MARKERS OF CANCER

Cancer can be thought of as a disease of the body, a disease of organs, and a disease of cells (Kahn 1994):

Since the clinical expression of cancer begins with cellular growth, cancer has been defined pathologically as a cellular disease. Clinicians are also aware that cancer is a disease of organs. Cell growth yields masses, which then invade the organ and disrupt its function. Cancer may also be defined as a disease process of the body in that the body’s response to the presence of the tumor affects the presentation and behavior of the patient. The usefulness of the laboratory in cancer management must reflect an understanding of the changes in the genes, cells, organs, and bodily responses to cancer.

A disease of the body

The response of the body to cancer is, in large part, mediated by cytokines released by inflammatory cells reacting to the cancer. Certain of the cytokines generate a persistent acute phase response, with its effects upon plasma proteins and circulating blood cells, and others are responsible for the constitution symptoms so typical of cancer, such as fatigue and weight loss (Tisdale 1997). The body also reacts to cancer by mounting an immune response to latent and neo-antigens expressed by the cancer cells. This contributes to the chronic inflammatory response and can result in the development of autoimmune disorders including a number of neurologic paraneoplastic syndromes (Dropcho 1998). In general, the activation of inflammatory and immune responses in cancer does not produce laboratory findings that are specific for cancer.

Other systemic consequences of cancer result from the release of bioactive substances from tumors. Hematologic abnormalities caused by cancer include cytopenias resulting from the secretion of hematopoietic growth factors by tumors and anemia due to the persistent inflammatory state, poor nutrition, and bleeding. Coagulation disorders are also frequently present in patients with cancer. Hemorrhagic conditions are sometimes due to thrombocytopenia, arising as an autoimmune process, but more frequently are a consequence of chronic activation of the plasminogen system (Carroll and Binder 1999). Thrombotic states are characterized by localized thrombosis and disseminated intravascular coagulation. They are caused by the release of tissue factor and cancer procoagulant from cancer cells (Falanga and Rickles 1999, Gordon and Mielicki 1997) and by the thrombocytosis and increased plasma concentrations of clotting factors (especially, fibrinogen and factor VIII) caused by the persistent inflammatory state. With the exception of cancer procoagulant, the substances responsible for these hematologic abnormalities are not specific for cancer so they have little role in the laboratory diagnosis of cancer. The laboratory evaluation of these processes is, however, a critical aspect of the care of the cancer patient.

Profound systemic effects are also caused by the hypersecretion of hormones by endocrine tumors. Similar effects are caused by peptide hormones secreted by some tumors of nonendocrine organs in what are referred to as the endocrine/metabolic syndromes of cancer (Odell 1997). In many cases of these syndromes, the hormone that is secreted is the normal product of a minor population of hormone-secreting cells within the tissue of origin of the cancer, such as the secretion of erythropoietin by renal carcinomas. In some cases, however, the hormone arises from a cancer in a tissue that is not believed to contain cells that elaborate the hormone, such as the secretion of erythropoietin by cerebellar hemangioblastomas. It is possible that in tumors of this sort, the hormone is expressed as a normal paracrine cytokine within the organ and that it assumes the status of a hormone simply by its overexpression by the tumor (Odell 1997).

A disease of organs

Cancers generate masses of tumor cells that displace and compress neighboring normal tissue leading to injury of the tissue and loss of tissue function. The extent of the effects on the nearby
tissue depend, in part, upon the size of the tumor. While still small, tumors tend to have little or no effect upon the surrounding tissue. As they grow in size, they affect a greater and greater amount of normal tissue. Even more important than size though is location. Cancers that are positioned to collapse or obstruct a tubular structure or large vessel will produce effects long before a tumor growing far from any vital structure. For instance, because it produces obstruction, a squamous carcinoma of the bronchus will cause much more in the way of clinical effects than a comparably sized adenocarcinoma in the periphery of the lung.

Another determinant of the extent of the effects of a tumor upon the surrounding tissue is the pathobiologic stage of the tumor. Before a cancer has begun to spread, all of its local effects are due to the volume of space it occupies. Once it becomes invasive, cancer cells infiltrate normal tissue, disrupting the parenchymal microarchitecture and obstructing the microvasculature. This is very injurious to the normal tissue and inevitably leads to an accelerated loss of tissue function. When a cancer becomes metastatic, it spreads to regional lymph nodes, where it causes local mass effects, and to distant sites, where it causes injury and loss of function in the seeded organs.

The interplay of tumor location and pathobiologic state is well illustrated by prostate cancer. Most prostate cancers arise in the posterior portion of the prostate, away from the urethra, and expand without causing any obstructive symptoms. If they are not detected early by clinical screening procedures, they will become invasive within the pelvis, in which case they can cause bladder dysfunction, ureteral obstruction with hydronephrosis, and local nerve and blood vessel damage usually manifesting as impotence. Prostate cancer can present with symptoms of metastatic disease if the tumor metastasizes early and the local spread has not produced symptoms. Bone pain, lymphadenopathy, and constitutional signs of cancer, most notably weight loss, constitute the usual symptoms of metastatic disease. In contrast, those cancers that develop more anteriorly tend to cause symptomatic urethral obstruction early in their evolution, often before they are widely invasive or metastatic.

Because of interindividual variability in the time course and pattern of tissue injury and function loss caused by tumors, markers of injury and function have a limited role in the diagnosis of cancer. Function markers are extremely useful, though, in providing quantitative indices of the functional status of involved organs.

A disease of cells
Cancer cells differ from normal cells in an immense number of ways, many of which are of central importance in the study of the biology of cancer (Hanahan and Weinberg 2000). Those differences that are clinically measurable can be used as cellular markers of cancer. These markers can be placed into two broad categorizes. The first category consists of markers that are present as components of whole cells. The laboratory measurement of these markers requires a cell sample. The second category consists of marker substances that are present in the body fluids as a result of having been secreted by living cancer cells or released from dead cancer cells. These markers are measured in body fluids, most often in plasma, and do not require access to a cell sample.

Cancer cells. Cancer cells are identified by studies of cellular phenotype and cellular genotype. Of the studies of phenotype, the characterization of morphology by microscopy is the most important. It remains the most specific of all laboratory markers of cancer. For this reason, the microscopic examination of cells or tissue obtained by biopsy of a tumor is considered essential for the diagnosis of cancer. Morphologic characteristics of interest include the cytologic features of individual cells, the microarchitecture of the cell population, and the invasion of normal tissue by the cells. Biopsy techniques that provide tissue specimens usually allow all of these characteristics to be evaluated. The cytologic features of some cancers are highly specific and allow for the diagnosis of cancer based on the examination of cells obtained by scrape or aspiration biopsy.

Genetic abnormalities are frequently demonstrable in cancer cells (Table 11.1). Many of the abnormalities that have been found are present so consistently among tumors of the same type that they probably represent genetic changes involved in the inception and malignant progression of that tumor type (e.g., Ried et al. 1996). Large-scale structural abnormalities of the chromosomes such as deletions and translocations can be studied by chromosome analysis (see Chapter 10). Chromosome analysis by G-banding is performed on tumor cells taken directly from biopsy material or on tumor cells grown in
short-term culture. Direct preparations tend to have only a few metaphase cells for study. Tumor cell cultures usually yield a satisfactory number of metaphase cells but culturing of the cells can select against the more dysfunctional cancer cells in the tumor population and can select for normal cells co-inoculated from the biopsy material (Ketter et al. 1996). This decreases the sensitivity of the method. Karyotyping by fluorescence in situ hybridization suffers from the same problems. The use of fluorescence in situ hybridization for the detection of defined chromosome abnormalities, on the other hand, has many advantages: it can be performed on interphase nuclei, so tumor cell culture is not necessary, and paraffin-embedded biopsy material prepared for histologic study can serve as the source of cells. It has proven to be an informative and accurate method for the directed cytogenetic analysis of solid tumors (Werner et al. 1997) and hematologic malignancies (Kearney 1999).

Chromosome deletions and translocations can also be studied in fresh, frozen, and even paraffin-embedded material using a variety of molecular diagnostic techniques (see Chapter 10). Consider the t(9;22) translocation that results in the Philadelphia chromosome of chronic myelogenous leukemia (Faderl et al. 1999). It consistently arises from translocation of the Abelson oncogene, abl, on chromosome 9 to the breakpoint cluster region, bcr, on chromosome 22. The presence of the \( bcr/abl \) fusion DNA sequence can be demonstrated by Southern blot hybridization or preferential PCR amplification. The translocation can also be demonstrated by detection of either the \( bcr/abl \) fusion protein or the \( bcr/abl \) fusion gene transcription product. The fusion mRNA is detected using Northern blot hybridization or reverse transcription-PCR specific for the \( bcr/abl \) mRNA sequence. Deletions and translocations that show extensive heterogeneity in the site of the chromosome abnormality are not readily studied using molecular diagnostic methods. Gene amplification is another genetic abnormality contributing to the development and progression of cancer. Amplification usually involves a region of DNA larger than one gene and can produce a structural alteration large enough to be detected by chromosome analysis either as double minutes (separate small DNA fragments) or as a homogeneously staining region (DNA segment inserted into a chromosome). Smaller amplifications can be demonstrated using quantitative modifications of fluorescence in situ hybridization, blot hybridization, and PCR amplification (Crotty et al. 1994).

Cancer-associated point mutations of oncogenes and tumor suppressor genes may be restricted to only a few base positions in the gene, such as ras oncogene mutations, but more typically they occur at diverse bases in the gene, as seen with the p53 tumor suppressor gene (Loda 1994). Allele-specific oligonucleotide hybridization is most commonly used to detect point mutations when the mutations in a gene consist of only a few recognized nucleotide substitutions. That technique is not applied to genes that have a large catalog of nucleotide substitutions because of the large number of probes that have to be used. A different technique, single strand conformational polymorphism analysis, is used then (Nollau and Wagener 1997). In this technique, the gene’s DNA is amplified by PCR using labeled primers or labeled nucleotides (so that all of the amplification products are labeled). The products are digested using restriction endonucleases to produce fragments smaller than 200 bp, denatured to yield single-stranded DNA, and electrophoretically separated in nondenaturing polyacrylamide gels. Single-stranded DNA fragments form secondary structures that depend upon the base composition of the fragment; the structure can vary as a result of even a single base alteration. The different single strand structures migrate with different speeds in the gel leading to their separation. Mutated DNA is recognized by the abnormal migration position of its single strand structure. Mutations at any position in the gene can be detected using this method.

### Table 11.1

<table>
<thead>
<tr>
<th>Alteration</th>
<th>Example</th>
<th>Cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chromosome deletion</td>
<td>del 8p12</td>
<td>prostate carcinoma</td>
</tr>
<tr>
<td>Chromosome translocation</td>
<td>t(15;17)</td>
<td>acute promyelocytic leukemia</td>
</tr>
<tr>
<td>Gene amplification</td>
<td>erbB2</td>
<td>breast carcinoma</td>
</tr>
<tr>
<td>Gene point mutation</td>
<td>K-ras</td>
<td>pancreatic carcinoma</td>
</tr>
</tbody>
</table>
A different way in which genetic alterations can be used in the laboratory evaluation of cancer is to detect a genomic change that is not abnormal in itself but which indicates a monoclonal origin of the tumor cells. The foremost example of this is in the evaluation of lymphoid cancer where monoclonality is revealed by demonstrating an identical V(D)J rearrangement in all of the cancer cells. Homogeneity of the immunoglobulin heavy chain gene rearrangement is found in tumors of B lymphocyte lineage, and of the T cell receptor $\alpha$ or $\beta$ genes in tumors of T cell lineage (Scarpa and Achille 1997).

**Marker substances.** The substances that serve as markers for cancer originate in cancer cells and enter the circulation following secretion from living cells or release from dead cells (Figure 11.1). Secreted substances enter the plasma directly and distribute in the extracellular fluids. They are removed by systemic processes. Substances that are released from dead cells enter the extracellular space in and around the tumor and are either catabolized locally or are removed in the lymph. Lymph-borne marker substances eventually enter the plasma, distribute in the extracellular fluids, and are eliminated. Marker substances released from cancerous blood cells are unusual in that they have direct access to the plasma. Marker substance concentrations can be measured in any of the body fluids transited by the marker substance. The most commonly studied fluid is plasma. Other fluids, such as pericardial fluid and peritoneal fluid, are studied when tumor involvement of the respective mesothelial linings is suspected. Urine may be studied if the marker substance or a metabolite of the substance is eliminated by renal clearance; concentrations in the urine are higher than in the plasma thereby allowing for easy analyte measurement.

The concentration of a marker substance in the plasma and extracellular fluids depends upon the rate of entry of the substance into the fluids and its rate of systemic elimination. Except in advanced disease, when many body functions are compromised, the systemic elimination of marker substances is fairly constant. That means that the predominant variable in the plasma concentration of marker substances is the rate of entry into the body fluids. For secreted substances, the rate of entry is determined by the individual cell rate of substance synthesis and secretion and by the number of cells, i.e. the size of the tumor. The individual cell secretion rate usually varies from cell to cell within a tumor due to the greater than normal degree of inter-cell phenotypic variability found in cancer. The rate also varies depending upon where the cancer cells are in their malignant evolution. Some marker substances will be expressed in early in the evolution of the cancer, when the cancer cells are fairly well differentiated, and not later, when the cells show poor differentiation. Other marker substances will be expressed in poorly differentiated cancer cells and not in well differentiated cells. Still other marker substances will be expressed throughout the phenotypic evolution of the cancer.

For released substances, the entry rate is determined by the individual cell content of the marker substance, by the number of cells, and by the turnover rate of the cells. As with secreted marker substances, released marker substances show
variability in the individual cell expression of marker substance within a tumor and variability in the level of expression of the substance at different stages in the malignant evolution of the cancer.

The clinical application of a marker substance depends upon two forms of tissue specificity: a specificity for cancerous tissue as opposed to noncancerous tissue and a specificity for a particular tissue or organ. Tissue specificity is achieved by utilizing substances that arise predominantly in cancers of the tissue of interest. Such substances are usually identified in one of two ways. One approach is to identify substances that show specificity for the normal tissue hoping that they will also be specific for cancer of the tissue. Not infrequently, however, the candidate substances are less tissue specific in cancer. One reason for this is that cancer cells of the tissue may express the substance at a lower level than normal tissue cells. Another reason is that cancer cells from other tissues, especially embryologically related tissues, may express the substance at higher levels than normal. The second approach is to identify substances expressed in cancer of the tissue of interest hoping that some will be specific for the cancer. This has often been done by inoculating animals with human tumor cells to raise antibodies to cancer cell substances. The substances so identified are typically referred to as tumor antigens (Sell 1980). In this approach, a lack of tissue specificity of the candidate substance can be due to expression of the substance in cancers of other tissues or even expression of the substance in other normal tissues.

As regards cancer specificity, in the few ways that the genome of a cancer cell differs from the normal genome, there is the potential for the production of a substance that is truly specific for the cancer producing it. The bcr/abl fusion protein produced by the t(9;22) translocation of chronic myelogenous leukemia is an example of such a substance. Unfortunately, no substances of this sort have yet been found that achieve clinically measurable concentrations in the body fluids.

The marker substances currently in clinical use are products of that portion of the cancer genome shared with the normal cell genome and, therefore, have the potential for being produced by normal cells (Table 11.2). That means that none of them is absolutely specific for cancer. The degree of cancer specificity they do attain depends in large part upon the relative specificity of the attribute of cancer reflected in the laboratory measurement of the substance. These attributes include: monoclonality, altered expression of cellular constituents and products, alteration in the dynamics of substance release, and increased cell turnover. Of these attributes, monoclonality is the most specific for cancer.

As mentioned previously, monoclonality can be demonstrated in B lymphocyte cancer by showing homogeneity of the V(D)J rearrangement of the immunoglobulin heavy chain gene in the cancerous B lymphocytes. Monoclonality can similarly be implied from the demonstration of structural homogeneity of the immunoglobulins secreted by cancerous B lymphocytes and plasma cells (Keren 1999).

Alpha-fetoprotein is an example of a marker substance that is expressed at much higher levels than normal in certain cancers (Abelev and Eraiser 1999). Alpha-fetoprotein is an albumin-like plasma protein expressed at high levels in the yolk sac during early embryonic development and in the liver during late embryonic and early fetal development (Deutsch 1991). In mature liver cells it is normally

<table>
<thead>
<tr>
<th>Class</th>
<th>Marker substance</th>
<th>Cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cellular constituents</td>
<td>carcinoembryonic antigen</td>
<td>colorectal carcinoma</td>
</tr>
<tr>
<td></td>
<td>CA-125</td>
<td>ovarian carcinoma</td>
</tr>
<tr>
<td>Secreted products</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hormones</td>
<td>β-chorionic gonadotropin</td>
<td>germ cell tumors, choriocarcinoma</td>
</tr>
<tr>
<td></td>
<td>calcitonin</td>
<td>medullary carcinoma of the thyroid</td>
</tr>
<tr>
<td>Enzymes</td>
<td>prostate-specific antigen</td>
<td>prostate carcinoma</td>
</tr>
<tr>
<td>Plasma proteins</td>
<td>immunoglobulin</td>
<td>multiple myeloma, B cell leukemia</td>
</tr>
<tr>
<td></td>
<td>alpha-fetoprotein</td>
<td>germ cell tumors, primary liver cancer</td>
</tr>
</tbody>
</table>

Cancer 11-5

Table 11.2
Some Selected Marker Substances of Cancer

<table>
<thead>
<tr>
<th>Class</th>
<th>Marker substance</th>
<th>Cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cellular constituents</td>
<td>carcinoembryonic antigen</td>
<td>colorectal carcinoma</td>
</tr>
<tr>
<td></td>
<td>CA-125</td>
<td>ovarian carcinoma</td>
</tr>
<tr>
<td>Secreted products</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hormones</td>
<td>β-chorionic gonadotropin</td>
<td>germ cell tumors, choriocarcinoma</td>
</tr>
<tr>
<td></td>
<td>calcitonin</td>
<td>medullary carcinoma of the thyroid</td>
</tr>
<tr>
<td>Enzymes</td>
<td>prostate-specific antigen</td>
<td>prostate carcinoma</td>
</tr>
<tr>
<td>Plasma proteins</td>
<td>immunoglobulin</td>
<td>multiple myeloma, B cell leukemia</td>
</tr>
<tr>
<td></td>
<td>alpha-fetoprotein</td>
<td>germ cell tumors, primary liver cancer</td>
</tr>
</tbody>
</table>
expressed at extremely low levels or not at all, so that it is present in only trace concentrations in the plasma. Its expression in cancers that have cells of yolk sac lineage (germ cell tumors) or of hepatocyte lineage (primary liver cancers) results in markedly elevated plasma concentrations. Its specificity for cancer is fairly high but increased plasma concentrations are sometimes seen in conditions associated with extensive hepatocellular injury, such as acute viral hepatitis and liver metastases. Elevated plasma concentrations can also be seen in cirrhosis, presumably as a result of alpha-fetoprotein expression in regenerative hepatocytes.

Carcinoembryonic antigen (CEA) and prostate specific antigen (PSA) are examples of marker substances that are expressed at somewhat higher than normal levels in certain cancers (Hammarström 1999, Stenman et al. 1999, respectively). CEA is an adhesive protein found in the glycocalyx of the microvillar surface of intestinal cells. PSA is a serine protease secreted by prostate epithelial cells into seminal fluid. Both proteins are expressed in mature tissue. They are normally present in low concentration in the plasma, lower than expected based on the level of expression of the proteins. The probable reason for this is that, following epithelial cell death, the majority of the released substance is lost into the respective organ lumen, CEA into the bowel lumen and PSA into the prostatic ducts, rather than entering the local extracellular space where it is available for lymphatic uptake.

Cancer cells express these marker substances at higher than normal levels. This cellular overexpression and the increased turnover typical of cancer cells increase the plasma concentrations of the substances. In addition, derangement of the glandular architecture in cancer contributes to the elevated plasma concentrations. Many of the glands do not have ductular connections. All of the marker substance released by cell turnover in such glands enters the extracellular space, with no loss into a lumen, and ends up contributing to the circulating pool of the substance. The same is true of cancer cells not forming glands. Because these marker substances are expressed in normal tissue, their plasma concentrations are increased in conditions causing injury to the respective organs. For CEA, injury to embryologically related organs that express the substance at low levels, such as the pancreas, can also elevate its plasma concentration. Benign proliferative disorders also lead to increased marker substance concentrations. Such behavior is characteristic of all marker substances that depend upon differential expression for their specificity. It is the major limitation to the specificity of such marker substances.

Hormones are also thought of as tumor markers but they are generally not at all specific for cancer. Cancers of the endocrine organs only infrequently secrete hormones; hypersecretion of hormones instead usually indicates benign disorders, either hyperplasia or an adenoma. Calcitonin secretion by medullary carcinoma of the thyroid is one exception (Giuffrida and Gharib 1998). Calcitonin is secreted by the thyroid C-cells, also called parafollicular cells. Proliferative disorders of C-cells are rare in the general population but are always present in individuals with one of the familial C-cell syndromes: multiple endocrine neoplasia types 2a, multiple endocrine neoplasia types 2b, and familial medullary carcinoma of the thyroid. In these syndromes, C-cell disease is invariably malignant (medullary carcinoma of the thyroid) or premalignant (C-cell hyperplasia). Calcitonin hypersecretion is usual in both malignant and premalignant disease making plasma calcitonin a sensitive and specific marker substance for C-cell cancer. Another hormone that, outside of pregnancy, is specific for cancer is β-chorionic gonadotropin which marks chorionicarcinomatous elements in germ cell tumors (Abelev and Eraiser 1999).

**DIAGNOSIS**

Markers of cancer are used in every aspect of medical care for cancer (Hayes et al. 1996). They are employed in screening for cancer and are very important in monitoring for post-therapeutic recurrence of disease. Markers have a lesser role in establishing the diagnosis of cancer; because of its superior reliability, pathologic examination is almost always considered necessary for the diagnosis of cancer. However, markers, especially cellular markers, can be useful in supporting a microscopic diagnosis of cancer, in elucidating the cellular lineage of the cancer, and in classifying therapeutic subgroups. Prognosis in cancer relates most directly to tumor type and to anatomic markers of invasiveness and metastatic spread but considerable efforts are currently being made, and some success has been achieved, in identifying markers that aid in refining prognostic classification.
Screening

The rationale for clinical screening programs is the detection of treatable disorders prior to their becoming clinically manifest. In consideration of the criterion for being treatable, this translates, at a minimum, into the detection of cancer before it has become metastatic and often it translates into the detection of cancer while it is still limited to the organ of origin. In order to achieve this end, the screening study must have acceptable diagnostic performance during some portion of the early phase of the growth of the cancer—that is, it must have a screening window—and the screening program must be designed so as to assure that the screening study is performed at least once during that screening window. The likelihood of a screening marker achieving the first goal depends upon the natural history of the cancer and the pattern of the marker’s expression within the context of that natural history. The likelihood of a screening program achieving the second goal depends upon the schedule on which the screening study is performed. These points will be illustrated by a consideration of screening for prostate cancer using PSA. PSA will also serve as the illustrative marker in several subsequent sections of this chapter. There is no small degree of controversy surrounding the use of PSA as a screening marker (Woolf and Rothemich 1999, Svetec and Thompson 1998, Moss and Melia 1998). The reader is strongly urged to refer to the cited articles as well as to more recent articles in the literature to appreciate the many viewpoints in this controversy.

The natural history of prostate cancer, according to the model proposed by Stenman et al. (1999), is depicted in Figure 11.2. There is an initial phase of intraepithelial neoplasia that can arise as early as the third or fourth decade of life (Franks 1954) but which may appear later in life. During this phase, which is of uncertain duration (in the figure, a duration of 15 years is shown), the cancer has a fairly slow rate of growth. Once the tumor becomes locally invasive it experiences a period of relatively rapid growth. When the cancer reaches a size at which vascularization is needed, its growth rate lessens with doubling times estimated to be on the order of two to three years (Schmid et al. 1993). This rate of growth persists until the cancer metastasizes.

Before prostate cancer becomes metastatic, the rate of entry of PSA into the body fluids appears to be directly related to the size of the cancer. A direct relationship between size and PSA entry rate is also found for benign prostatic hyperplasia (BPH). The magnitude of the entry rates differs considerably between the two diseases, however. Stamey et al. (1987) estimate that, on average, normal prostate contributes to plasma PSA concentration at a rate of 0.2 µg/L/g, BPH at a rate of 0.6 ug/L/g, and prostate cancer at a rate of 2 µg/L/g. Statistical modeling of the large data set in the report by Collins et al. (1993) indicates that BPH increases plasma PSA concentrations with a median value of 0.13 µg/L/g and that normal prostate makes only an extremely small contribution (Noe, unpublished; based on a model of direct proportionality between tissue volume and plasma PSA concentration and a lognormal distribution of PSA values). Modeling of the data in Partin et al. (1990) yields median values of 0.14 µg/L/g and 2.64 µg/L/g, respectively, for BPH and prostate cancer (Noe, unpublished).

In men with prostate cancer but otherwise normal prostates, prostate cancer would be detectable as soon as the plasma PSA concentration reached the limit of detection of the assay used to measure PSA concentration. A desirable value for this limit is 0.2 µg/L (Stenman et al. 1995). Based on a median value of 2.64 µg/L PSA concentration rise per 1 g of cancer tissue, half of the men would have detectable PSA concentrations while their cancers weighed 0.075 g or less. If the standard deviation in the lognormal distribution of the relationship between PSA concentration and tissue mass in prostate cancer is similar to that in BPH, ninety-five percent of the cancers would be detectable while their mass was less than 0.22 g. This is well shy of the mass associated with capsular
invasion (3 to 10 g, Stenman et al. 1999), indicating that most cancers could be detected while they were organ limited. But would they be detected? That depends upon the screening program schedule. If, for example, the schedule called for a single PSA determination at the age of 50 years, prostate cancers would be detected only in those men who had a cancer that had already reached a detectable size. They would represent only a fraction of the individuals who would eventually have symptomatic prostate cancer. If, on the other hand, the PSA concentration were measured yearly after the age of 50 years, almost every cancer would be detected before capsular invasion. This is so because it takes about 13 years for a prostate cancer to grow from 0.22 g to 3 grams (see Figure 11.2; 0.22 g and 3 g equate to diameters of 7.5 and 18 mm, respectively). Thus, many screening studies would be performed on every man harboring a prostate cancer while the cancer was detectable and pre-invasive. An optimal screening schedule is one in which the screening study is performed once during the screening window. If a sensitivity of 0.95 were to be considered the minimal acceptable screening performance, the screening window for PSA would be 13 years, from a tumor mass of 0.22 g to a mass of 3 g. Performing the screening study on a schedule of once every 13 years would therefore be optimal. To account for possible lapses in the regularity of obtaining a screening study, a PSA determination every 10 years might be recommended. A problem with this analysis is, of course, that it applies to only a few men, those with normal prostate glands. By the time they reach their fifties, an age at which starting to screen for prostate cancer is reasonable based on the natural history of the disease, most men have developed some degree of BPH. This means that there is an appreciable background plasma PSA concentration. Figure 11.3 shows population frequency distributions for plasma PSA concentration. The PSA concentrations in men without cancer are due entirely to the presence of BPH; the PSA concentrations in men with cancer are due to cancer and concomitant BPH (Partin et al. 1990). There is a considerable overlap in the frequency distributions of the two populations. Consequently, screening for prostate cancer in the male population as a whole necessitates a tradeoff between screening sensitivity and specificity. For unscheduled, or one-time, screening among men 50 to 59 years old, the tradeoff is as shown in Figure 11.4 (for a similar analysis of a different data set read the article by Meigs et al. 1996). To determine the critical screening value for the plasma PSA concentration in this setting, the threshold likelihood ratio for followup must be calculated. This is done using the formula,

\[
\text{threshold likelihood ratio for followup} = \frac{(1 - \text{prevalence}) \cdot P[\text{rejection}]}{\text{prevalence} \cdot (1 - P[\text{rejection}])}
\]

where P[rejection] is the threshold posterior probability for rejection of the diagnosis of prostate cancer.

Figure 11.3 Reference frequency distributions for plasma PSA concentration. The curves represent lognormal frequency distributions fit to the data reported by Catalona et al. (1994) for men 50 to 59 years old.

Figure 11.4 ROC curves for screening for prostate cancer using plasma PSA concentration. The squares represent the data reported by Catalona et al. (1994) for men 50 to 59 years old. The continuous line is the curve constructed from lognormal frequency distributions of the data.
cancer. Using the value derived for the threshold likelihood ratio for followup, the critical screening value is determined from the likelihood ratio curve (shown in Figure 11.5). To calculate the threshold likelihood ratio for followup, the prevalence of prostate disease in the clinical population must be estimated and the threshold posterior probability for rejection of the diagnosis of prostate cancer must be selected. In their study, Catalona et al. (1994b) found that, in men aged 50 to 59 years, the prevalence of prostate cancer is approximately 0.04. This value serves as the prevalence estimate. Any of a number of values could be selected for the threshold posterior probability for rejection of the diagnosis. For instance, it might be felt to be reasonable to reject the diagnosis if the posterior probability of cancer is less than 0.06 (one and one-half times the prevalence). Substituting this threshold probability into the formula yields a threshold likelihood ratio of 1.5. The corresponding critical value for the plasma PSA concentration is 4.0 µg/L. Using the modeled reference frequency distributions (Figure 11.3) and a critical value 4.0 µg/L, the sensitivity of the screening study is predicted to be 0.67. Catalona et al. report a sensitivity of 0.75 in their study sample at this critical value (Catalona et al. 1994b). They also report that approximately 70 percent of the cancers detected using this critical value were confined to the prostate and thus highly treatable. The specificity at this critical value is predicted to be 0.74 and was found by Catalona et al. (1994b) to be 0.63.

Typically, a screening program gives better performance than unscheduled screening: better specificity can be had at a comparable rate of detection of organ-limited cancer or an improved rate of detecting organ-limited cancer can be obtained at a comparable specificity. Consider the case in which the specificity of the PSA screening program in men aged 50 to 59 years is targeted to be 0.75. This specificity is the same as that predicted in the foregoing discussion of unscheduled PSA screening. For a single performance of the study, the stipulated specificity is obtained using a critical PSA concentration of 4.0 µg/L. In a screening program, the study is performed multiple times so, in general, using the same critical value would result in a lower overall specificity. However, overall specificity remains nearly the same as single-study specificity when the repeated measurements are highly correlated, as they would be in a PSA screening program because the PSA concentration in each individual reflects the mass of BPH in that individual and the mass changes only slowly over time. Therefore, a critical value of 4.0 µg/L can be used for the screening program. The schedule of screening studies is calculated based on the target detection rate. Consider, for example, a fairly ambitious target detection rate, 0.9. At the stipulated critical value, this detection rate is found to correspond to a cancer mass of 1.74 g. (This calculation is based on an assumption made earlier, that the standard deviation in the lognormal distribution of the relationship between PSA concentration and tissue mass in cancer is similar to that in BPH, and takes into account the concomitant presence of BPH). This mass is about half that at which prostate cancer first shows capsular invasion, which is 3 g (Stenman et al. 1999), so the screening window is approximately one doubling time which, for tumors of this size, is two to three years (Schmid et al. 1993). Therefore, a cautious screening schedule would consist of a PSA determination every two years.

**Improving screening performance.** The performance of plasma PSA concentration as a screening study for prostate cancer is not outstanding. This can be appreciated by comparing the ROC curve for PSA concentration (Figure 11.4) with the ROC curve for a truly excellent diagnostic marker such as ferritin concentration, which is used to detect iron deficiency in adults with anemia (Figure 3.14). The area under the ROC curve for PSA concentration is 0.78 while that for ferritin

![Figure 11.5 The likelihood ratio of prostate cancer in men 50 to 59 years old as a function of their plasma PSA concentration constructed from lognormal frequency distributions (Figure 11.3). Also shown is a reference line for a likelihood ratio of 1.](image-url)
concentration is 0.95. Clearly, it would be better if the screening performance for PSA concentration were more like the performance of ferritin concentration.

One way to improve the discriminatory power of a screening marker is to reduce the preanalytic and analytic components of the variability in the measurement of the marker. In the case of PSA concentration, preanalytic variability could be reduced, for example, by stipulating that blood specimens be obtained only after a 2 day abstinence from sexual activity, as ejaculation is associated with a 24-hour elevation in plasma PSA concentration (Simak et al. 1993). Because digital rectal examination can sometimes raise the plasma PSA concentration (Bruel et al. 1992), it is already recommended that blood specimens be drawn before the physical examination. There are many ways to reduce the analytic variability of a laboratory study, but the most important way is to use a more specific method of measurement. Immunoassays are used to measure PSA concentration. Some of the available assays utilize polyclonal antibodies which can be expected to recognize various epitopes on the PSA molecule. Standardization in terms of utilizing monoclonal antibodies to a single epitope could reduce analytic variability and improve interlaboratory transportability of study results (Stenman et al. 1995).

Another way to improve screening performance is to take account of known sources of biologic variability in the measurement of the screening marker. As discussed in Chapter 6, such sources of variability typically include gender, age, and race. Age has been shown to contribute to the variability of PSA concentrations. Figure 11.6 shows the frequency distribution for PSA concentration as a function of age, as modeled by Oesterling et al. (1993). The frequency distribution shifts rightward, i.e., toward higher values of PSA concentration, with increasing age. Much of this age effect is due to an increase in prostate size in men over the age of 55 years; the increase being attributable to progressive BPH (Collins et al. 1993). However, there is also an independent contribution of age alone (Oesterling et al. 1993, Collins et al. 1993).

The frequency distribution of PSA concentration in patients with prostate cancer also shifts rightward with increasing age. The magnitudes of the rightward shifts of the two distributions are nearly equal. Consequently, the separation between the two distributions remains constant and the ROC curve for screening for prostate cancer using PSA concentration retains roughly the same shape and area under the curve (Catalona et al. 1994b). Note, however, that the PSA values associated with given points on the ROC curve change with increasing age. For instance, Catalona et al. (1994b) found that the sensitivity and specificity achieved with a PSA concentration of 4.0 µg/L in men aged 50 to 59 are approximately 0.75 and 0.65, respectively. In men aged 60 to 69, this screening performance is found with a PSA concentration of 4.5 µg/L, and in men 70 years and older with a PSA concentration of 5.0 µg/L.

The age-specific frequency distributions shown in Figure 11.6 were derived from a population of white males living in Minnesota. Race is another source of biologic variability in PSA concentration, so the distributions cannot be applied to men of other races. In particular, on a decade-by-decade basis, median PSA concentrations in black men are comparable to those in white men, but the upper limits of the normal range are much higher in black men (Morgan et al. 1996).

A more explicit way in which to account for the contribution of hyperplastic prostate tissue to the plasma PSA concentration is to adjust for the size of the prostate. Prostate volume can be measured clinically by transrectal ultrasound. The method is relatively easy to perform but unfortunately suffers from appreciable measurement variability (Catalona et al. 1994a). Figure 11.7 shows the reference range.
Figure 11.7 Plasma PSA concentration as a function of prostate volume. The dots represent the observed data. The lines are contours of the frequency distribution as calculated based on a lognormal statistical model of PSA values. Reprinted Oesterling JE, Jacobsen SJ, Chute CG, Guess HA, Girman CJ, Panser LA, and Lieber MM. 1993. Serum prostate-specific antigen in a community-based population of healthy men. JAMA 270:860.

for PSA concentration as a function of prostate size as modeled by Oesterling et al. (1993). The lines in the figure can be used like linear discriminant functions for screening classification. For instance, to achieve a screening specificity of 0.90, PSA concentration and prostatic volume measurement pairs above the “90 percentile” line would be considered screen-positive and those below would be considered screen-negative. Alternatively, the PSA concentration and prostatic volume measurements can be used to derive a discriminant score that is used to classify result pairs. The most appropriate algebraic form for the discriminant score is a linear equation. The algebraic form that is actually used clinically, however, is the ratio of PSA concentration to prostate volume, referred to as the PSA density. Evaluation of this ratio in a modest number of men indicates that a ratio of 0.05 has a sensitivity of 0.95 for cancer confined to the prostate and a specificity of 0.70; a ratio of 0.10 has a sensitivity of 0.80 and a specificity of 0.95 (Benson et al. 1992). This is far better screening performance than that achieved by the PSA concentration. The specificity estimates are consistent with the data reported by Oesterling et al. (1993) in their large study of PSA concentration and prostate volume (Figure 11.7; the 75th percentile line approximates a PSA density of 0.06 and the 95th percentile line approximates a PSA density of 0.11). The sensitivity estimates are similar to those found for PSA density in a large study reported by Catalona et al. (1994a): 0.71 and 0.90 for PSA densities of 0.05 and 0.10, respectively. The specificity estimates reported by Catalona et al. (1994a) are lower than those reported by Benson et al. (1992). However, the study by Catalona et al. (1994a) was confined to men with PSA concentrations above 4.0 µg/L or with abnormal digital rectal examinations, resulting in a strong bias toward higher values of PSA density among men with BPH, and thereby leading to significant underestimation of the specificity. (Unfortunately, this design limitation is not widely appreciated and the low specificity estimates found in the study are cited as evidence that the PSA density is not a better screening study than PSA concentration).

It is sometimes possible to increase the clinical information provided by a laboratory measurement by interpreting the result in the context of a physiologic model. An example of this would be the interpretation of arterial blood gas measurements using a quantitative model of acid-base metabolism. Model-based interpretation can also be used to improve the discriminatory power of a screening study. In cancer screening, a model of tumor growth is a logical choice. As a tumor grows, the amount of marker substance released by it usually increases; although this will not be so if the growth results from, or is accompanied by, malignant evolution of the cancer with reduction in the expression of the marker. As the rate of entry of the marker into the circulation increases, the plasma concentration of the marker increases. If the marker is followed over time, the steady increase in its concentration can be demonstrated and the presence of the growing tumor inferred. An advantage of this approach is that an upward trend in marker concentration may be detectable at concentrations of the marker that are below what would otherwise be the critical value for the marker. Hence, a cancer can be detected when it is smaller. A disadvantage of the approach is that an individual must be tested a number of times over an interval of time long enough to allow for a trend in marker concentration to be detected.

In screening for prostate cancer, the trend in PSA concentrations over time is called the PSA velocity. It is usually evaluated as the average of two consecutive measurements of the rate of change of the PSA concentration. Studies of the variability of PSA concentration in healthy men and in men with BPH indicate that an interval of at least a year between concentration determinations is desirable when computing the rate of change in the concentration (Carter et al. 1995, Kadmon et al. 1996).
means that three PSA determinations equally spaced over two years is the minimum testing schedule.

A nonzero PSA velocity is expected in men with prostate cancer but also in men with BPH, as BPH is also a progressive process. Because the rate of release of PSA from hyperplastic prostate tissue is usually small and that from cancerous prostate tissue is generally large, the PSA velocity would be expected to be lower in men with BPH than in men with tumors. However, in men with PSA concentrations less than 4 µg/L, the PSA velocity in men with prostate cancer is similar to that in men with BPH (Carter et al. 1992). A dramatic increase in PSA velocity is found in most men with prostate cancer only after their PSA concentration exceeds 4 µg/L (Carter et al. 1992). It appears, therefore, that PSA velocity does not improve the lead time for detection of prostate cancer when compared to simple screening based on a critical PSA concentration of 4 µg/L.

**Alternative markers.** The approaches to improving screening performance considered in the preceding paragraphs do not address the issue that is often at the heart of the failure of a plasma substance to serve as a satisfactory screening marker which is that the would-be marker does not have a high degree of cancer specificity. Markers that are not specific for cancer cannot perform well in a setting in which there is a high prevalence of benign disease. In such circumstances it is necessary to identify a screening marker that is more cancer specific. Usually this entails finding another marker substance altogether but, in some cases, a certain form of the original marker can be identified that is more highly cancer specific. The cancer specificity of the form may arise from its preferential synthesis in cancer cells or from some cancer-specific post-translational modification of the marker. No plasma substance has yet been found to be more specific for prostate cancer than PSA.

There is, however, a form of PSA that appears to be somewhat more cancer specific, protein-bound PSA (Polascik et al. 1999). PSA exists in the circulation in four forms: (1) proPSA (Mikolajczyk et al. 1997), which is the enzymatically inactive precursor form of PSA, (2) active PSA, (3) nicked PSA, which is the product of the action of nicking enzyme on either proPSA or active PSA, and (4) other degradation products of PSA (Hilz et al. 1997). In tissue (Jung et al. 2000) and at the time of release into the circulation, all of these forms of PSA exist as free species. In the circulation, active PSA is rapidly inactivated by binding to the plasma protease inhibitors, α₂-macroglobulin and α₁-antichymotrypsin. Essentially all of the circulating active PSA is inhibitor-bound. Both forms are cleared from the circulation by the liver. The α₂-macroglobulin-bound form is removed very rapidly (half-life of minutes) while the α₁-antichymotrypsin-bound form is removed slowly (half-life of days). Consequently, even though active PSA is more avidly bound by α₂-macroglobulin than by α₁-antichymotrypsin, the α₁-antichymotrypsin-bound form is the predominant plasma form. Nicked PSA binds very slowly to α₂-macroglobulin and not at all to α₁-antichymotrypsin. As a result, it is present in plasma predominantly in the free form. The other degradation products, as well as proPSA, are present entirely in the free state.

It has been found that the percentage of PSA present in the free state in plasma is lower in prostate cancer than in BPH. This finding means that, in prostate cancer, a greater fraction of the circulating PSA is present in the α₁-antichymotrypsin-bound form. One way that this could come about is if the PSA entering the circulation from cancerous tissue were enriched in the active form. Then, of that not rapidly eliminated by binding to α₂-macroglobulin, most would bind to α₁-antichymotrypsin. Alternatively, the release of human kallikrein 2, the prostatic enzyme that activates proPSA (Rittenhouse et al. 1998), could be greater from cancerous tissue than from BPH, resulting in a greater fraction of the PSA being activated and ending up bound to α₁-antichymotrypsin. It is currently not known if these or other mechanisms explain the clinical findings (Stenman et al. 1999).

**Screening for a genetic predisposition to cancer.** The development and evolution of cancer appear to result from an accumulation of genetic lesions within the cell of origin (Kinzler and Vogelstein 1996). These lesions include gain-of-function mutations of oncogenes, loss-of-function mutations of tumor-suppressor genes, and loss-of-function mutations of DNA repair genes (Lynch et al. 1997). When a lesion with oncogenic potential is present as a germline mutation, the total number of somatic genetic lesions that need to occur for the development of cancer is fewer. Consequently, there is an increased probability that the individual bearing the mutation will develop cancer. This predisposition may be predominantly to cancer of a particular organ or it may be to cancers of several different organs.
Laboratory screening for germline mutations that predispose to cancer is an appealing idea but screening for a mutation is only of practical value when four conditions are met: (1) the mutation must confer an appreciable risk of cancer, (2) screened individuals must have a significant probability of harboring the mutation, (3) the screening laboratory study must have acceptable diagnostic performance, and (4) there must be something that can done to substantially reduce the risk of developing invasive cancer in individuals who have the mutation (Ponder 1997).

Unless the risk of cancer is increased to an unacceptable level by the presence of an oncogenic mutation, individuals will not be willing to undergo a test for the mutation, especially if the test is expensive. If the baseline risk for a cancer is high, such as with the most common cancers, a modest mutation-related increase in the risk of the cancer may yield an level of risk that is unacceptable to most individuals. Thus, for these cancers, even mutations with relatively minor oncogenic effects may be desirable targets for screening. If the baseline risk for a cancer is low, a mutation needs to have a major oncogenic effect to be considered for screening because only a dramatic mutation-related increase in the risk of the cancer will result in a unacceptable level of risk for the cancer.

The prevalence of oncogenic germline mutations is very low. Screening random individuals for one of these mutations would therefore be inefficient and, if the screening test did not have perfect specificity, there would be a large number of false-positive results. To minimize these problems, screening studies are performed only on individuals who have an increased likelihood of having the mutation as indicated by a family history of the type or types of cancer associated with the mutation.

Screening for oncogenic mutations relies primarily on molecular diagnostic studies. These studies normally have excellent specificity. They are, however, usually limited to the detection of well-characterized mutations. Consequently, the diagnostic sensitivity of a study will be very high among members of families with well-characterized mutations but will be it will be very low among individuals who come from families with novel mutations. The overall sensitivity of a study will be determined by the proportion of families who have a predisposition to cancer due to well-characterized mutations. For those cancers in which the sensitivity of the laboratory screening test is unsatisfactory, the practical solution is to use the family history as the screening tool.

Screening for a predisposition to cancer is clearly justified when there is an intervention that can be undertaken to lessen or eliminate the risk of developing the cancer. This may mean something as simple as a change in lifestyle designed to lessen the exposure to a carcinogen implicated in the development of the cancer but, more often, it means prophylactic removal of the organ or organs likely to develop the cancer. In the absence of intervention options, screening can be useful for identifying those individuals who should undergo intensified cancer surveillance. This usually consists of early institution of cancer screening and shorter screening intervals to account for the tendency for an early age of appearance in mutation-related cancers and for their rapid evolution into invasive and metastatic disease.

Colorectal cancer is an example of a cancer that has hereditary forms. Two of these are familial adenomatous polyposis and hereditary nonpolyposis colorectal cancer (HNPCC). HNPCC is responsible for approximately 5 percent of the cases of colorectal cancer. It can be caused by loss-of-function mutations in a number of DNA mismatch repair genes; about half of the cases are caused by mutations in hMLH1 and hMSH2 (Boland 2000, Lynch and de la Chapelle 1999). Mutations in either hMLH1 or hMSH2 result in an 80 percent probability of developing colorectal cancer and, in females, a 40 to 60 percent probability of developing endometrial cancer (Vasen et al. 1996). Carriers of mutations in hMSH2 also have an increased risk of cancer of a number of other organs. There are two clinical approaches available to reduce the burden of this very high probability of developing cancer. In the first approach, the risk of developing invasive cancer is lowered by beginning a program of annual colonoscopic surveillance when an individual at risk for HNPCC is in his or her twenties (Lynch and Lynch 2000). This program addresses the earlier age of onset of the hereditary form of the cancer. The program also accounts for the preference for proximal colon (colonoscopy) and the shorter adenoma-to-carcinoma transition time (annual procedures). In the second approach, prophylactic subtotal colectomy is performed, reducing the risk of developing cancer by removing most of the tissue at risk (Lynch and Lynch 2000).

Individuals at increased risk for a predisposition to colorectal cancer are identified by a family history
of colorectal cancer among first-degree relatives. Because of the high rate of cancer development among individuals with HNPCC, kindreds affected by this disease often have multiple family members, across several generations, who have had nonpolypsis colorectal cancer or endometrial cancer. Additionally, a history of onset of the disease before 50 years of age is common. These findings have been formalized as the Amsterdam criteria for the clinical diagnosis of HNPCC. When the Amsterdam criteria are met, family members can assume they come from an affected kindred. The question for them is if they are carriers of the causative mutation. This question can be answered currently only in those families with mutations in genes known to cause an appreciable fraction of the cases of HNPCC. Half of the cases are caused by mutations in hMLH1 or hMSH2 and laboratory studies are available for mutation analysis of these genes. An individual who is found to carry a mutation is advised to enter a program of intensive cancer surveillance. If it is felt to be unlikely that he or she will be able to comply with the demands of such a program, the option of a subtotal colectomy may be offered. Family members who are found not to carry a mutation can simply be advised to enroll in a routine program of screening for colorectal cancer once they reach the age of 50 years. Members of families in which HNPCC is caused by other mutations cannot have their individual cancer risk determined; they are left knowing only that they have a 50:50 chance of carrying a causative mutation. Intensive cancer surveillance is recommended for all family members in this circumstance.

When the Amsterdam criteria are not met, there is a possibility that a familial clustering of cancers results from a different form of hereditary colorectal cancer—one with a lower rate of development of cancer—or else the clustering could be a chance aggregation of sporadic tumors. Mutation analysis of hMLH1 or hMSH2 would be informative in these individuals in whom the result was positive, but a negative result would not exclude the possibility of HNPCC, as half of the kindreds with this disorder have a different genetic basis, nor would it exclude the possibility of other familial cancer predispositions. Clinical management in this situation is guided by the epidemiologic finding that, in individuals who have two or more first-degree family members with colorectal cancer, the risk of developing colorectal cancer is multiplied about 2.75 times (Fuchs et al. 1994). Because colorectal cancer is common, there being a lifetime risk of 0.06 in the general population, even a modest increase in risk is significant and it is usually felt prudent to offer intensive cancer surveillance as a clinical option.

Individuals with colorectal cancer who have a family history of colorectal cancer in a first-degree relative, who have previously had colorectal cancer or endometrial cancer, who are younger than 55 years, or who have a tumor with characteristic pathologic findings, are candidates for examination of their tumor tissue for microsatellite instability and mutation analysis (Jass 2000). Mutations of the DNA mismatch repair enzymes that cause HNPCC result in insertion and deletion mutations at microsatellite sequences, referred to as microsatellite instability (Boland 2000). Hypermethylation of the promoter region of hMLH1, which is not associated with a familial predisposition to cancer, also causes microsatellite instability but occurs only in elderly patients (Jass 2000). In cases in which tumors show microsatellite instability, histochemical staining of the tumor tissue or laboratory mutation analysis of white cells from the patient can be used to identify individuals with germline mutations in hMLH1 or hMSH2. Demonstration of such a mutation establishes a diagnosis of HNPCC. Given this finding, first-degree family members should undergo evaluation of their carrier status. 

**Diagnosis**

Because of the serious medical and psychological implications of being diagnosed with cancer, the diagnosis must be established with certainty. With very few exceptions, this means that microscopic examination of the cancerous tissue is required. The microscopic examination may take the form of surgical pathologic review of a biopsy specimen, cytopathologic examination of cells exfoliated from the lesion, or hematologic evaluation of blood or bone marrow. Despite the generally high specificity of microscopic examination, it is not usual for the diagnosis of cancer to be based solely on the microscopic findings. To achieve even greater diagnostic specificity, clinical, imaging, and laboratory findings are also considered. Typically, it is only if there is consistency in all of the findings that the diagnosis of cancer is made.

Whenever possible, diagnosis includes classification of the cancer in terms of its cell type. Classification is usually quite easy given the location and
microscopic appearance of a tumor but, occasionally, it is necessary to perform special studies in order to make an assignment as to the cell type (Table 11.3). Electron microscopy can be used to extend the resolution of the microscopic examination and special stains can be employed to identify certain cellular contents. Cell-type specific enzymes can be detected by enzyme histochemistry and enzymes and other proteins can be demonstrated by immunocytochemistry. Characteristic chromosomal abnormalities can be detected by microscopic techniques such as fluorescence in situ hybridization (FISH) or by molecular biologic analysis, such as Southern blot hybridization or polymerase chain reaction (PCR). Chromosomal abnormalities have proven to be especially useful markers in the classification of hematologic malignancies (Frizzera et al. 1999). Among solid tumors, the t(11;22) translocation of Ewings sarcoma has been found to be a reliable classification marker.

<table>
<thead>
<tr>
<th>Cellular constituent</th>
<th>Technique</th>
<th>Example</th>
<th>Cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Membrane protein</td>
<td>Immunocytochemistry</td>
<td>Leukocyte common antigen</td>
<td>Lymphoma, lymphocytic leukaemia</td>
</tr>
<tr>
<td>Secretory granules,</td>
<td>Electron microscopy</td>
<td>Premelanosome</td>
<td>Melanoma</td>
</tr>
<tr>
<td>granule contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cytoskeletal protein</td>
<td>Electron microscopy</td>
<td>Desmosomes</td>
<td>Squamous carcinoma</td>
</tr>
<tr>
<td>Cytosolic enzyme</td>
<td>Enzyme histochemistry</td>
<td>Chloroacetate esterase</td>
<td>Acute myelogenous leukemia</td>
</tr>
<tr>
<td>Cytosolic protein</td>
<td>Immunocytochemistry</td>
<td>Myoglobin</td>
<td>Rhabdomyosarcoma</td>
</tr>
<tr>
<td>Chromosomes</td>
<td>FISH, PCR, Southern blot hybridization</td>
<td>t(11;22)</td>
<td>Ewings sarcoma</td>
</tr>
</tbody>
</table>

**MANAGEMENT**

The management of a patient with cancer includes establishing a prognosis, selecting appropriate therapy, and clinical monitoring. Markers of cancer are used in all of these aspects of care.

**Prognosis and prediction**

In cancer medicine, prognosis can be thought of as an expression of the probability that the cancer will progress. This probability is usually expressed as the likelihood of progression within a specific period of time, say a certain number of years. Obviously, for many tumors, the prognosis depends upon whether therapy is undertaken; it is possible for the prognosis of the untreated cancer to be dismal while the prognosis with treatment may be good. To account for the effects of therapeutic intervention on prognosis, the concept of prediction has been developed. Prediction is the probability that the cancer will respond to a specified therapy.

Prognosis depends largely upon the type of cancer and upon the extent and location of spread of the cancer at the time of presentation. These and other features of a cancer that relate to the prognosis in individual patients are referred to as prognostic factors. Prognostic factors that relate to the type of tumor include histopathologic attributes that correlate with tumor proliferation rate, aggressiveness, or metastatic potential; genetic alterations that relate to the stage of malignant evolution of the cancer; and cell and marker substance findings that relate to phenotypic features of the cancer associated with disease progression. The extent of spread of a cancer is assessed from clinical examination, tumor biopsy or surgical removal, and imaging studies. Prognosis usually worsens in the order: in situ cancer, tumor confined to the organ of origin, local spread of tumor beyond the organ of origin, involvement of regional lymph nodes, and distant metastatic spread. The prognostic importance of each of these categories varies somewhat among cancer types as does the prognostic implication of the exact extent of intra-organ and local spread. As an aid to the clinician, staging systems have been devised that list the varying criteria of tumor spread and order their importance. The most popular of these are the TNM (Tumor Node Metastasis) staging systems. A crucial use of staging designations beyond that of establishing a prognosis is to aid in identifying patients in whom the extent of tumor spread is such that...
local-regional therapy (surgery and radiotherapy) may be effective. In prostate cancer, for instance, radical prostatectomy is recommended only for patients with stage T1 (clinically inapparent tumor neither palpable nor visible by imaging), T2 (tumor confined within prostate), or T3a (unilateral extracapsular extension of tumor through the prostate capsule) cancer that is N0 (no regional node metastasis) or NX (regional lymph nodes cannot be assessed) and M0 (no distant metastases) or MX (presence of metastasis cannot be assessed).

Predictive factors are clinical and laboratory features that allow accurate prediction of the response to a therapy; they include cell and marker substance findings that relate to the molecular target of therapy or to features of the cancer associated with resistance to therapy. Because prediction applies to a specific therapy, the predictive factors for a cancer may vary among the therapies available for treatment of the cancer. For example, overexpression of erbB2 in breast cancer is not a useful predictive factor for response to endocrine therapy but it is a predictive factor for response to trastuzumab (Yamauchi et al. 2001). The predictive factors in use for a particular therapy can also be expected to change over time as our understanding of the molecular and cell biology of cancer allows for the development of new and more accurate markers. It is worthwhile noting that although some prognostic factors have a role only in prognosis and some predictive factors only in prediction, often the same marker may serve as a prognostic and a predictive factor (Hayes et al. 1998).

Prognostic and predictive classification can be either dichotomous or quantitative. Dichotomous classification is based upon the use of a critical value, or combination of values (for combination testing), that identify the prognostic or predictive category in which an individual likely belongs. If the measured value of the prognostic or predictive factor exceeds the critical value, he or she is assigned to one category and if the measured value is less than the critical value, he or she is assigned to another category. The probability of belonging to the assigned category can be calculated using Bayes formula (as discussed in Chapter 5); when there are two classification categories,

$$P[\text{post}] = \frac{\text{prevalence} \cdot FCC_1}{\text{prevalence} \cdot FCC_1 + (1 - \text{prevalence})(1 - FCC_2)}$$

where \text{prevalence} is the prevalence of the assigned category and the performance characteristics \(FCC_1\) and \(FCC_2\) are the fractions of patients correcting classified in each of the two categories. Dichotomous classification provides less information about the individual than quantitative classification, in which the probability of an individual belonging to a particular prognostic or predictive group is calculated using the likelihood ratio associated with the measured value of the prognostic or predictive factor; for two classification categories,

$$P[\text{post}] = \frac{\text{prevalence} \cdot \text{likelihood ratio} - \text{prevalence}}{\text{prevalence} \cdot \text{likelihood ratio} + (1 - \text{prevalence})}$$

As discussed in Chapter 3, both of these classification approaches can be extended to take into account multiple classification categories and combinations of factor values. An example of this has been reported by Partin et al. (1997). The authors present tables that list the probability of the pathologic stage of prostate cancer, as determined using a quantitative classification approach, in patients with localized disease. Four pathologic stages are considered: organ-confined disease, isolated capsular penetration, seminal vesicle involvement, and pelvic lymph node involvement. The probabilities depend upon three prognostic factors: the plasma PSA concentration, the TNM stage, and the Gleason score (a histologic grading system for assessing aggressiveness of prostate cancer). These same authors have also generated a nomogram (Figure 11.8) based on the same three prognostic factors that can be used to calculate the probability of recurrence of cancer within five years following radical prostatectomy (Kattan et al. 1998).

**Ex vivo drug sensitivity testing.** The prediction of response to therapy based on plasma markers and tumor cell predictive factors can be very useful in planning chemotherapy for a patient but, until all of the cellular factors that confer susceptibility to any particular drug are known, there will remain uncertainty in the prediction. Directly testing the susceptibility of living tumor cells to drugs—something similar to in vitro antimicrobial susceptibility testing, but for cancer cells—could potentially circumvent this problem and allow for highly reliable individualized chemotherapy. Unfortunately, the very process of studying cancer cells while keeping them alive ex vivo can lead to alterations in the cells or preferential survival of unrepresentative cells, thereby lessening
the reliability of the result of the study. Additional problems with ex vivo studies include variability in the tumor cell sampling process due to intra-tumoral cancer cell heterogeneity and loss of the stromal and vascular setting that defines the tumor's microenvironment in the patient. Despite these obstacles, a number of ex vivo drug sensitivity assays have been developed and new ones continue to be developed (Bellamy 1992, Cree and Kurbacher 1997) although, to date, none of the assays has been shown to be reliable enough for clinical use.

Monitoring
Clinical monitoring of a cancer patient consists of monitoring the patient’s tumor and monitoring the physiologic status of the patient. Tumor monitoring may be undertaken to evaluate the response to therapy, to detect the recurrence of a cancer following successful therapy, or to assess the progress of an established tumor. Monitoring is accomplished primarily by clinical examination, imaging studies, and, for some tumor types, by serial measurement of the plasma concentration of a marker substance. In leukemias, monitoring relies on serial counts of cancer cells in blood. A number of techniques are currently being developed for the detection of micrometastatic carcinoma cells in blood and bone marrow and of cancer cell DNA in plasma (Pantel and von Knebel Doeberitz 2000). Laboratory studies based on these techniques will allow much greater analytic sensitivity in the monitoring of tumors.

Conventional treatment has aimed for eradication of cancer but newer approaches may produce growth control rather than cell death. Therapy directed at suppressing tumor growth may be expected to stabilize the plasma concentration of a tumor marker. If the intent of therapy is to reduce the number of cancer cells, with the hope of totally eliminating the cancerous clone, the plasma concentration of a tumor marker will be expected to decline over time. The rapidity of the decline will be determined by the rate of cell loss in the tumor (in response to the
therapy) and by the plasma half-life of the marker (Bidart et al. 1999). In the extreme, when a tumor has been surgically excised in its entirety, the plasma concentration of the marker will fall exponentially until a new steady-state is reached. As a rule-of-thumb, this requires five or more half-lives of the marker. If the marker is highly tumor-specific, the plasma concentration of the marker will fall to zero (that is, below the lower limit of quantification). The same is true for a marker that is highly organ-specific if the entire organ is removed as a surgical approach to treatment of a tumor. In both cases, if the plasma concentration of the marker remains measurable, it can be inferred that the tumor has not been completely excised. These considerations apply, for instance, to PSA following radical prostatectomy for prostate cancer. PSA has a plasma half-life of two to three days and becomes undetectable in the plasma by 21 days (seven to ten half-lives) after a successful prostatectomy (Bidart et al. 1999). Persistence of a measurable concentration of PSA in the plasma indicates the presence of residual tumor. It is important, when evaluating the success of a surgical tumor excision, not to measure the marker concentration soon after surgery while the marker present in the circulation at the time of surgery has still not been completely cleared from the plasma. Similarly, the plasma concentration of a marker should not be measured too early after treatment by chemotherapy when evaluating the antitumor effect of the therapy because the marker concentration will not have had time to decline to its new steady-state. It is even possible to see early increases in the plasma concentration of a marker if the marker is released into the circulation from dying cancer cells. It is also possible to see modest, transient increases in marker concentration after several months of therapy, if the chemotherapeutic agents reduce the clearance rate of the marker due to liver dysfunction. This happens with CEA occasionally following chemotherapy of colorectal cancer.

Patients may be monitored following therapy in order to detect recurrence of their cancer if early detection of tumor recurrence and institution of therapy will alter the outcome. This is not unlike periodic screening for the appearance of the cancer although the monitoring program may have a different schedule and dissimilar critical values from those used in screening because of the higher probability of disease and because prior therapy can alter the performance of a plasma marker in detecting cancerous tissue. In the case of monitoring PSA following radical prostatectomy, the removal of all hyperplastic (and normal) prostate tissue eliminates the high background plasma PSA concentrations that make the interpretation of PSA concentration in the screening setting so difficult.

In general, when using a tumor-specific marker or an organ-specific marker following excision surgery of the organ, tumor recurrence is indicated by the return of measurable plasma concentrations of the marker. The rapidity with which this happens depends upon the amount of residual tumor from which the recurrent tumor arises and upon the rate of growth of the tumor. The amount of residual tumor is always hoped to be zero, but it can be as large as that associated with a marker concentration just below the limit of quantification. Because of the possibility that, if present, the amount of residual tumor is just shy of that needed for detection by the marker, monitoring is begun soon after the completion of therapy, when even a small amount of tumor growth will have produced a measurable plasma marker concentration. This allows for the earliest possible detection of recurrent disease. The subsequent monitoring schedule is based on clinical experience with the time of appearance of residual tumor following the particular therapy employed. If recurrences tend to appear soon after therapy, intensive post-therapy monitoring is appropriate. If recurrences usually arise sporadically for years following therapy, monitoring may be less frequent but will need to continued for many years.

Most markers used in cancer monitoring are not tumor-specific or even organ-specific, so the detection of tumor recurrence requires identifying a plasma marker concentration, or concentration profile, that can reliably be distinguished from the background marker concentration arising from the non-tumor or non-organ sources of the marker. There are three approaches that can be used in this setting: (1) establishment of a marker concentration that serves as a critical value above which the diagnosis of tumor recurrence is made, (2) use of a significant difference rule, or (3) use of a clinical monitoring rule such as the 2s rule and the 4s rule which can be used to evaluate consecutive monitoring measurements when the findings from the measurements are of concern but do not exceed the significance criterion for a concentration difference. Figure 11.9 (identical to Figure 5.2) shows CEA concentrations in a patient being monitored for
recurrent breast cancer after surgery (data from Winkel et al. 1982). The plasma concentration of the marker was measured monthly and for the first 13 months the concentration was fairly constant. This suggests that the concentrations during that interval represent the background CEA concentration in this patient. The CEA concentration at month 14 is 3.5 µg/L, which appears to be an increase above background but does not exceed the population-based critical value of 7.4 µg/L. The critical value was determined using postoperative CEA concentrations in patients who, after long-term follow-up, didn’t experience a tumor recurrence. If the intraindividual variability in CEA concentration in this particular patient is used as the basis for the calculation of a significant difference in marker concentration, a value of 0.51 µg/L is found (var\_indiv is 0.032 µg²/L²). This formula for a significant difference is based on a specificity of 0.95, which is probably not appropriate given the gravity of the diagnosis here; a specificity of 0.999 is more justified, in which case, a significant difference ≥ 3.3 \sqrt{\text{var\_indiv}} At this level of specificity, a significant difference in CEA concentration is 0.83 µg/L for this patient. This value is less than the observed difference of 2.1 µg/L, so the CEA concentration at month 14 represents a significant increase. Because an increase in CEA concentration can result from conditions affecting any tissue that releases CEA into the circulation (for example, pneumonia or hepatitis), it cannot be assumed that the increase seen in month 14 necessarily implies tumor recurrence. In this way the significant difference rule differs from the population-based critical value approach where CEA concentrations above the critical value specifically imply tumor recurrence and not disorders of other CEA-producing tissues. Because tumor recurrence is one of the possible causes of a significant increase in CEA concentration, clinical evaluation of the patient may be initiated. The continued increases in CEA concentration seen over the subsequent months in this patient are convincing evidence of tumor recurrence. Indeed, she developed clinical signs of tumor soon after the last monitoring specimen was obtained.

**REFERENCES**


Catalona WJ, Richie JP, deKernion JB, Ahmann FR, Ratliff TL, Dalkin BL, Kavoussi LR, MacFarlane MT, and Southwick PC. 1994a. Comparison of prostate specific antigen concentration versus prostate specific antigen density in the early detection of


