

480 never become clinically significant or cause death. This problem is compounded by the fact that the most common cancers appear most frequently at ages when competing causes of death are more frequent.

Selection bias must be considered in assessing the results of any screening effort. The population most likely to seek screening may differ from the general population to which the screening test might be applied. In general, volunteers for studies are more health conscious and likely to have a better prognosis or lower mortality rate, irrespective of the screening result. This is termed the *healthy volunteer effect*.

Potential Drawbacks of Screening Risks associated with screening include harm caused by the screening intervention itself, harm due to the further investigation of persons with positive tests (both true and false positives), and harm from the treatment of persons with a true-positive result, whether or not life is extended by treatment (e.g., even if a screening test reduces relative cause-specific mortality by 20–30%, 70–80% of those diagnosed still go on to die of the target cancer). The diagnosis and treatment of cancers that would never have caused medical problems can lead to the harm of unnecessary treatment and give patients the anxiety of a cancer diagnosis. The psychosocial impact of cancer screening can also be substantial when applied to the entire population.

Assessment of Screening Tests Good clinical trial design can offset some biases of screening and demonstrate the relative risks and benefits of a screening test. A randomized controlled screening trial with cause-specific mortality as the endpoint provides the strongest support for a screening intervention. Overall mortality should also be reported to detect an adverse effect of screening and treatment on other disease outcomes (e.g., cardiovascular disease). In a randomized trial, two like populations are randomly established. One is given the usual standard of care (which may be no screening at all) and the other receives the screening intervention being assessed. The two populations are compared over time. Efficacy for the population studied is established when the group receiving the screening test has a better cause-specific mortality rate than the control group. Studies showing a reduction in the incidence of advanced-stage disease, improved survival, or a stage shift are weaker (and possibly misleading) evidence of benefit. These latter criteria are early indicators but not sufficient to establish the value of a screening test.

Although a randomized, controlled screening trial provides the strongest evidence to support a screening test, it is not perfect. Unless the trial is population-based, it does not remove the question of generalizability to the target population. Screening trials generally involve thousands of persons and last for years. Less definitive study designs are therefore often used to estimate the effectiveness of screening practices. However, every nonrandomized study design is subject to strong confounders. In descending order of strength, evidence may also be derived from the findings of internally controlled trials using intervention allocation methods other than randomization (e.g., allocation by birth date, date of clinic visit); the findings of analytic observational studies; or the results of multiple time series studies with or without the intervention.

Screening for Specific Cancers Screening for cervical, colon, and breast cancer is beneficial for certain age groups. Depending on age and smoking history, lung cancer screening can also be beneficial in specific settings. Special surveillance of those at high risk for a specific cancer because of a family history or a genetic risk factor may be prudent, but few studies have assessed the influence on mortality. A number of organizations have considered whether or not to endorse routine use of certain screening tests. Because these groups have not used the same criteria to judge whether a screening test should be endorsed, they have arrived at different recommendations. The American Cancer Society (ACS) and the U.S. Preventive Services Task Force (USPSTF) publish screening guidelines (Table 100-3); the American Academy of Family Practitioners (AAFP) generally follow/endorse the USPSTF recommendations; and the American College of Physicians (ACP) develops recommendations based on structured reviews of other organizations' guidelines.

BREAST CANCER Breast self-examination, clinical breast examination by a caregiver, mammography, and magnetic resonance imaging (MRI) have all been variably advocated as useful screening tools.

A number of trials have suggested that annual or biennial screening with mammography or mammography plus clinical breast examination in normal-risk women older than age 50 years decreases breast cancer mortality. Each trial has been criticized for design flaws. In most trials, breast cancer mortality rate is decreased by 15–30%. Experts disagree on whether average-risk women age 40–49 years should receive regular screening (Table 100-3). The U.K. Age Trial, the only randomized trial of breast cancer screening to specifically evaluate the impact of mammography in women age 40–49 years, found no statistically significant difference in breast cancer mortality for screened women versus controls after about 11 years of follow-up (relative risk 0.83; 95% confidence interval 0.66–1.04); however, <70% of women received screening in the intervention arm, potentially diluting the observed effect. A meta-analysis of eight large randomized trials showed a 15% relative reduction in mortality (relative risk 0.85; 95% confidence interval 0.75–0.96) from mammography screening for women age 39–49 years after 11–20 years of follow-up. This is equivalent to a number needed to invite to screening of 1904 over 10 years to prevent one breast cancer death. At the same time, nearly half of women age 40–49 years screened annually will have false-positive mammograms necessitating further evaluation, often including biopsy. Estimates of overdiagnosis range from 10 to 40% of diagnosed invasive cancers. In the United States, widespread screening over the last several decades has not been accompanied by a reduction in incidence of metastatic breast cancer despite a large increase in early-stage disease, suggesting a substantial amount of overdiagnosis at the population level.

No study of breast self-examination has shown it to decrease mortality. A randomized controlled trial of approximately 266,000 women in China demonstrated no difference in breast cancer mortality between a group that received intensive breast self-exam instruction and reinforcement/reminders and controls at 10 years of follow-up. However, more benign breast lesions were discovered and more breast biopsies were performed in the self-examination arm.

Genetic screening for *BRCA1* and *BRCA2* mutations and other markers of breast cancer risk has identified a group of women at high risk for breast cancer. Unfortunately, when to begin and the optimal frequency of screening have not been defined. Mammography is less sensitive at detecting breast cancers in women carrying *BRCA1* and *BRCA2* mutations, possibly because such cancers occur in younger women, in whom mammography is known to be less sensitive. MRI screening may be more sensitive than mammography in women at high risk due to genetic predisposition or in women with very dense breast tissue, but specificity may be lower. An increase in overdiagnosis may accompany the higher sensitivity. The impact of MRI on breast cancer mortality with or without concomitant use of mammography has not been evaluated in a randomized controlled trial.

CERVICAL CANCER Screening with Papanicolaou (Pap) smears decreases cervical cancer mortality. The cervical cancer mortality rate has fallen substantially since the widespread use of the Pap smear. With the onset of sexual activity comes the risk of sexual transmission of HPV, the fundamental etiologic factor for cervical cancer. Screening guidelines recommend regular Pap testing for all women who have reached the age of 21 (prior to this age, even in individuals that have begun sexual activity, screening may cause more harm than benefit). The recommended interval for Pap screening is 3 years. Screening more frequently adds little benefit but leads to important harms, including unnecessary procedures and overtreatment of transient lesions. Beginning at age 30, guidelines also offer the alternative of combined Pap smear and HPV testing for women. The screening interval for women who test normal using this approach may be lengthened to 5 years.

An upper age limit at which screening ceases to be effective is not known, but women age 65 years with no abnormal results in the previous 10 years may choose to stop screening. Screening should be