3

Technologies To
Detect Prostate Cancer

*t*he most commonly used technologies for detecting and diagnosing prostate cancer are digital rectal examination (DRE), prostate-specific antigen (PSA) measurement, transrectal ultrasound (TRUS), and transrectal needle biopsy of the prostate (TRNB). For primary-care based case-finding and mass screening, TRUS and TRNB would be logistically difficult to include as primary screening tests given their relative complexity and invasive nature. Moreover, the marginal value of TRUS above DRE and PSA seems to be small (18, 91, 215), and the risk and discomfort of TRNB would seem to obviate its use as a primary screening test. Therefore, this chapter considers the use of DRE and/or PSA as primary screening tests, and TRUS and TRNB as followup, confirmatory tests.

To analyze the impact of screening, it is necessary to know the "operating characteristics" of each screening technology. In general, the operating characteristics, which refer to the ability of a test to find all cancers that would cause harm and to find only those cancers, are expressed in terms of the sensitivity and specificity of the test. (Box 3-1 describes these concepts.) Unfortunately, the "true" operating characteristics of DRE and PSA cannot be defined since few studies have evaluated them in populations where the true underlying prevalence of

clinically-significant prostate cancer is known. The fact that small volume, well-differentiated cancers should be considered as "nondisease" and that it is relatively easy to detect advanced cancer which may offer no therapeutic benefit further complicates the design and analysis of these studies.
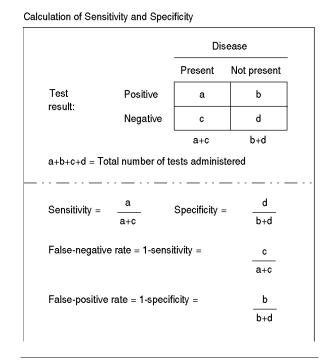
What are usually available are studies of the "positive predictive value" of tests, the proportion of positive or suspicious test results that ultimately turn out to be cancer (see box 3-1); in these studies, patients with "negative" test results do not receive followup TRNB (even though they may harbor significant prostate cancers that the screening test did not find). Furthermore, these studies use different combinations of primary screening tests and different strategies of followup evaluation. Finally, the studies do not uniformly provide age-specific predictive values, which are important to an analysis of screening older men.

To overcome these problems, this analysis presents "likelihood ratios" of disease (292) for DRE and for PSA. These likelihood ratios are estimates of how many times more likely a patient with a particular test result is to have a given type of cancer than if the patient did not have the test. The probabilities of cancer with no test are the prevalence estimates found in table 2-5. Appendix C

BOX 3-1: DESCRIBING THE ACCURACY OF SCREENING TESTS

To analyze the impact of a screening program, it is necessary to understand the accuracy of each screening technology, sometimes referred to as the "operating characteristics" of the test. These operating characteristics, which include the ability of a test to find all existing disease and to find only disease, are usually expressed in terms of the test's sensitivity and specificity. **Sensitivity** is the percentage of all screened people with disease who test positive, while **specificity** is the percentage of all healthy screened people who test negative. In other words, sensitivity is the ability of a test to find people with disease, while specificity represents the test's ability to label healthy people correctly. These characteristics relate inversely to the **false-positive** rate (the percentage of people free of disease who test positive) and the false-negative rate (the percentage of people afflicted by the disease whose screening results are negative). For example, a test with sensitivity between 70 and 95 percent would have a false-negative rate of 5 to 30 percent. The figure below displays the calculation of sensitivity and specificity and the relationship of these indicators to false-positive and false-negative rates.

CALCULATION OF SENSITIVITY AND SPECIFITY

Calculation of Sensitivity and Specificity

		Disease	
		Present	Not present
Test result:	Positive	a	b
	Negative	c	d
		a+c	b+d

a+b+c+d = Total number of tests administered

$$\text{Sensitivity} = \frac{a}{a+c} \qquad \text{Specificity} = \frac{d}{b+d}$$

$$\text{False-negative rate} = 1 - \text{sensitivity} = \frac{c}{a+c}$$

$$\text{False-positive rate} = 1 - \text{specificity} = \frac{b}{b+d}$$

SOURCE: Office of Technology Assessment, 1990.

Calculating sensitivity and specificity requires that one know the true underlying prevalence of disease in the screened population, regardless of screening test results. In other words, it would require performing definitive followup tests on all screenees, even those whose screening test is negative. This is usually not done in studies of prostate cancer screening because of the invasiveness, costs, and risks of such followup procedures (usually transrectal needle biopsies). Hence, most studies report a less useful measure of a screening technology's accuracy, **the positive predictive value (PPV).** The PPV is the percentage of people with positive test results who ultimately turn out to have cancer. Conversely, the **negative predictive value (NPV)**, is the percentage of people with negative test results who ultimately turn out to be free of disease. Calculation of PPV does not require knowing the true underlying prevalence of disease among all people screened. The PPV for a specific condition is directly related to the prevalence of the condition being screened for and, all else being equal, is inversely related to the false-positive rate. A low PPV usually indicates a high false-positive rate, although it is sometimes possible to have both a low PPV and a low false-positive rate. This occurs if the disease is rare. With rare conditions, because the prevalence of a previously undetected disease would decrease as the frequency of testing increases, prolonged studies implementing periodic rescreening normally yield declining PPVs as the studies progress.

The PPV is a limited measure of screening accuracy. In most circumstances a low PPV indicates that for every cancer detected a substantial number of individuals undergo the risks and costs associated with followup testing. However, policymakers or clinicians may decide that reductions in mortality and morbidity associated with screening in a population are large enough to justify the risks and costs associated with screening and followup among healthy individuals. The uncertainty concerning whether this is true for prostate cancer screening is a major issue in the analysis presented in this background paper.

SOURCE: Office of Technology Assessment, 1995.

discusses the methods used in making these estimates. The estimates themselves are presented in the sections on DRE and PSA respectively below.[1]

A potential problem with these estimates is that the positive predictive value in different studies depends heavily on the aggressiveness of the followup strategy employed for a suspicious test. Studies tend to find more cancer by performing multiple systematic biopsies (and even repeated sets of multiple systematic biopsies) in response to a suspicious primary test (70). Using this methodology, a test that has poor sensitivity and specificity but is "positive" in a large proportion of the population will appear to perform well if one examines only the predictive value of the strategy. For example, a strategy of performing multiple sets of biopsies on all men with brown eyes would probably have a rather high "yield" in terms of the number of prostate cancers detected, despite eye color having no information value as a test for prostate cancer. Eye color, in essence, becomes a lottery for receiving the more accurate diagnostic test, TRNB. A recent study of DRE and PSA suggests that this phenomenon occurs with prostate cancer screening (72, 123). Although the predictive value of a suspicious DRE in this study was about 22 percent (72), the percentage of palpably suspicious quadrants of the prostate that yielded cancer was only about 11 percent, implying that roughly half the cancers found as a result of selecting patients for biopsy based on a suspicious DRE were actually found elsewhere in the prostate as a result of the systematic biopsy.

DIGITAL RECTAL EXAMINATION

The digital rectal examination, in which the clinician attempts to feel abnormalities in the size or shape of the prostate gland through the rectum, is a time-honored test for the early detection of prostate cancer despite very weak agreement among published guidelines about its value (100). The DRE is limited in sensitivity because of an inability to detect tumors deep within the prostate gland. Because larger tumors are easier to feel, DRE is unlikely to detect insignificant cancers (although this risk will increase if a suspicious DRE triggers a set of systematic biopsies in addition to a biopsy of the suspicious area). The detection of larger cancers also means that a relatively high percentage of DRE-detected tumors (half or more) will have already spread beyond the confines of the prostatic capsule (139, 279, 271). Many investigators have been concerned about variation among physicians in their ability to detect cancers by DRE (271), especially the possibility that DREs performed by primary care physicians may not be as discriminating as urologists' exams. However, little empirical evidence exists to address this concern (354).

Appendix C lists studies of primary DRE screening for prostate cancer, with brief descriptions of study methods and results. Comparisons are difficult given different patient populations, different thresholds for calling a DRE "suspicious," and different strategies of followup testing. One study by Chodak and colleagues (79) provides the most detailed presentation, and allows es-

---

[1] This method is methodologically inferior to knowing the underlying disease state of all individuals in each study, but probably superior to the alternative methods used in the screening literature, such as screening a population with multiple modalities (often DRE, PSA, and TRUS) and assuming all clinically significant cancers have been detected, or testing only patients with documented clinical disease status (e.g., men scheduled for radical prostatectomy for known cancer). The former method overestimates sensitivity and specificity since some clinically significant cancers would likely be undetected by all modalities; the latter method overestimates sensitivity if cancers in the tested population are more advanced than those that would be identified by screening, or if the screening test were actually used in the process of identifying them in the first place.

TABLE 3-1: ESTIMATED LIKELIHOOD RATIOS FOR RESULTS OF DIGITAL RECTAL EXAMINATION CHANGING THE ODDS OF SIGNIFICANT[a] PROSTATE CANCER (>0.5mL) OF DIFFERENT PATHOLOGIC EXTENTS[b]

DRE result	Likelihood Ratio	
	Intracapsular cancer	Extracapsular cancer
"Suspicious"		
Chodak (1989)[c]	1.5	8.6
Richie (1993)[d]	2.0	2.7
"Nonsuspicious"		
Chodak (1989)	0.96	0.53
Richie (1993)	0.83	0.72

[a] Probability of prostate cancer <0.5mL = 11% based on J.E. Oesterling, V.J. Suman, H. Zincke et al., "PSA-Detected (Clinical Stage T1c or BO) Prostate Cancer: Pathologically Significant Tumors," *Urologic Clinics of North America* 17:719-737, 1990.

[b] See appendix C for methods deriving these estimates.

[c] G.W. Chodak, P. Keller, and H.W. Schoenberg, "Assessment of Screening for Prostate Cancer Using the Digital Rectal Examination," *Journal of Urology* 141:1136-1138, 1989.

[d] J.P. Richie, W.J. Catalona, F.R. Ahmann, et al., "Effect of Patient Age on Early Detection of Prostate Cancer with Serum Prostate-Specific Antigen and Digital Rectal Examination," *Urology* 42:365-374, 1993.

SOURCE: Office of Technology Assessment, 1995. Based on information from M.J. Barry, C.M. Coley, C. Fleming et. al, "The Safety, Effectiveness, and Cost of Early Detection and Treatment of Prostate Cancer Among Older Men: A Report to the Congressional Office of Technology Assessment" OTA contract paper no. K3-0546.0 Massachusetts General Hospital, Boston, MA June 30, 1994.

timation of the likelihood of cancers with and without capsular penetration (table 3-1) for each DRE test result.[2] Appendix B discusses the methods used to produce these estimates. No clinical trials of the use of DRE alone for the early detection of prostate cancer are available. However, neither a case-control study (129) nor a decision model (241) has suggested an important survival benefit for men screened with DRE.

PROSTATE-SPECIFIC ANTIGEN

Prostate-specific antigen is a glycoprotein produced in the prostate gland with a probable role in the transport of semen. Because cancerous prostate tissue, gram for gram, produces greater quantities of PSA than does normal or benignly enlarged tissue, and because prostate cancer may increase the likelihood that PSA "leaks" into the general circulatory system, serum (blood) PSA levels appear to have some discriminating capacity for prostate cancer (99, 257). Preliminary evidence suggests prostate cancers need to be greater than 1 mL in volume before they cause an increase in serum PSA (49).

Three PSA assays have been commonly used clinically and described in the literature (172). Hybritech's Tandem PSA assays detect PSA with monoclonal antibody

[2] In a more recent study, with a policy of systematic biopsy for abnormal DRE results, 15 percent of 6,630 male volunteers over age 50 had an abnormal DRE, and 21 percent of the men with an abnormal DRE had cancer at biopsy; the overall detection rate of cancer for DRE in this series was 3.2 percent, reflecting the more aggressive use of biopsies (72). A new followup study has suggested better outcomes for men diagnosed at initial rather than followup screening with DRE (139); this finding may represent the effect of length bias with one-time screening (discussed in chapter 2).

probes; these assays use radioactive antibodies and enzymatic reactions to perform the measurement. The Tandem PSA tests are currently the only assays approved by the U.S. Food and Drug Administration (FDA) for use in conjunction with DRE as an aid in the detection of prostate cancer in men over age 50.[3] Abbott's IMx PSA assay uses a microparticle enzyme immunoassay technique. Yang's Pros-Check PSA assay uses a polyclonal antibody probe to measure PSA (356). The levels of PSA measured by the Hybritech and Abbott assays appear roughly similar (190, 355), while the polyclonal assay runs values about 1.6-fold higher (148, 339). However, investigators have recently raised concerns about the calibration of the Hybritech and Abbott assays (48, 149, 226, 266), which together dominate the PSA assay market. Clinicians need to know which test their laboratory uses, and to consider a switch in assays in the "differential diagnosis" of a changing PSA in a given patient.

One potential difficulty with this screening test is that factors other than prostate cancer can temporarily elevate PSA levels for several weeks: acute inflammation of the prostate (prostatitis), acute urinary retention, a diagnostic medical procedure called rigid cystoscopy, TRUS, TRNB, or prostate surgery (193, 262). A recent study has also found temporary elevations in PSA following ejaculation (250). However, several studies have now documented that there is no clinically important elevation in PSA values following routine DRE (95, 371), an important finding since physicians often perform DRE and PSA at the same visit.

TABLE 3-2: PROPOSED AGE-SPECIFIC NORMAL REFERENCE RANGES FOR PROSTATE-SPECIFIC ANTIGEN MEASUREMENTS

Age	Normal reference range (ng/mL)	
	Oesterling, 1993c [a]	Dalkin, 1993[b]
40-49	0 - 2.5	-
50-59	0 - 3.5	0 - 3.5
60-69	0 - 4.5	0 - 5.4
70-79	0 - 6.5	0 - 6.3

[a] J.E. Oesterling, S.J. Jacobsen, C.G. Chute, et al., "Serum Prostate-Specific Antigen in a Community-Based Population of Healthy Men: Establishment of Age-Specific Reference Ranges," *Journal of the American Medical Association.* 270:860-864, 1993.

[b] B.L. Dalkin, F.R. Ahmann, and J.B. Kopp, "Prostate Specific Antigen Levels in Men Older Than 50 Years Without Clinical Evidence of Prostatic Carcinoma," *Journal of Urology* 150:1837-1839, 1993.

SOURCE: Office of Technology Assessment, 1995. Based on information from M.J. Barry, C.M. Coley, C. Fleming et. al, "The Safety, Effectiveness, and Cost of Early Detection and Treatment of Prostate Cancer Among Older Men: A Report to the Congressional Office of Technology Assessment", OTA contract paper no. K3-0546.0, Massachusetts General Hospital, Boston, MA June 30, 1994.

Most studies consider an Abbott or Hybritech PSA level up to 4.0 nanograms per milliliter of serum (ng/mL) (equivalent to a Yang PSA level up to 7 ng/mL) as nonsuspicious (148, 339).[4] However, "normal" PSA values increase as a man ages, reflecting the increasing size of the prostate with age (88). Two recent articles have proposed age-specific reference ranges for normal PSA values (table 3-2). One study used the 95th percentile of serum PSA among men without evidence of prostate cancer as the upper boundary of the reference range

[3] The FDA approved the Tandem PSA assays for detection on August 25, 1994. The Tandem tests, the Abbott IMx, the Toschmedix, AIA pack, and the Ciba-Corning ACS assays are all approved for monitoring men with previous prostate problems (228).

[4] Some investigators prefer a lower threshold on the Abbott or Hybritech assays of 3.0 ng/mL to improve test sensitivity (201). For a given underlying prevalence of true cancer, lowering the threshold increases the proportion of all true cancers found by screening, but at the cost of having to do more biopsies (which, as described later in this paper carries cost and risk in itself) and an increased number of false-positive screening results. In other words, in setting the threshold for conducting a biopsy, there is a tradeoff between false-negative and false-positive test results.

(260, 261), while the other used a slightly different, but methodologically similar approach[5] to define the upper limit (101).[6] Another recent study compared the performances of several PSA test kits as part of an international PSA standardization conference (329).

Appendix D lists published studies that use PSA as the primary screening tool to detect prostate cancer (DRE used only to followup a suspicious PSA).[7] Although these studies generally have a somewhat higher proportion of subjects with a cancer detected than do the studies of primary DRE, these proportions are likely underestimates of the maximal attainable yield since patients were often not biopsied unless a followup DRE or TRUS was also suspicious. Using data from the Catalona and Brawer studies, likelihood ratios for Hybritech PSA results of different categories were calculated as described in appendix B and are provided in table 3-3 (44, 66, 70).[8]

Variations in the use of PSA for screening have been proposed to improve the operating characteristics of this test for prostate cancer (96, 182). These variations, each of which has its own drawbacks, include: 1) *PSA density* (PSAD), a method of correcting the raw PSA value by the volume of the prostate, as measured by TRUS (32, 33, 284); 2) a *predicted PSA* (pPSA) based on gland volume against which measured PSA is compared to make decisions about proceeding to biopsy (206); and 3) *PSA velocity*, the rate of change of PSA over time (63, 64).[9] Research currently underway may lead to a test for more specific types of PSA (36, 37, 106, 211, 212, 213) or other types of biological substances (171, 298) that more precisely identify men with prostate cancer.

One-Time Versus Repeated PSA Screening

Much less is known about the results of repeated screening with PSA than about one-time screening. This gap in our knowledge is significant since a Medicare prostate cancer screening benefit would most likely cover periodic screenings, not one screening per lifetime. The few studies that are available suggest a decrease in the proportion of screenees with cancer over repeated screenings (46, 47), while the proportion of patients with cancer confined to the prostate capsule appears to increase: 71 percent as opposed to 63 percent in one series (13, 70), and 87 percent versus 56 percent in another series (46). Appendix E summarizes these studies.

---

[5] Dalkin and colleagues (101) selected two standard deviations above the mean of the distribution of log-transformed age-specific PSA values to define the upper limit of the reference range.

[6] If the references ranges in table 3-2 are interpreted as age-dependent thresholds for conducting followup tests, they implicitly assume that the costs of a false-positive relative to a false-negative test increase with age. This assumption makes conceptual sense, as older men have a greater risk of treatment complications, and fewer years of life expectancy over which to reap the benefits of screening (on the other hand, younger men also have more years of life to live with any complications engendered by treatment). However, some clinicians are unwilling to trade sensitivity for specificity, regardless of age (255).

[7] In addition, a single, recent case-control study published just prior to publication of this report suggests that PSA may actually preferentially identify aggressive cancers early with relatively high sensitivity and specificity (130).

[8] In a recent study, a group of 72 men underwent systematic sextant biopsies despite a PSA less than 4 ng/mL and a normal digital rectal exam; these men had lung masses on chest radiography and were being evaluated to rule out metastatic prostate cancer as a cause. Prostate cancer was discovered in 3 out of 72 men (4 percent), compared to 9 out of 77 men (12 percent) with a normal digital rectal examination but an elevated PSA (160). This data yields a **likelihood ratio of 0.51** for a normal PSA and 1.51 for an elevated PSA (assuming these cases were consecutive), not inconsistent with the **likelihood ratios** presented in table 3-3.

[9] Because of normal fluctuations in PSA values within a given patient, a PSA velocity based on only two measurements probably has little value in clinical decisionmaking (280). Most recently, the concept of adjusting serum PSA by transition zone volume, rather than whole prostate volume, has been introduced (181).

TABLE 3-3: ESTIMATED LIKELIHOOD RATIOS FOR DIFFERENT RESULTS OF PROSTATE-SPECIFIC ANTIGEN TESTING CHANGING THE ODDS OF SIGNIFICANT (>0.5 mL)[a] PROSTATE CANCER[b]

PSA result[c]	Likelihood ratio	
	Intracapsular cancer	Extracapsular cancer
Pooled Catalona, 1991[d] and Brawer, 1992[e]		
<4.0 ng/mL	0.98	0.09
4.1-10 ng/mL	1.4	5.1
>10 ng/mL	0.4	49.6
Richie, 1993[f]		
<4.0 ng/mL	0.7	0.4
≥4.1 ng/mL	3.0	4.6
Catalona, 1993c[g]		
<4.0 ng/mL	0.8	0.5
4.1-10 ng/mL	2.8	3.2
>10 ng/mL	3.0[h]	23.7

[a] As described in appendix C, probability of a detected cancer <0.5 mL is assumed to be 11% based on J.E. Oesterling, V.J. Suman, H. Zincke, et al., "PSA-Detected (Clinical Stage T1c or BO) Prostate Cancer: Pathologically Significant Tumors," *Urologic Clinics of North America* 17:719-737, 1990.

[b] See appendix C for methods of deriving these estimates.

[c] Results based on Hybritech assay.

[d] W.J. Catalona, D.S. Smith, T.L.. Ratliff, et al., "Measurement of Prostate-Specific Antigen in Serum as a Screening Test for Prostate Cancer," *New England Journal of Medicine* 324:1156-1161, 1991.

[e] M.K. Brawer, M.P Chetner,, J. Beatie, et al., "Screening for Prostatic Carcinoma with Prostate Specific Antigen," *Journal of Urology* 147:841-845, 1992.

[f] J.P Richie, W.J. Catalona, F.R. Ahmann, et al., "Effect of Patient Age on Early Detection of Prostate Cancer with Serum Prostate-Specific Antigen and Digital Rectal Examination," *Urology* 42:365-374, 1993.

[g] W.J. Catalona, D.S. Smith, T.L. Ratliff, et al., "Detection of Organ-Confined Prostate Cancer Is Increased Through Prostate-Specific Antigen-Based Screening," *Journal of the American Medical Association* 270:948-954, 1993.

[h] The discrepancy between this value and the corresponding derivation (0.4) from the pooled earlier studies is explained by the observed difference in probability of pathological localization for cancers (>0.5 mL) detected by PSA >10 ng/mL (32% vs. 5%).

SOURCE: Office of Technology Assessment, 1995. Based on information from M.J. Barry, C.M. Coley, C. Fleming, et. al, "The Safety, Effectiveness, and Cost of Early Detection and Treatment of Prostate Cancer Among Older Men: A Report to the Congressional Office of Technology Assessment," OTA contract paper no. K3-0546.0, Massachusetts General Hospital, Boston, MA June 30, 1994.

PSA Screening Among Men with Symptoms of BPH

As noted earlier, benign prostatic hyperplasia (BPH) can raise PSA levels complicating PSA measurement. Given the widespread prevalence of urinary symptoms indicative of BPH among older men, PSA screening for prostate cancer among this large group may yield little useful information. Men with symptoms of BPH do not appear to be at much greater risk of harboring cancer (ex-cept as conferred by their age) (235) and in one large study, when controlling for age, men with symptoms of prostatism actually had a *lower* chance of being found to have cancer through DRE and PSA screening (72). In addition, because BPH and prostate cancer share symptoms and the likelihood of elevated PSA levels, the specificity of PSA deteriorates to 50 to 79 percent among men with clinical evidence of BPH (173, 309). Furthermore, there appears to be a great degree of overlap

among men with *localized* (intracapsular) prostate cancer and BPH, further limiting the value of PSA testing among men with these symptoms (309).[10]

COMBINATION OF DRE AND PSA

Although combination screening with both DRE and PSA may currently, be the most popular strategy of aggressive office-based early detection of prostate cancer among U.S. urologists, studies of the predictive value of this strategy are only just becoming available for low-risk populations. DRE and PSA each detect some cancers not identified by the other modality; therefore, the yield of a screening program (the percentage of screenees who ultimately have a cancer confirmed) can be increased (to roughly 4 percent) by combining both tests. In addition, the studies of combination testing reported recently have generally performed a set of systematic biopsies if either test is suspicious, as well as biopsies of suspicious lesions noted on followup TRUS; this more aggressive use of TRNB also contributes to the higher yield seen in these studies.

However, these more aggressive strategies result in performing biopsies on up to a third of all screenees; the additional cancers detected must be weighed against the cost and risk of biopsy. Furthermore, these studies were conducted among volunteers, and some data suggest that volunteers may have a higher "prior probability" of prostate cancer than unselected men in the community (261).[11]

The newest studies where DRE and PSA are performed in the same men make it clear that PSA is a better single test than DRE in terms of detecting cancers and of detecting cancers still confined within the prostatic capsule (28, 72, 119, 263, 279).

FOLLOWUP TESTING

Increasingly, followup strategies for a suspicious DRE or PSA include both TRUS and TRNB. Most investigators use TRUS to guide biopsies of areas determined to be suspicious by DRE or TRUS. Many clinicians now perform multiple systematic (four to six) biopsies of the prostate (in a single procedure) in addition to biopsies of suspicious areas, since a patient with a normal TRUS may actually harbor cancer 12 to 33 percent of the time (depending on the PSA level) (157). Others base decisions about whether to perform systematic biopsies on raw PSA values or PSAD values (29, 99, 306). Although some investigators advocate simply following men with mild PSA elevations (i.e., in the 4.1 to 10.0 ng/mL range) if the DRE and TRUS are negative, when aggressively evaluated, this group yields the highest percentage of intracapsular cancers, the real targets of screening.

There is also variability in how clinicians follow men who have a negative set of biopsies after a suspicious PSA test. Some urologists recommend repeating the systematic biopsies at least once (particularly for a PSA greater than 10 ng/mL); others perform followup PSA

---

[10] As mentioned in chapter 2, the FDA has approved the drug finasteride for treatment of BPH. It reduces PSA levels through its intended physiological effects. However, it is not clear, given the need to expect lower PSA levels when screening men on finasteride for prostate cancer, that this drug reduces the (already fairly low) information value of PSA among men with BPH (145, 154, 155, 289). Because of a trend toward less invasive management of BPH, the issue of adequate pretreatment screening of men with a diagnosis of BPH for coexistent prostate cancer is becoming a hotly debated issue (179).

[11] When Oesterling (261) applied the same screening strategy to randomly selected men in the community, only 1 percent were found to have prostate cancer compared with 4 percent in the volunteer studies.

tests more frequently than annually and rebiopsy for either persistent elevations or a rising PSA value. Often then, a suspicious screening test, even if followed by a negative biopsy, will lead to heightened surveillance for prostate cancer and further tests and biopsies in the future. On the other hand, this more intensive surveillance in turn increases the yield of screening to some degree.

Transrectal Ultrasound

Because of the anatomy of the prostate gland itself, TRUS has much better sensitivity for cancers found in certain parts of the prostate than for others (334). Appendix F lists studies that use TRUS as a primary means for early prostate cancer detection. In one of these studies, a demonstration project of the American Cancer Society, about 14 percent of men had a suspicious TRUS, and 15 percent of these men had cancer, a lower predictive value than studies of DRE or PSA alone (Mettlin, 1991). In the absence of a suspicious DRE or elevated PSA, the predictive value in this series dropped to 5.4 percent (19, 215). In a study based in a urologic practice where the prevalence of cancer was especially high (detection rate of 14.6 percent), and where about half of the men were biopsied based on results of combined screening (DRE, PSA, and TRUS), Cooner and associates found that if men had a PSA less than 4 ng/mL and a nonsuspicious rectal exam, the yield of ultrasonographic screening was about 2 percent. Put in another way, the overall yield of the testing strategy only increased from 13.5 to 14.6 percent through the performance of TRUS in addition to DRE and PSA (91).

Several studies provide more direct evidence about the true sensitivity and specificity of TRUS than is available for DRE and PSA. Two studies were able to estimate the operating characteristics of preoperative TRUS performed on men already scheduled for radical prostatectomy for cancer or BPH. The study on men scheduled for prostatectomy for cancer showed a TRUS sensitivity of 52 percent and a specificity of 68 percent (61), and the study of men with BPH showed a sensitivity of 30 percent (315). These relatively low sensitivity estimates for TRUS are a major reason for the increasing tendency to perform systematic biopsies for suspicious DRE or PSA results, even if TRUS does not indicate anything suspicious. Furthermore, these and other studies (337) suggest that TRUS tends to underestimate the size of cancers that are detected, making it a problematic technology for identifying men with small cancers who may not need aggressive treatment. Finally, evidence also suggests that BPH may also erode the ability of TRUS to detect cancer (74).

TRUS itself does not appear to pose any risk for patients, although it does pose costs to patients or their health insurers. In 1992, Medicare reimbursements were $89 for a diagnostic TRUS by itself and $189 for a TRUS-guided biopsy.

Transrectal Needle Biopsy

Modern transrectal needle biopsies (TRNBs) are usually done with ultrasound guidance using a needle mounted in a spring-loaded biopsy "gun." Biopsies can be directed toward areas deemed suspicious by DRE or TRUS, or performed systematically to sample the entire prostate; often six biopsies are taken in a sextant pattern from different parts of the prostate gland (326). TRNB is uncomfortable and can be complicated by infection or bleeding (89). Complications of biopsy include urinary tract infections in 0.5 to 5 percent of patients and urosepsis in an estimated 0.5 percent (no deaths), despite routine antibiotic prophylaxis (16, 91, 109, 160). Some patients also experience bleeding (less than 1 percent) with very few (one out of 835 biopsies in one study) requiring transfusion (91, 109).

TRNB is often considered the "gold standard" test for the diagnosis of prostate cancer; however, it is increasingly clear that the gold standard is "tarnished" to some degree. In terms of the sensitivity of TRNB, investigators from Washington University have found that when men are found to have a persistent mild elevation in PSA (4 to 9.9 ng/mL), repeated biopsies find a large number of cancers presumably missed by previous biopsies. In one preliminary report, 25 percent of these men with one previously negative biopsy had cancer, as well as 14 percent with two previously negative biopsies and 10 percent with three previously negative biopsies (187). Although many of these patients had original biopsies that were directed by abnormal DRE or TRUS results instead of multiple, systematic biopsies, simulation modeling has also suggested systematic biopsies may be relatively insensitive (103).

In terms of specificity, TRNB can detect "incidental" cancers of less than 0.5 mL in volume, which (as discussed in chapter 2) may likely pose no threat to the patient's health, making them conceptually equivalent to "false positives." This risk increases as more biopsies are performed, and particularly with repeated systematic biopsies. Terris and colleagues recently estimated that the probability of finding an incidental cancer on a set of six biopsies was approximately 4 percent (338).

SCREENING THE MEDICARE POPULATION

Age has a complex effect on the results of screening for prostate cancer. The prior probability of cancer increases with age, but the percentage of organ-confined cancers decreases. Furthermore, the specificity of PSA, and probably DRE as well, deteriorates as more men in the population have greater amounts of BPH. Richie and colleagues (279) present the net effect of these factors using data from their large, six-center study of screening:

- The deteriorating specificity of the tests with age resulted in a steeply increasing number of patients with suspicious results on either DRE or PSA that would generate a recommendation for biopsy: 15 percent at ages 50 to 59, 28 percent at ages 60 to 69, and 40 percent at ages 70 to 79.

- The rising prevalence of cancer maintained the predictive value relatively constant, so that cancer was detected in 2, 4, and 7 percent of these age groups, respectively.

- Among men whose cancers were pathologically staged, the percentages that were organ confined (definition not specified) by age groups were 74, 76, and 60 percent.

- In this study, for men ages 60 to 69, adding PSA increased the percentage of men with a suspicious screening evaluation from 16 percent (with DRE alone) to 28 percent; interestingly, the percentage of patients with pathologically localized cancer did not decrease with the addition of PSA in this age group. For men ages 70 to 79, adding PSA to DRE increased the percentage of suspicious evaluations from 20 41 percent, with an increase in the resulting percentage of organ- confined cancers detected from 45 to 60 percent.[12]

All of these data suggest that as screening programs, especially those employing PSA as one screening technology, are directed toward older populations, the number of patients requiring more costly, invasive, and

---

[12] The proportion of men with organ-confined cancers in this study is much higher than in previous studies, presumably because of the performance of systematic biopsies in all patients, rather than only screenees with an abnormal DRE or TRUS. The high proportion of screenees with an abnormal DRE in this study also suggests a very low threshold for considering this exam suspicious.

riskier followup also increases, with a larger number of the cancers ultimately found being confined within the prostate and quite possibly not destined to cause health problems. For policymakers, the decision about whether to support screening depends on the number of followup tests and incidental cancers they are willing to endure in order to find more cancers that may threaten patients' health or lives. This balance may depend on medicine's ability to cure more aggressive prostate cancers, the question addressed in chapter 4.