In this issue of the Journal, Andrae et al. (1) present the first audit of a national screening program for cervical cancer, and they should be congratulated for their efforts. In the United Kingdom, such audits were proposed in 1986 (2) and pilot audits were first reported in 1996 (3), but audits have yet to be translated into a routine national activity. Screening is a large-scale, repetitive “industrial process” and, as with all other such processes, one can learn most about its performance by examining the failures—here defined as a woman who is eligible for screening and who develops a potentially fatal cervical cancer—and then retuning the process accordingly.

To this end, the screening histories of women who developed cervical cancer can be compared with those of a population-based age-matched sample of women who were eligible for screening but did not develop cancer. Three broad groups of screening failures can be identified: 1) women who were not screened within the recommended interval; 2) women who were screened and found to have an abnormality, but who subsequently developed cancer; and 3) women who were adequately screened within the recommended interval with apparently normal results, but who subsequently developed cancer. Each of these broad groups represents a different type of failure and indicates the need for a different sort of remedial action.

The first group of screening failures is a result of inadequate coverage and needs to be examined separately for women of screening age and for those beyond the recommended age for routine screening. For women younger than 65 years (5 years older than the age at which the last screen is recommended in Sweden), coverage within the recommended interval was reported by Andrae et al. (1) to be 68% for population-based control subjects but only 49% for subjects who were diagnosed with cervical cancer (case subjects), leading to an odds ratio of 2.21. Little difference in the protection afforded by screening was seen between women younger than 30 years and those aged 30–65 years. Not unexpectedly, the rates of screening within the past 5 years were much lower for women older than 65 years (ie, only 8% for case subjects and 19% for control subjects). The fact that 32% of all cancers—and an even greater percentage of late-stage tumors—appeared in this age group suggests that it might be useful to extend routine screening to women who are older than 60 years.

Because a large proportion of microinvasive (ie, stage IA) cancers are screen detected and have a very favorable prognosis, a detailed analysis of screening performance that excludes stage IA case subjects (and their matched control subjects) would also be of interest. Women diagnosed with microinvasive cancer can arguably be considered screening successes because these cancers are virtually 100% curable with minimal morbidity and very rarely lead to fatality. Andrae et al. (1) have provided some analyses indicating that women who were screened within the recommended interval have greater protection against localized (stage IB) and especially more advanced (stage II or higher) cancers than that obtained for microinvasive (stage IA) cancers [odds ratios for being screened within the recommended interval = 0.59, 0.48, and 0.21 for stage IA, IB, and II or higher, respectively; Table 3 (1)]. Ideally, this analysis should be extended to examine different age groups separately because many of the late-stage cancers occur at ages beyond that recommended for screening. Also, as suggested by figure 1 of Andrae et al. (1), the proportion of microinvasive cancers decreases dramatically with increasing age, from approximately 50% of the cancers in women aged 20–29 years to less than 10% of those in women older than 50 years, with a concomitant increase in the proportion of late-stage tumors. This association of microinvasive tumors with young age could explain why the overall effectiveness screening for women aged 20–29 was found to be similar to that for older women, in contradistinction to the findings in the UK study (4) and elsewhere (5). In the UK case–control study (4), only stage IB or more advanced tumors were considered, and screening was not found to be very effective in the youngest women. It would be of great interest to see the Swedish data broken down by age for stage IB or higher (and stage II) tumors to see if the protection obtained against these more often fatal cancers is apparent in this youngest age group.

The second group identified above consists of women who developed cancer despite having an abnormal smear more than 6 months before diagnosis. This situation occurred among 11% of the case subjects (35% of whom had a biopsy) but only 3% of the population-based control subjects, corresponding to a relative risk of 3.67. Here the failure was in follow-up and was of two main types: those with a subsequent biopsy and those without one. For women who did not have a biopsy, the failure reflects a mixture of women who did not attend an invitation for a repeat screening, women whose repeat smears were negative, and women who underwent colposcopy but for whom no lesion was visualized. It would be useful to identify how many women were in each group and, particularly for the last group, to know more about the length of time between the abnormal smear and the diagnosis of cancer.
as well as the number of subsequent positive and negative smears. Information on the grade of cytological abnormality would also be useful. For women with a biopsy, failures need to be split according to the histological outcome. For women with a negative (ie, normal) biopsy, the failure is in detecting clinically significant disease; for women with a positive biopsy, it indicates a failure or lack of appropriate treatment.

The third major group consists of women who developed cancer despite having a normal smear within the recommended interval and represents a failure of the screening test. In the study by Andrae et al. (1), this group included 24% of all cases, but the percentage is likely to be closer to 40% for women aged 65 years or less because for women in that age group 51% of the cases occurred in women who were screened within the recommended interval (regardless of outcome). Different questions arise here that are best answered by a review of the cytology slides from these case subjects (and their matched control subjects) to see if abnormalities were missed or if nothing was detectable. This is also an ideal group to see if new technologies—such as automated cytological reading or human papillomavirus testing of the cells on the slide after electronic preservation of the cytological image—would have identified an abnormality.

Audits, such as the one described by Andrae et al. (1), need to become routine within screening programs if screening is to achieve its full potential. As such, they need to be streamlined and to focus on key issues. If audits are to be done routinely, histological review of the cancers may not be necessary, but a review of the apparently normal cytology slides of women who develop cancer (and for matched control subjects) would be highly desirable. Audits should incorporate stage information and exclude microinvasive cancers from the main analysis. Because mortality reduction is the ultimate goal of screening, analytic methods that weight incident cases according to their mortality potential (eg, using stage), as proposed (6) for the analysis of randomized trial data, need to be extended to case–control studies. This approach can be achieved, for example, by weighting case–control sets according to their 10-year probability that the cancer will result in death. Such an analytic approach would more fully account for benefits of screening in situations where screening leads to both a reduction in cancer incidence and a downstaging of the cancers detected. Case–control studies are not as reliable as randomized trials because of the potential bias that women who attend for screening may have a different underlying cancer risk than those who do not. However, this problem appears to be minimal in the analysis of screening efficacy, and it is clear from this and other studies that much useful information can be obtained by this approach.

This important paper by Andrae et al. (1) further emphasizes the value of conducting screening via organized programs, so that a full evaluation of the degree of success can be determined and areas for further improvement can be identified. In places where population-based cancer registration exists, even nonorganized programs could achieve some degree of evaluation by requiring details of all screening episodes in the last 5 years to be provided to the registry as part of the reporting process.

This audit process allows evaluation of routine service screening, as opposed to extrapolation from clinical trials, and should be widely emulated for all types of mass screening programs and not only restricted to cervical screening.

References