund Quarterly/Health and Society* 55 (Winter): 25–59.

Kern, W. H., and Zivolich, M. R. 1977. The Accuracy and Consistency of the Cytologic Classification of Squamous Lesions of the Uterine Cervix. *Acta Cytologica* 21 (4): 519–523.

Kinlen, L. J., and Doll, R. 1973. Trends in Mortality from Cancer of the Uterus in Canada and in England and Wales. *British Journal of Preventive and Social Medicine* 27: 146–149.

———, and Spriggs, A. I. 1978. Women with Positive Cervical Smears But Without Surgical Intervention. *Lancet* 2: 463–465.

Kirkland, J. A. 1963. Atypical Epithelial Changes in the Uterine Cervix. *Journal of Clinical Pathology,* 16: 150–154.

Knox, E. G. 1968. Cervical Cancer. In Nuffield Provincial Hospitals Trust, *Screening for Medical Care.* pp. 43–54. London: Oxford University Press.

———. 1976. Ages and Frequencies for Cervical Cancer Screening. *British Journal of Cancer* 34: 444–452.

Kolstad, P. 1970. Diagnosis and Management of Pre-cancerous Lesions of the Cervix Uteri. *International Journal of Gynaecology and Obstetrics* 8: 551–560.

———, and Klem, V. 1976. Long-term and Follow-up of 1121 Cases of Carcinoma in-situ. *Obstetrics and Gynecology* 48 (2): 125–129.

Koss, L. G., and Phillips, A. J. 1974. Summary and Recommendations of the Workshop on Uterine-Cervical Cancer. *Cancer* 33 (6, Suppl.): 1753–1754.

Lambourne, A., and Lederer, H. 1973. Effects of Observer Variations in Population Screening for Cervical Carcinoma. *Journal of Clinical Pathology* 26: 564–569.

Lancet. 1972. Editorial. The Value of Cervical Cytology. 2: 1236–1237.

———. 1976. Editorial. Walton Report on Cervical Screening. 1: 1337.

———. 1978. Editorial. Screening for Cervical Cancer in Young Women. 2: 1029–1030.

Lave, J. R., and Lave, L. B. 1977. Measuring the Effectiveness of Prevention. I. *Milbank Memorial Fund Quarterly/Health and Society* 55 (Spring): 273–289.

Lyon, J. L., and Gardner, J. W. 1977. The Rising Frequency of Hysterectomy: Its Effect on Uterine Cancer Rates. *American Journal of Epidemiology* 105 (5): 439–443.

MacGregor, J. E. 1974. Letter. Screening for Cervical Cancer. *British Medical Journal* 1 (Jan. 12): 77.

———. 1976. Evaluation of Mass Screening Programs for Cervical Cancer in NE Scotland. *Tumori* 62: 287–295.

———, and Teper, S. 1978. Mortality from Carcinoma of Cervix Uteri in Britain. *Lancet* 2 (Oct. 7): 774–776.

McKeown, T. 1968. Validation of Screening Procedures. In Nuffield Provincial Hospitals Trust, *Screening for Medical Care.* pp. 1–13. London: Oxford University Press.

McIndoe, A. A., and Green, G. H. 1969. Vaginal Carcinoma in situ following Hysterectomy. *Acta Cytologica* 13: 158–162.

Miller, A. B., Lindsay, J., and Hill, G. B. 1976. Mortality from Cancer of the Uterus in Canada and its Relationship to Screening for Cancer of the Cervix. *International Journal of Cancer* 17: 600–612.

National Panel of Consultants on the Conquest of Cancer. 1971. *Report prepared for the Committee on Labor and Public Welfare, U. S. Senate.* Document 92-9. Washington, D. C.: U. S. Government Printing Office.

Nelson, J. H. Jr., Averette, H. E., and Richart, R. 1975. *Dysplasia and Early Cervical Cancer.* New York: American Cancer Society.

The New England Journal of Medicine. 1976. Editorial. Papanicolaou Testing—Are We Screening the Wrong Women? 294 (Jan. 22): 223.

Nolen, W. A. 1978. Cutting Your Medical Costs in Half. *McCall's* 105 (August): 78, 172.

Notman, M. T. 1976. Elective Hysterectomy: Pro and Con. *The New England Journal of Medicine* 295 (5): 266–267.

Pedersen, E., Høeg, K., and Kolstad, P. 1971. Mass Screening for Cancer of the Uterine Cervix in Ostfold County, Norway: An Experiment. *Acta Obstetrica et Gynecologica Scandinavica* 11 (Suppl.): 1–18.

Peterson, O. 1956. Spontaneous Course of Cervical Precancerous Conditions. *American Journal of Obstetrics and Gynecology* 72: 1063–1071.

Preston, T. A. 1977. *Coronary Artery Surgery: A Critical Review.* New York: Raven Press.

Randall, K. J. 1974. Cancer Screening by Cytology. *Lancet* 2: 1303–1304.

Rettig, R. A. 1977. *Cancer Crusade: The Story of the National Cancer Act of 1971.* Princeton: Princeton University Press.

Richards, D. H. 1974. A Post-Hysterectomy Syndrome. *Lancet* 2: 983–985.

Robbins, L. 1977. Interview by Lester and Devra Breslow. In Breslow, L. *A History of Cancer Control in the U.S. with Emphasis on the Period 1946–1971.* Vol. 4, Appendix 6, p. 35, Los Angeles, Calif.: UCLA School of Public Health.

Rogers, P. G. 1972. The Need to Emphasize Preventive Medicine in Future U. S. Health Care Programs. *Preventive Medicine* 1 (Dec.): 493–498.

———. 1978. Personal communication. Washington, D.C.: U.S. House of Representatives.

Rosen, G. 1975. *Preventive Medicine.* New York: Science History Publications.

Sackett, D. L. 1975. Screening for Early Detection of Disease: To What Purpose. *Bulletin of the New York Academy of Medicine* 51: 39–52.

Schmidt, O. 1977. Cervical Cancer Screening Program: The SOGC's View. *Canadian Medical Association Journal* 116: 971–972, 975.

Schneider, J., and Twiggs, L. B. 1972. The Costs of Carcinoma of the Cervix. *Obstetrics and Gynecology* 40: 851–859.

Sedlis, A., Walters, A. T., Balin, H. et al. 1974. Evaluation of Two Simultaneously Obtained Cervical Cytological Smears. *Acta Cytologica* 18 (4): 291–296.

Seybolt, J. F., and Johnson, W. D. 1971. Cervical Cyto-diagnostic Problems — A Survey. *American Journal of Obstetrics and Gynecology* 109: 1089–1103.

Shapiro, S. 1975. Screening for Early Detection of Cancer and Heart Disease. *Bulletin of the New York Academy of Medicine* 51: 80–95.

———. 1977. Measuring the Effectiveness of Prevention II. *Milbank Memorial Fund Quarterly/Health and Society.* 55 (Spring): 291–313.

Shulman, J. J., Leyton, M., and Reed, H. 1974. The Papanicolaou Smear: An Insensitive Procedure. *American Journal of Obstetrics and Gynecology* 120: 446–451.

Siegler, E. F. 1956. Microdiagnosis of Carcinoma in-situ of the Uterine Cervix: A Comparative Study of Pathologists' Diagnoses. *Cancer* 9: 463–469.

Silverberg, E. 1975. *Gynecologic Cancer: Statistical and Epidemiological Information.* New York: American Cancer Society.

Spitzer, W. D., and Brown, B. P. 1975. Unanswered Questions About the Periodic Health Examination. *Annals of Internal Medicine* 83: 257–263.

Spriggs, A. I. 1971. Follow-up of Untreated Carcinoma in-situ of Cervix Uteri. *Lancet* 2: 599.

Stern, E., and Neely, P. M. 1963. Carcinoma and Dysplasia of the Cervix: A Comparison of Rates for New and Returning Population. *Acta Cytologica* 7: 357–361.

———, Misczynski, M., Greenland, S. et al. 1977. Pap Testing and Hysterectomy Prevalence: A Survey of Communities with High and Low Cervical Cancer Rates. *American Journal of Epidemiology* 106: 296–305.

Strickland, S. P. 1972. *Politics, Science, and Dread Disease.* Cambridge, Mass: Harvard University Press.

Talebian, F. Shayan, A. Krumholz, B. A. et al. Colposcopic Evaluation of Patients With Abnormal Cervical Cytology. *American Journal of Obstetrics and Gynecology* 49 (6): 670–674.

Timonen, S., Niemenen, V., and Kauraniemi, T. 1974. Cervical Screening. *Lancet* 1: 401.

Truman, D. B. 1958. *The Governmental Process.* New York: Alfred A. Knopf.

U.S. Department of Commerce. 1978. Bureau of the Census, *Current Population Reports.* Report P-25, No. 721 (April). Table 3. Washington, D. C.: U. S. Government Printing Office.

U.S. Department of Health, Education, and Welfare. 1975. Public Health Service, National Center for Health Statistics. *Characteristics of Females Ever Having a Pap Smear and Interval Since Last Pap Smear, United States, 1973.* Monthly Vital Statistics Report, Health Interview Survey, Provisional Data (HRA) 76-1120, Vol. 24, No. 7. (Suppl., October).

———. 1966a. Program Analysis: Disease Control Programs. Cancer. (Cited by Dickinson, L. M. 1972c. Evaluation of the Effectiveness of Cytologic Screening for Cervical Cancer. III. Cost Benefit Analysis. *Mayo Clinic Proceedings* 47 (8): 550–555.

———. 1966b. Public Health Service, National Center for Health Statistics. *Age at Menopause, United States, 1960–1962.* DHEW Publication 1000, Series 11, No. 19. Washington, D.C.: U.S. Government Printing Office.

———. 1978a. SEER Program (Surveillance, Epidemiology and End Results). *Cancer Incidence and Mortality in the United States,*

1973–1976. DHEW Publication No. (NIH) 78-1837. Washington, D.C.: U.S. Government Printing Office.

———. 1978b. *U.S. Vital Statistics: Mortality.* Unpublished Data: Trend C, Tables 292 and 292A, Death and Death Rates, United States, 1976. Washington, D.C.

Wakefield, J., Yule, R., Smith, A. et al. 1973. Relation of Abnormal Cytological Smears and Carcinoma of Cervix Uteri to Husband's Occupation. *British Medical Journal* 2: 142.

Wells, C. K., and Feinstein, A. R. 1977. Routine Radiographic Measurement and Prognostic Importance of Rate of Growth (Auxometry) in Patients With Lung Cancer (Abstract). *Clinical Research* 25: 266A.

Whitby, L. 1974. Definitions and Criteria. *Lancet* 2: 819–821.

Wilson, J. M. G. 1963. Multiple Screening, *Lancet* 2: 51–54.

———. 1968. The Worth of Detecting Occult Disease. In Sharp, C. L. E. H. and Keen, H., eds., *Presymptomatic and Early Diagnosis.* pp. 141–178. London: Pittman.

Wilson, J. M. G., and Jungner, G. 1968. *Principles and Practice of Screening for Disease.* Geneva: World Health Organization.

World Health Organization. 1978. *Steroid Contraception and the Risk of Neoplasia.* Report of a WHO Scientific Group. WHO Technical Report Series G19. Geneva. Switzerland: World Health Organization.

Yule, R. 1972. The Frequency of Cytologic Testing. (Abstract). *Acta Cytologica* 16: 389–390.

An earlier abbreviated version of this paper was presented at the 106th Annual Meeting of the American Public Health Association, Los Angeles, California, October 17, 1978.

Acknowledgments: We wish to thank Alvan R. Feinstein for review of an earlier draft of this paper, and *Milbank Memorial Fund Quarterly* referees for helpful comments.

Address correspondence to: Anne-Marie Foltz, M.P.H., Graduate School of Public Administration, New York University, 40 West 4th Street, New York, New York 10003.