Cervical Cancer Prevention
Making Programs More Appropriate and Pragmatic

Paul D. Blumenthal, MD, MPH
Lynne Gaffikin, DrPH

Cervical cancer is unique for several reasons. It has recognized, well-described risk factors; there is an identifiable precancerous condition; the transition from precancer to cancer occurs over an extended period (10 years on average); screening tests for detecting cancer and precancer exist; and outpatient treatments for precancer are safe, effective, and relatively inexpensive.

Cervical cancer is also a rare, but confirmed, sequela of a sexually transmitted, high-risk human papillomavirus infection. Each year almost 500,000 new cases of cervical cancer are diagnosed and 230,000 deaths occur, 80% of which are in low-resource countries. Thus, more deaths result from human papillomavirus infection and cervical cancer than from sequelae of all other sexually transmitted infections combined, with the recent exceptions of human immunodeficiency virus (HIV) and AIDS.

Although the Papanicolaou test was developed almost 80 years ago, and cytology-based screening programs were initiated more than 50 years ago, cervical cancer still remains the principal cause of cancer deaths among women in developing countries. Even in developed countries, despite long-standing screening programs and technological improvements in both cytological (Papanicolaou) screening and outpatient treatment of precancer that have reduced the burden of disease, elimination of this preventable cancer remains elusive.

Simply put, the majority of cervical cancer cases worldwide occur among women who are never screened or treated. These women tend to be from mostly low-income groups or have other characteristics that place them at highest risk for disease, such as early sexual debut (age < 20 years), high parity (> 4 live births), and a history of multiple sexual partners (> 5). Although cervical cancer screening has received much emphasis and refinements in screening tests render them ever more accurate, screening alone has no intrinsic preventive value. It is only when a positive test result is linked to treatment that disease prevention can occur. Indeed, it is telling that many regional and national programs are called cervical cancer screening as opposed to prevention programs.

In developed countries, cervical cancer “screening” programs commonly mandate multiple, separate visits for screening, confirmation, diagnosis, and treatment. As might be expected, the higher the number of visits required to complete the process, the more likely the cancer prevention loop will not be closed.

Three articles published in this issue of JAMA provide new and practical insights into how cervical cancer prevention efforts could be improved among women at highest risk for disease. The studies by Denny and colleagues and by Brewster and colleagues primarily address never screened or underscreened women who have a relatively high (prior) probability of disease. These reports focus on approaches to strengthen the critical link between testing and treatment of patients with positive test results in settings with higher disease prevalence. In the third article in this issue, Sawaya discusses clinical protocols applied to women with relatively low (prior) probability of disease and focuses on how to improve the clinical utility of testing in a relatively low-risk (or incident disease) population. The themes of program practicality, focusing on high-risk groups, and developing better and more appropriate links between screening and treatment are common features of these 3 articles.

Working with a high-risk, largely Latina population in Southern California, Brewster et al recognized that cervical cancer prevention was severely limited by the mandates of a diagnosis-oriented, multiple-visit approach. In conceptualizing their approach, they emphasize that in a time when treatment of precancerous lesions is available, it is unacceptable for women to develop this disease. The results of their randomized trial testing the feasibility and acceptability of a single-visit program demonstrate that pro...
viding both screening and treatment in 1 visit can enhance the preventive value of a program. Specifically, Brewster et al demonstrate that a single-visit program consisting of a Papanicolaou test and a large loop electrosurgical excision procedure for women with abnormal Papanicolaou test results can successfully recruit women at high risk of cervical cancer for screening and move the highest-risk women through screening, treatment, follow-up, and retreatment if necessary. Importantly, the study demonstrated high levels of satisfaction with an approach that did not require a diagnostic step, such as colposcopy and, if indicated, biopsy.

An important finding of Brewster et al was that by providing treatment based on precancerous findings in a single visit, the women felt better informed about their health status. Among the subgroup of women informed of precancerous findings and treated in the same visit for a high-grade lesion, follow-up (a critical component of prevention) was significantly better than among women who were randomly assigned to usual care (63% vs 21%, respectively; \( P = 0.02 \)). The authors acknowledge that the higher rate of follow-up may be indirectly related to the fact that the women had received treatment. This link is probably more direct and positive. Latina women, the authors note, generally seek health services when they experience symptoms. However, dysplasia is usually asymptomatic. Once in the health care system, these women likely returned because they were well informed about and treated for a potentially precancerous condition. Receiving same-day treatment likely increased their perception that a health problem existed, and they recognized that the preventive loop needed to be closed.

The principal question arising from this study is whether a cytology-based approach such as this is sustainable. Papanicolaou test results for women enrolled in this study were available within 2 hours. Could all patients presenting for Papanicolaou tests obtain their result in 2 hours? Or, would women more likely be triaged in the clinic and Papanicolaou tests processed relative to a woman’s risk status? Triaging by risk status could help improve program efficiency. Furthermore, the advent of convenient, inexpensive, point-of-care, molecular-based tests for detecting cervical precancer will significantly reduce the role of cytology and make a single-visit program similar to that described by Brewster et al eminently more sustainable.

The issue of sustainability has been especially problematic in low-resource settings such as developing countries. In the past several years, resource-limited, developing nations have explored alternative means of cervical cancer prevention through approaches similar to the single-visit program described by Brewster et al without cytology. Cytology-based services involve a level of cost and infrastructure that is neither practical nor sustainable in low-resource settings. As a result, the focus has been on identifying tests that are less costly, have the potential of determining the next management step in a single visit (real-time testing), and are locally sustainable. A demonstration project from rural Thailand indicated that a single-visit approach using visual inspection with acetic acid (VIA) coupled with an immediate link to treatment (cryotherapy) was safe, acceptable, and feasible. A perceived shortcoming of that project was the lack of an observation-only control group. Also, because the project attempted to measure the feasibility of a single-visit protocol, no diagnosis was obtained for women offered cryotherapy. Therefore, the efficacy of this approach compared with alternatives could not be measured.

In the study by Denny et al, women at 3 clinics in peri-urban South Africa were randomly assigned to delayed evaluation or cryotherapy several days after initial evaluation by VIA or human papillomavirus DNA testing, and all women had colposcopy with biopsy of abnormal areas at a 6-month follow-up visit. Thus, Denny et al were able to demonstrate that either VIA or human papillomavirus DNA testing followed up by treatment had a significant effect on the prevalence of high-grade cervical intraepithelial neoplasia (CIN 2+) among participating women. Demonstration of short-term (6-month) efficacy was based on the fact that, after randomization, the 3 study groups (2 treatment, 1 observation) had almost identical risk and demographic profiles and thus, likely identical pretreatment disease profiles. Therefore, differences in the pathological characteristics of individuals in the treated and nontreated groups after treatment could rationally be attributed to the intervention.

Denny et al also provide data on an important safety issue about which there has been little evidence. Although cryotherapy generally is considered safe, it is not clear whether a woman is at increased risk of acquiring HIV infection during the time of cervical healing after cryotherapy. In this study, women were tested for HIV at baseline and at 6-month follow-up, and the rate of HIV acquisition among women who received cryotherapy was not significantly different than among those in the observation-only group (under their power assumption, a meaningful difference was a 2 times greater seroconversion rate). While these results relate only to an observation period of 6 months, subsequent HIV acquisition is unlikely to be ascribable to an effect of cryotherapy based on physiological grounds because the cervix is almost completely reepithelialized by this time. Thus, for clinicians considering a screen-and-treat or single-visit program in areas where HIV prevalence is high, this study provides some important reassurance.

Critics of single-visit programs are concerned that without a diagnostic step, considerable numbers of women who do not actually have disease would be offered treatment. If the treatment was expensive or dangerous, this could be a significant program limitation. However, given the safety of outpatient treatments such as cryotherapy and, to a slightly lesser extent, the loop electrosurgical...
excision procedure, the results of a decision analysis, and the outcomes reported by both Brewster et al and Denny et al, it seems reasonably clear that a single-visit program could be safe, acceptable, and effective. As noted by Denny and colleagues, screen-and-treat is an approach that may prevent a common but preventable cancer, particularly among women in low-resource settings. The evidence from this study supports the implementation of large-scale intervention projects to refine operational aspects of these programs.

The article by Sawaya presents the case of a 21-year-old woman with suspected cervical dysplasia and reviews current, evidence-based protocols for the management of an abnormal Papanicolaou test result. In commenting on both the strengths and challenges of current protocols, however, a number of other messages emerge.

Sawaya aptly observes that “cervical cancer screening in young women can, and often does, elicit anxiety and a cascade of clinical interventions of no clinical value. Until we have more specific screening and triage tests, our best defense against the diagnostic dilemma posed by ASC-US [atypical squamous cells of undetermined significance] in young women is ‘primary prevention of uncertainty’ by exercising restraint and prudence in screening initiation.” Although a policy of restraint and prudence is primarily directed at young women, such as the patient under discussion, this perspective could be applied to any program in which risk status does not play a role in determining screening priorities and management strategies.

Sawaya urges clinicians (and by implication, those designing national programs) to be sensible about when to start screening and how to avoid overmanagement. He also emphasizes that a focus on evolving technology or new clinical protocols will not resolve the problem. What especially links Sawaya to the work of Brewster et al and Denny et al is the point that having screening programs appropriate for the population they address is equally or even more important than the technology applied. Specifically, as noted by Sawaya, attention must be paid to the clinical utility and psychosocial implications of the tests. Test availability does not mandate that the test must be used.

When faced with a challenge, identifying and seizing an opportunity can be the first step toward triumph. That is what public health approaches are often about. An important lesson from the work of Brewster et al and Denny et al is that whether in southern California or southern Africa, safe, acceptable, effective, and pragmatic public health approaches to cervical cancer prevention can be designed. Furthermore, as Sawaya noted, “cervical cancer incidence might differ between strategies if substantial proportions of women managed in certain ways do not follow-up for diagnosis and treatment.” This observation, although not stated as such, provides the rationale underlying single-visit or screen-and-treat approaches.

The advent of a vaccine to prevent human papillomavirus infection, widely perceived to be just around the corner, is eagerly awaited to advance preventive capabilities. Recent data reported by Skjeldstad and colleagues made this development increasingly likely. However, even if a vaccine became available tomorrow, several generations of women worldwide would still need conventional care. To provide care to all women, future cervical cancer prevention programs worldwide will need to be designed to reach women who are at highest risk for disease and to ensure that the critical components of testing, treatment, and follow-up are realized as appropriate for every woman. Programs that build on the experience reported in these articles will have a good start in achieving those goals.

Financial Disclosures: None reported.

REFERENCES
When (Not) to Stop a Clinical Trial for Benefit

Stuart J. Pocock, PhD

In this issue of JAMA, Montori and colleagues1 provide a valuable extensive and critical systemic review of clinical trials that were stopped early for benefit. Readers of the reports of such trials often feel a sense of excitement, especially when phrases such as “a major treatment advance,” “ethical need to stop the inferior treatment,” and “vital to tell the world immediately” are used. However, experience suggests that early results and enthusiasm, especially for modestly sized trials terminated early for apparent major benefit, are often moderated as subsequent reports arise.2

The skeptic should ask first whether correct and appropriate structures were in place for analyzing and reviewing, and making decisions based on, the trial’s accumulating interim data. Having the members of an effective independent data monitoring committee (DMC) or data and safety monitoring board as the only individuals accessing and interpreting interim data split by treatment group is now considered an essential part of good practice for major randomized trials.3,4 Still, a substantial minority of reported major trials appear not to have a DMC in place.5

Second, with or without a formal DMC recommendation, another question is whether the decision to stop a trial early and report the results was an appropriate judgment. This decision should be aided by a predefined statistical stopping boundary for a primary outcome,6,7 but some trials have no such guideline. It is important that such a boundary is sufficiently stringent (eg, very strong evidence of a treatment difference with a very small P value) to match the ethical and public health implications of a decision to stop the trial. In a spirit of requiring proof beyond reasonable doubt that a treatment difference is sufficient to affect future clinical practice, some lenient statistical boundaries are not a sensible choice in the direction of benefit. For instance, the so-called Pocock boundary8 and the O’Brien-Fleming boundary’s last interim look9 both typically require values around P= .02 for stopping, which is usually insufficient strength of evidence to stop a trial for benefit. Both boundaries can be made more appropriate if the overall type I error is set at 1% rather than the conventional 5%.

Many complex methods exist for statistical stopping boundaries, whereas in practice there is considerable merit in the simple Haybittle-Peto boundary,9 which requires P<.001 as evidence required to consider stopping a trial early for benefit. Even so, such a boundary should not be applied too soon, when few outcome events have been observed.

Decisions on early stopping (or not) need to be based on wise judgments interpreting the totality of available evidence, both in the current trial (considering primary and other efficacy outcomes and safety issues) and in other external evidence (especially from related trials).10 Accordingly, a statistical stopping boundary is only one useful objective component in an inevitably more challenging decision-making process. The ethical dilemma is to safeguard the interests of patients randomized in the current trial while also protecting society from overzealous premature claims of treatment benefit.11 For instance, if a trial is evaluating a treatment meant to be given long-term for conditions such as hypertension or chronic arthritis, short-term benefits, no matter how statistically significant, may not merit early stopping. If a trial is for regulatory approval, the sponsor and trialists should be encouraged not to stop early unless there is overwhelming evidence of treatment superiority, since the regulators require substantial evidence of both efficacy and safety monitoring board as the only individuals accessing and interpreting interim data. Having the members of an effective independent data monitoring committee (DMC) or data and safety monitoring board as the only individuals accessing and interpreting interim data split by treatment group is now considered an essential part of good practice for major randomized trials. Still, a substantial minority of reported major trials appear not to have a DMC in place.

Second, with or without a formal DMC recommendation, another question is whether the decision to stop a trial early and report the results was an appropriate judgment. This decision should be aided by a predefined statistical stopping boundary for a primary outcome, but some trials have no such guideline. It is important that such a boundary is sufficiently stringent (eg, very strong evidence of a treatment difference with a very small P value) to match the ethical and public health implications of a decision to stop the trial. In a spirit of requiring proof beyond reasonable doubt that a treatment difference is sufficient to affect future clinical practice, some lenient statistical boundaries are not a sensible choice in the direction of benefit. For instance, the so-called Pocock boundary and the O’Brien-Fleming boundary’s last interim look both typically require values around $P = .02$ for stopping, which is usually insufficient strength of evidence to stop a trial for benefit. Both boundaries can be made more appropriate if the overall type I error is set at 1% rather than the conventional 5%.

Many complex methods exist for statistical stopping boundaries, whereas in practice there is considerable merit in the simple Haybittle-Peto boundary, which requires $P<.001$ as evidence required to consider stopping a trial early for benefit. Even so, such a boundary should not be applied too soon, when few outcome events have been observed.

Decisions on early stopping (or not) need to be based on wise judgments interpreting the totality of available evidence, both in the current trial (considering primary and other efficacy outcomes and safety issues) and in other external evidence (especially from related trials). Accordingly, a statistical stopping boundary is only one useful objective component in an inevitably more challenging decision-making process. The ethical dilemma is to safeguard the interests of patients randomized in the current trial while also protecting society from overzealous premature claims of treatment benefit. For instance, if a trial is evaluating a treatment meant to be given long-term for conditions such as hypertension or chronic arthritis, short-term benefits, no matter how statistically significant, may not merit early stopping. If a trial is for regulatory approval, the sponsor and trialists should be encouraged not to stop early unless there is overwhelming evidence of treatment superiority, since the regulators require substantial evidence of both efficacy and safety.