Cellular Adaptation, Injury, and Death

5

Sheila Grossman

CELLULAR ADAPTATION
- Atrophy
- Hypertrophy
- Hyperplasia
- Metaplasia
- Dysplasia
- Intracellular Accumulations
- Pathologic Calcifications
  - *Dystrophic Calcification*
  - *Metastatic Calcification*

CELL INJURY AND DEATH
- Causes of Cell Injury
  - *Injury from Physical Agents*
  - *Radiation Injury*
  - *Chemical Injury*
  - *Injury from Biologic Agents*
  - *Injury from Nutritional Imbalances*
- Mechanisms of Cell Injury
  - *Free Radical Injury*
  - *Hypoxic Cell Injury*
  - *Impaired Calcium Homeostasis*
- Reversible Cell Injury and Cell Death
  - *Reversible Cell Injury*
  - *Programmed Cell Death*
  - *Necrosis*
- Cellular Aging

When confronted with stresses that endanger its normal structure and function, the cell undergoes adaptive changes that permit survival and maintenance of function. It is only when the stress is overwhelming or adaptation is ineffective that cell injury and death occur. This chapter focuses on cellular adaptation, injury, and death.

CELLULAR ADAPTATION

After completing this section of the chapter, you should be able to meet the following objectives:

- Cite the general purpose of changes in cell structure and function that occur as the result of normal adaptive processes.
- Describe cell changes that occur with atrophy, hypertrophy, hyperplasia, metaplasia, and dysplasia, and state general conditions under which the changes occur.
- Compare the pathogenesis and effects of dystrophic and metastatic calcifications.

Cells adapt to changes in the internal environment, just as the total organism adapts to changes in the external environment. Cells may adapt by undergoing changes in size, number, and type. These changes, occurring singly or in combination, may lead to

- Atrophy
- Hypertrophy
- Hyperplasia
- Metaplasia
- Dysplasia

Adaptive cellular responses also include intracellular accumulations and storage of products in abnormal amounts.[1]

There are numerous molecular mechanisms mediating cellular adaptation, including factors produced by other cells

or by the cells themselves. These mechanisms depend largely on signals transmitted by chemical messengers that exert their effects by altering gene function. In general, the genes expressed in all cells fall into two categories:

- Operating genes that are necessary for normal function of a cell
- Genes that determine the differentiating characteristics of a particular cell type

In many adaptive cellular responses, the expression of the differentiation genes is altered, whereas that of the operating genes remains unaffected. Thus, a cell is able to change size or form without compromising its normal function. Once the stimulus for adaptation is removed, the effect on expression of the differentiating genes is removed and the cell resumes its previous state of specialized function. Whether adaptive cellular changes are normal or abnormal depends on whether the response was mediated by an appropriate stimulus. Normal adaptive responses occur in response to need and an appropriate stimulus. After the need has been removed, the adaptive response ceases.

KEY POINTS

CELLULAR ADAPTATIONS

- Cells are able to adapt to increased work demands or threats to survival by changing their size (atrophy and hypertrophy), number (hyperplasia), and form (metaplasia).
- Normal cellular adaptation occurs in response to an appropriate stimulus and ceases once the need for adaptation has ceased.

Atrophy

When confronted with a decrease in work demands or adverse environmental conditions, most cells are able to revert to a smaller size and a lower and more efficient level of functioning that is compatible with survival. This decrease in cell size is called **atrophy** and is illustrated in Figure 5.1 regarding atrophy of the endometrium. Cells that are atrophied reduce their oxygen consumption and other cellular functions by decreasing the number and size of their organelles and other structures. There are fewer mitochondria, myofilaments, and endoplasmic reticulum structures. When a sufficient number of cells are involved, the entire tissue or muscle atrophies.

Cell size, particularly in muscle tissue, is related to workload. As the workload of a cell declines, oxygen consumption and protein synthesis decrease. Furthermore, proper muscle mass is maintained by sufficient levels of insulin and insulin-like growth factor-1 (IGF-1).[2] When insulin and IGF-1 levels are low or catabolic signals are present, muscle atrophy occurs by mechanisms that include reduced synthetic processes, increased proteolysis by the ubiquitin–proteasome system, and apoptosis or programmed cell death.[3] In the ubiquitin–proteasome system, intracellular proteins destined for destruction are covalently bonded to a small protein called *ubiquitin* and then degraded by small cytoplasmic organelles called *proteasomes*.[3]

The general causes of atrophy can be grouped into five categories:

1. Disuse
2. Denervation
3. Loss of endocrine stimulation
4. Inadequate nutrition
5. Ischemia or decreased blood flow

Disuse atrophy occurs when there is a reduction in skeletal muscle use. An extreme example of disuse atrophy is seen

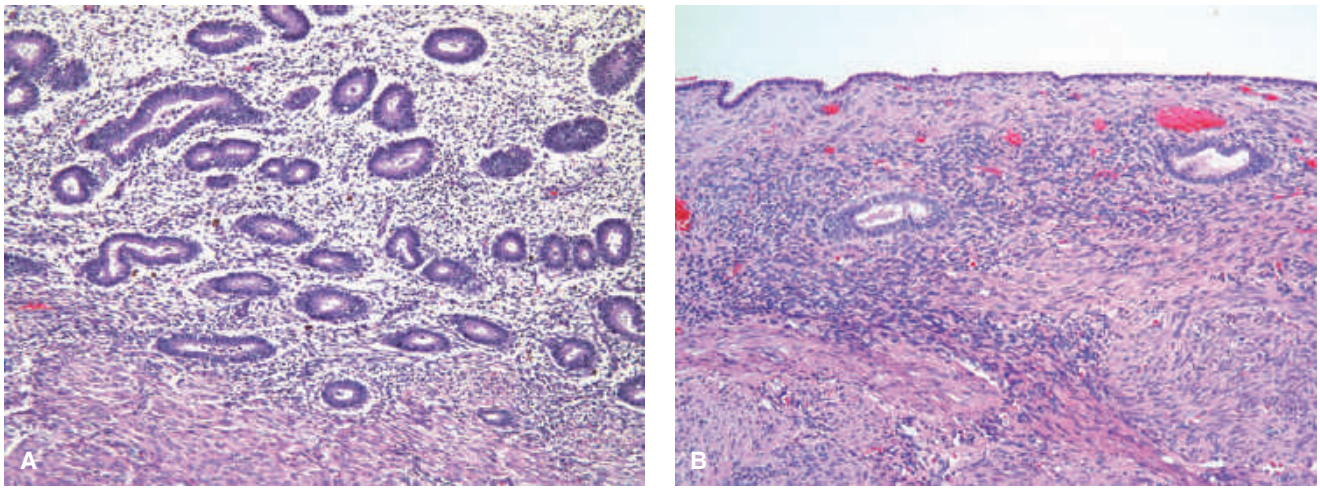


FIGURE 5.1 • Atrophy of cells in endometrium. (**A**) This illustrates a piece of a woman of reproductive age who has a normal thick endometrium. (**B**) This section of endometrium is of a 75-year-old woman that shows atrophic cells and cystic glands. (Both slides are taken at the same magnification.) (From Rubin R., Strayer D. (2012). *Rubin's pathology: Clinicopathologic foundations of medicine* (6th ed., Fig. 1–2, p.3). Philadelphia, PA: Lippincott Williams & Wilkins.)

in the muscles of extremities that have been encased in plaster casts. Because atrophy is adaptive and reversible, muscle size is restored after the cast is removed and muscle use is resumed. Denervation atrophy is a form of disuse atrophy that occurs in the muscles of paralyzed limbs. Lack of endocrine stimulation produces a form of disuse atrophy. In women, the loss of estrogen stimulation during menopause results in atrophic changes in the reproductive organs. With malnutrition and decreased blood flow, cells decrease their size and energy requirements as a means of survival.

Hypertrophy

Hypertrophy represents an increase in cell size and with it an increase in the amount of functioning tissue mass (Fig. 5.2). It results from an increased workload imposed on an organ or body part and is commonly seen in cardiac and skeletal muscle tissue, which cannot adapt to an increase in workload through mitotic division and formation of more cells.[1] Hypertrophy involves an increase in the functional components of the cell that allows it to achieve equilibrium between demand and functional capacity. For example, as muscle cells hypertrophy, additional actin and myosin filaments, cell enzymes, and adenosine triphosphate (ATP) are synthesized.[1,4]

Hypertrophy may occur as the result of normal physiologic or abnormal pathologic conditions. The increase in muscle mass associated with exercise is an example of physiologic hypertrophy. Pathologic hypertrophy occurs as the result of disease conditions and may be adaptive or compensatory. Examples of adaptive hypertrophy are the thickening of the urinary bladder from long-continued obstruction of urinary outflow and the myocardial hypertrophy that results from valvular heart disease or hypertension. Compensatory hypertrophy is the enlargement of a remaining organ or tissue after a portion has been surgically removed or rendered inactive. For instance, if one kidney is removed, the remaining kidney enlarges to compensate for the loss.

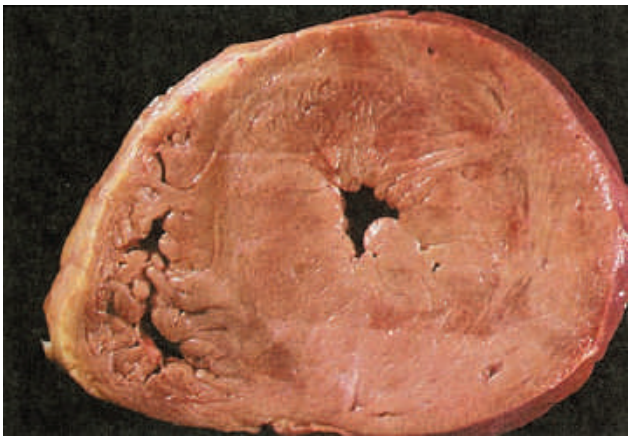


FIGURE 5.2 • Myocardial hypertrophy. Cross section of the heart with left ventricular hypertrophy. (From Rubin R., Strayer D. (2012). *Rubin's pathology: Clinicopathologic foundations of medicine* (6th ed., p. 4). Philadelphia, PA: Lippincott Williams & Wilkins.)

The initiating signals for hypertrophy appear to be complex and related to ATP depletion, mechanical forces such as stretching of the muscle fibers, activation of cell degradation products, and hormonal factors.[4] In the case of the heart, initiating signals can be divided into two broad categories:

1. Biomechanical and stretch-sensitive mechanisms
2. Neurohumoral mechanisms that are associated with the release of hormones, growth factors, cytokines, and chemokines[5]

Internal stretch-sensitive receptors for the biochemical signals and an array of membrane-bound receptors for the specific neurohumoral ligands, such as IGF-1 and epidermal growth factor (EGF), activate specific signal transduction pathways.[5] These pathways control myocardial growth by altering gene expression to increase protein synthesis and reduce protein degradation, thereby causing hypertrophic enlargement of the heart. A limit is eventually reached beyond which further enlargement of the tissue mass can no longer compensate for the increased work demands. The limiting factors for continued hypertrophy might be related to limitations in blood flow. In hypertension, for example, the increased workload required to pump blood against an elevated arterial pressure in the aorta results in a progressive increase in left ventricular muscle mass and need for coronary blood flow.

There continues to be interest in the signaling pathways that control the arrangement of contractile elements in myocardial hypertrophy. Research suggests that certain signal molecules can alter gene expression controlling the size and assembly of the contractile proteins in hypertrophied myocardial cells. For example, the hypertrophied myocardial cells of well-trained athletes have proportional increases in width and length. This is in contrast to the hypertrophy that develops in dilated cardiomyopathy, in which the hypertrophied cells have a relatively greater increase in length than width. In pressure overload, as occurs with hypertension, the hypertrophied cells have greater width than length.[5] It is anticipated that further elucidation of the signal pathways that determine the adaptive and nonadaptive features of cardiac hypertrophy will lead to new targets for treatment.

Hyperplasia

Hyperplasia refers to an increase in the number of cells in an organ or tissue. It occurs in tissues with cells that are capable of mitotic division, such as the epidermis, intestinal epithelium, and glandular tissue.[1] Certain cells, such as neurons, rarely divide and therefore have little capacity, if any, for hyperplastic growth. There is evidence that hyperplasia involves activation of genes controlling cell proliferation and the presence of intracellular messengers that control cell replication and growth. As with other normal adaptive cellular responses, hyperplasia is a controlled process that occurs in response to an appropriate stimulus and ceases after the stimulus has been removed.

The stimuli that induce hyperplasia may be physiologic or nonphysiologic. There are two common types of physiologic hyperplasia: hormonal and compensatory. Breast and uterine enlargements during pregnancy are examples of a physiologic hyperplasia that results from estrogen stimulation. The regeneration of the liver that occurs after partial hepatectomy (*i.e.,* partial removal of the liver) is an example of compensatory hyperplasia. Hyperplasia is also an important response of connective tissue in wound healing, during which proliferating fibroblasts and blood vessels contribute to wound repair. Although hypertrophy and hyperplasia are two distinct processes, they may occur together and are often triggered by the same mechanism.[1] For example, the pregnant uterus undergoes both hypertrophy and hyperplasia as the result of estrogen stimulation.

Most forms of nonphysiologic hyperplasia are due to excessive hormonal stimulation or the effects of growth factors on target tissues.[2] The public seems to appreciate that a laboratory finding including the term hyperplasia generally is something to take seriously. For example, excessive estrogen production can cause endometrial hyperplasia and abnormal menstrual bleeding. Endometrial hyperplasia is considered a high risk for developing endometrial cancer and is a condition that is monitored carefully.[6] Benign prostatic hyperplasia (BPH), which is a common disorder of men older than 50 years of age, is related to the action of androgens. BPH is a nonmalignant condition that causes lower urinary tract symptoms. BPH sometimes develops into prostate cancer.[2,7] Women with atypical hyperplasia of the breast are also monitored carefully since they have a fourfold higher risk of developing ductal carcinoma in situ or invasive breast cancer.[8] Skin warts are another example of hyperplasia caused by growth factors produced by certain viruses, such as the papillomaviruses.

Metaplasia

Metaplasia represents a reversible change in which one adult cell type (epithelial or mesenchymal) is replaced by another adult cell type. Metaplasia is thought to involve the reprogramming of undifferentiated stem cells that are present in the tissue undergoing the metaplastic changes.[1]

Metaplasia usually occurs in response to chronic irritation and inflammation and allows for substitution of cells that are better able to survive under circumstances in which a more fragile cell type might succumb. However, the conversion of cell types never oversteps the boundaries of the primary tissue type (*e.g.,* one type of epithelial cell may be converted to another type of epithelial cell, but not to a connective tissue cell). An example of metaplasia is the adaptive substitution of stratified squamous epithelial cells for the ciliated columnar epithelial cells in the trachea and large airways of a habitual cigarette smoker. Barrett esophagus is a premalignant condition that occurs in the esophagus of people with chronic gastroesophageal reflux disease (GERD). It is characterized by normal squamous epithelium in the lower esophagus transforming into columnar-lined epithelium. Barrett esophagus is the primary risk factor for developing esophageal adenocarcinoma.[9]

Dysplasia

Dysplasia is characterized by deranged cell growth of a specific tissue that results in cells that vary in size, shape, and organization. Minor degrees of dysplasia are associated with chronic irritation or inflammation. The pattern is most frequently encountered in areas of metaplastic squamous epithelium of the respiratory tract and uterine cervix. Although dysplasia is abnormal, it is adaptive in that it is potentially reversible after the irritating cause has been removed. Dysplasia is strongly implicated as a precursor of cancer.[1] In cancers of the respiratory tract and the uterine cervix, dysplastic changes have been found adjacent to the foci of cancerous transformation. Through the use of the Papanicolaou (Pap) smear, it has been documented that cancer of the uterine cervix develops in a series of incremental epithelial changes ranging from severe dysplasia to invasive cancer. However, dysplasia is an adaptive process and as such does not necessarily lead to cancer.

Preterm babies who are ventilated for long periods of time due to their prematurity and lack of surfactant, and term infants who require intubation and ventilated oxygen in the first month of life, often develop bronchopulmonary dysplasia (BPD).[10] In fact there are more preterm babies surviving today, so more BPD is evident. Approximately 20% of infants born at less than 30 weeks' gestation and weighing less than 1500 g develop BPD.[10] Although there has been some excellent therapy that has decreased some of the negative lung disease experienced by infants with BPD, many infants who develop BPD experience the long-term effects of alveolar destruction the rest of their lives.[1,10]

Intracellular Accumulations

Intracellular accumulations represent the buildup of substances that cells cannot immediately use or eliminate. The substances may accumulate in the cytoplasm (frequently in the lysosomes) or in the nucleus. In some cases the accumulation may be an abnormal substance that the cell has produced, and in other cases the cell may be storing exogenous materials or products of pathologic processes occurring elsewhere in the body. An example would be the accumulation of beta amyloid fragments, which progress to a skeletal muscle disorder called myositis.[11]

These substances may accumulate transiently or permanently, and they may be harmless or, in some cases, toxic. These substances can be grouped into three categories:

1. Normal body substances, such as lipids, proteins, carbohydrates, melanin, and bilirubin, that are present in abnormally large amounts
2. Abnormal endogenous products, such as those resulting from inborn errors of metabolism
3. Exogenous products, such as environmental agents and pigments, that cannot be broken down by the cell[2]

The accumulation of normal cellular constituents occurs when a substance is produced at a rate that exceeds its metabolism or removal. An example of this type of process is fatty changes in the liver due to intracellular accumulation of triglycerides. Liver cells normally contain some fat, which is either oxidized and used

for energy or converted to triglycerides. This fat is derived from free fatty acids released from adipose tissue. Abnormal accumulation occurs when the delivery of free fatty acids to the liver is increased, as in starvation and diabetes mellitus, or when the intrahepatic metabolism of lipids is disturbed, as in alcoholism.

Intracellular accumulation can result from genetic disorders that disrupt the metabolism of selected substances. A normal enzyme may be replaced with an abnormal one, resulting in the formation of a substance that cannot be used or eliminated from the cell, or an enzyme may be missing, so that an intermediate product accumulates in the cell. For example, there are at least 10 genetic disorders that affect glycogen metabolism, most of which lead to the accumulation of intracellular glycogen stores. In the most common form of this disorder, von Gierke disease, large amounts of glycogen accumulate in the liver and kidneys because of a deficiency of the enzyme glucose-6-phosphatase. Without this enzyme, glycogen cannot be broken down to form glucose. The disorder leads not only to an accumulation of glycogen but also to a reduction in blood glucose levels. In Tay-Sachs disease, another genetic disorder, abnormal lipids accumulate in the brain and other tissues, causing motor and mental deterioration beginning at approximately 6 months of age, followed by death at 2 to 5 years of age. In a similar manner, other enzyme defects lead to the accumulation of other substances.

Pigments are colored substances that may accumulate in cells. They can be endogenous (*i.e.,* arising from within the body) or exogenous (*i.e.,* arising from outside the body). Icterus, also called *jaundice*, is characterized by a yellow discoloration of tissue due to the retention of bilirubin, an endogenous bile pigment. This condition may result from increased bilirubin production from red blood cell destruction, obstruction of bile passage into the intestine, or toxic diseases that affect the liver's ability to remove bilirubin from the blood. Lipofuscin is a yellow-brown pigment that results from the accumulation of the indigestible residues produced during normal turnover of cell structures (Fig. 5.3). The accumulation of lipofuscin increases with age, and it is sometimes referred to as the *wear-and-tear pigment*. It is more common in heart, nerve, and liver cells than other tissues and is seen more often in conditions associated with atrophy of an organ.

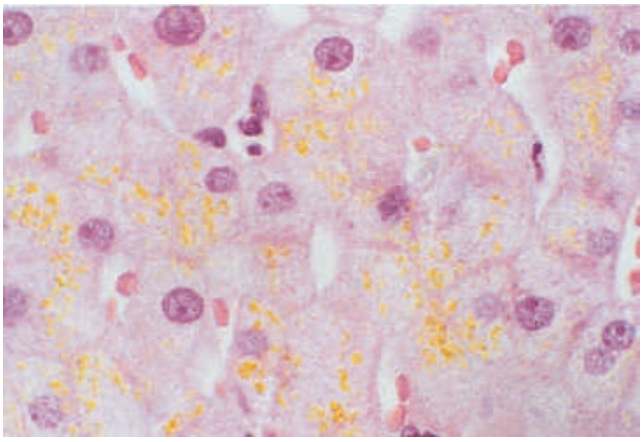


FIGURE 5.3 • Accumulation of intracellular lipofuscin. A photomicrograph of the liver of an 80-year-old man shows golden cytoplasmic granules, which represent lysosomal storage of lipofuscin. (From Rubin R., Strayer D. (2012). *Rubin's pathology: Clinicopathologic foundations of medicine* (6 ed., p. 121). Philadelphia, PA: Lippincott Williams & Wilkins.)

One of the most common exogenous pigments is carbon in the form of coal dust. In coal miners or people exposed to heavily polluted environments, the accumulation of carbon dust blackens the lung tissue and may cause serious lung disease. The formation of a blue lead line along the margins of the gum is one of the diagnostic features of lead poisoning. Tattoos are the result of insoluble pigments introduced into the skin, where they are engulfed by macrophages and persist for a lifetime.

The significance of intracellular accumulations depends on the cause and severity of the condition. Many accumulations, such as lipofuscin and mild fatty change, have no effect on cell function. Some conditions, such as the hyperbilirubinemia that causes jaundice, are reversible. Other disorders, such as glycogen storage diseases, produce accumulations that result in organ dysfunction and other alterations in physiologic function.

Pathologic Calcifications

Pathologic calcification involves the abnormal tissue deposition of calcium salts, together with smaller amounts of iron, magnesium, and other minerals. It is known as *dystrophic calcification* when it occurs in dead or dying tissue and as *metastatic calcification* when it occurs in normal tissue.

Dystrophic Calcification

Dystrophic calcification represents the macroscopic deposition of calcium salts in injured tissue.[12,13] It is often visible to the naked eye as deposits that range from gritty, sandlike grains to firm, hard rock material. The pathogenesis of dystrophic calcification involves the intracellular or extracellular formation of crystalline calcium phosphate. The components of the calcium deposits are derived from the bodies of dead or dying cells as well as from the circulation and interstitial fluid.

Dystrophic calcification is commonly seen in atheromatous lesions of advanced atherosclerosis, areas of injury in the aorta and large blood vessels, and damaged heart valves. Although the presence of calcification may only indicate the presence of previous cell injury, as in healed tuberculosis lesions, it is also a frequent cause of organ dysfunction. For example, calcification of the aortic valve is a frequent cause of aortic stenosis in older adults (Fig. 5.4).

Metastatic Calcification

In contrast to dystrophic calcification, which occurs in injured tissues, metastatic calcification occurs in normal tissues as the result of increased serum calcium levels (hypercalcemia). Almost any condition that increases the serum calcium level can lead to calcification in inappropriate sites such as the lung, renal tubules, and blood vessels. The major causes of hypercalcemia are hyperparathyroidism, either primary or secondary to phosphate retention in renal failure; increased mobilization of calcium from bone as in Paget disease, cancer with metastatic bone lesions, or immobilization; and vitamin D intoxication.[14]

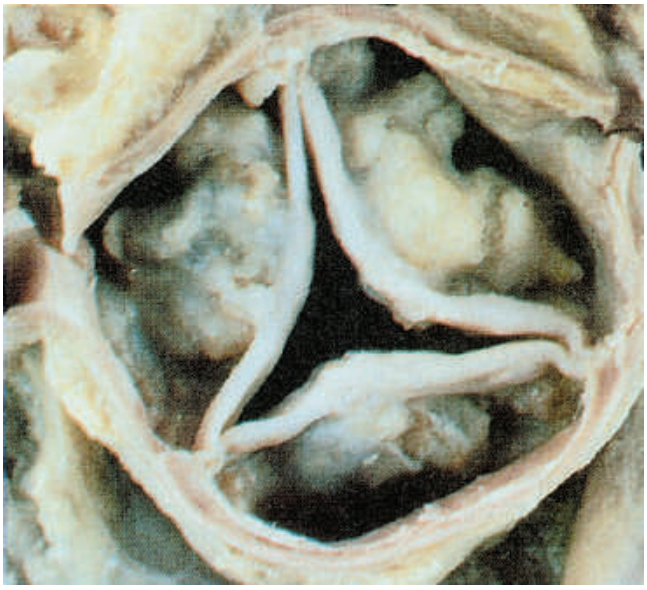


FIGURE 5.4 • Calcific aortic stenosis. Large deposits of calcium salts are evident in the cusps and free margins of the thickened aortic valve as viewed from above. (From Strayer D. S., Rubin E. (2008). Cell injury. In Rubin R., Strayer D. S. (Eds.), *Rubin's pathology: Clinicopathologic foundations of medicine* (5th ed., p. 13). Philadelphia, PA: Lippincott Williams & Wilkins.)

IN SUMMARY

Cells adapt to changes in their environment and in their work demands by changing their size, number, and characteristics. These adaptive changes are consistent with the needs of the cell and occur in response to an appropriate stimulus. The changes are usually reversed after the stimulus has been withdrawn.

When confronted with a decrease in work demands or adverse environmental conditions, cells atrophy or reduce their size and revert to a lower and more efficient level of functioning. Hypertrophy results from an increase in work demands and is characterized by an increase in tissue size brought about by an increase in cell size and functional intracellular components. An increase in the number of cells in an organ or tissue that is still capable of mitotic division is called *hyperplasia*. Metaplasia occurs in response to chronic irritation and represents the substitution of cells of a type that is better able to survive under circumstances in which a more fragile cell type might succumb. Dysplasia is characterized by deranged cell growth of a specific tissue that results in cells that vary in size, shape, and appearance. It is often a precursor of cancer.

Under some circumstances, cells may accumulate abnormal amounts of various substances. If the accumulation reflects a correctable systemic disorder, such as the hyperbilirubinemia that causes jaundice, the accumulation is reversible. If the disorder cannot be corrected, as often occurs in many inborn errors of metabolism, the cells become overloaded, causing cell injury and death.

Pathologic calcification involves the abnormal tissue deposition of calcium salts. Dystrophic calcification occurs in dead or dying tissue. Although the presence of dystrophic calcification may only indicate the presence of previous cell injury, it is also a frequent cause of organ dysfunction (*e.g.,* when it affects the heart valves). Metastatic calcification occurs in normal tissues as the result of elevated serum calcium levels. Almost any condition that increases the serum calcium level can lead to calcification in inappropriate sites such as the lung, renal tubules, and blood vessels.

CELL INJURY AND DEATH

After completing this section of the chapter, you should be able to meet the following objectives:

- Describe the mechanisms whereby physical agents such as blunt trauma, electrical forces, and extremes of temperature produce cell injury.
- Differentiate between the effects of ionizing and nonionizing radiation in terms of their ability to cause cell injury.
- State the mechanisms and manifestations of cell injury associated with lead poisoning.
- Relate free radical formation and oxidative stress to cell injury and death.

Cells can be injured in many ways. The extent to which any injurious agent can cause cell injury and death depends in large measure on the intensity and duration of the injury and the type of cell that is involved. Cell injury is usually reversible up to a certain point, after which irreversible cell injury and death occur. Whether a specific stress causes irreversible or reversible cell injury depends on the severity of the insult and on variables such as blood supply, nutritional status, and regenerative capacity. Cell injury and death are ongoing processes, and in the healthy state, they are balanced by cell renewal.

KEY POINTS

CELL INJURY

- Cells can be damaged in a number of ways, including physical trauma, extremes of temperature, electrical injury, exposure to damaging chemicals, radiation damage, injury from biologic agents, and nutritional factors.
- Most injurious agents exert their damaging effects through uncontrolled free radical production, impaired oxygen delivery or utilization, or the destructive effects of uncontrolled intracellular calcium release.

Causes of Cell Injury

Cell damage can occur in many ways. For purposes of discussion, the ways by which cells are injured have been grouped into five categories:

1. Injury from physical agents
2. Radiation injury
3. Chemical injury
4. Injury from biologic agents
5. Injury from nutritional imbalances

Injury from Physical Agents

Physical agents responsible for cell and tissue injury include mechanical forces, extremes of temperature, and electrical forces. They are common causes of injuries due to environmental exposure, occupational and transportation accidents, and physical violence and assault.

Mechanical Forces. Injury or trauma due to mechanical forces occurs as the result of body impact with another object. The body or the mass can be in motion or, as sometimes happens, both can be in motion at the time of impact. These types of injuries split and tear tissue, fracture bones, injure blood vessels, and disrupt blood flow.

Extremes of Temperature. Extremes of heat and cold cause damage to the cell, its organelles, and its enzyme systems. Exposure to low-intensity heat (43°C to 46°C), such as occurs with partial-thickness burns and severe heat stroke, causes cell injury by inducing vascular injury, accelerating cell metabolism, inactivating temperature-sensitive enzymes, and disrupting the cell membrane. With more intense heat, coagulation of blood vessels and tissue proteins occurs. Exposure to cold increases blood viscosity and induces vasoconstriction by direct action on blood vessels and through reflex activity of the sympathetic nervous system. The resultant decrease in blood flow may lead to hypoxic tissue injury, depending on the degree and duration of cold exposure. Injury from freezing probably results from a combination of ice crystal formation and vasoconstriction. The decreased blood flow leads to capillary stasis and arteriolar and capillary thrombosis. Edema results from increased capillary permeability.

Electrical Injuries. Electrical injuries can affect the body through extensive tissue injury and disruption of neural and cardiac impulses. Voltage, type of current, amperage, pathway of the current, resistance of the tissue, and interval of exposure determine the effect of electricity on the body.[15]

Alternating current (AC) is usually more dangerous than direct current (DC) because it causes violent muscle contractions, preventing the person from releasing the electrical source and sometimes resulting in fractures and dislocations. In electrical injuries, the body acts as a conductor of the electrical current.[15] The current enters the body from an electrical source, such as an exposed wire, and passes through the body and exits to another conductor, such as the moisture on the ground or a piece of metal the person is holding. The pathway that a current takes is critical because the electrical energy disrupts impulses in excitable tissues. Current flow through the brain may interrupt impulses from respiratory centers in the brain stem, and current flow through the chest may cause fatal cardiac arrhythmias.

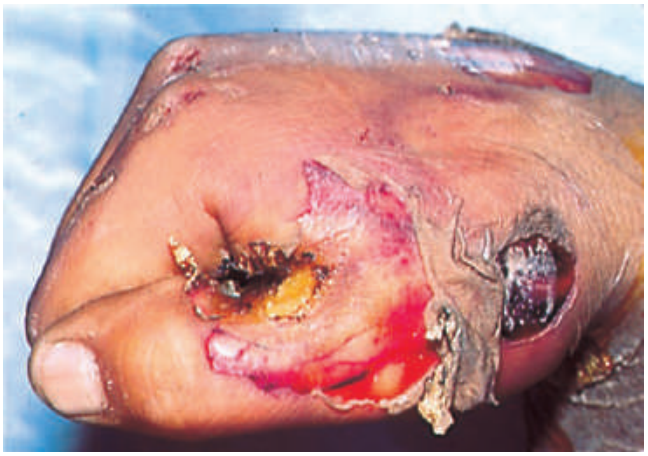


FIGURE 5.5 • Electrical burn of the skin. The victim was electrocuted after attempting to stop a fall from a ladder by grasping a high-voltage line. (From McConnell T., Hull K. (2011). *Human form human function: Essentials of anatomy & physiology* (p. 158). Philadelphia, PA: Lippincott Williams & Wilkins.)

The resistance to the flow of current in electrical circuits transforms electrical energy into heat. This is why the elements in electrical heating devices are made of highly resistive metals. Much of the tissue damage produced by electrical injuries is caused by heat production in tissues that have the highest electrical resistance. Resistance to electrical current varies from the greatest to the least in bone, fat, tendons, skin, muscles, blood, and nerves. The most severe tissue injury usually occurs at the skin sites where the current enters and leaves the body (Fig. 5.5). After electricity has penetrated the skin, it passes rapidly through the body along the lines of least resistance—through body fluids and nerves. Degeneration of vessel walls may occur, and thrombi may form as current flows along the blood vessels. This can cause extensive muscle and deep tissue injury. Thick, dry skin is more resistant to the flow of electricity than thin, wet skin. It is generally believed that the greater the skin resistance, the greater is the amount of local skin burn, and the less the resistance, the greater are the deep and systemic effects.

Radiation Injury

Electromagnetic radiation comprises a wide spectrum of wave-propagated energy, ranging from ionizing gamma rays to radiofrequency waves (Fig. 5.6). A photon is a particle of radiation energy. Radiation energy above the ultraviolet (UV) range is called *ionizing radiation* because the photons have enough energy to knock electrons off atoms and molecules. *Nonionizing radiation* refers to radiation energy at frequencies below those of visible light. *UV radiation* represents the portion of the

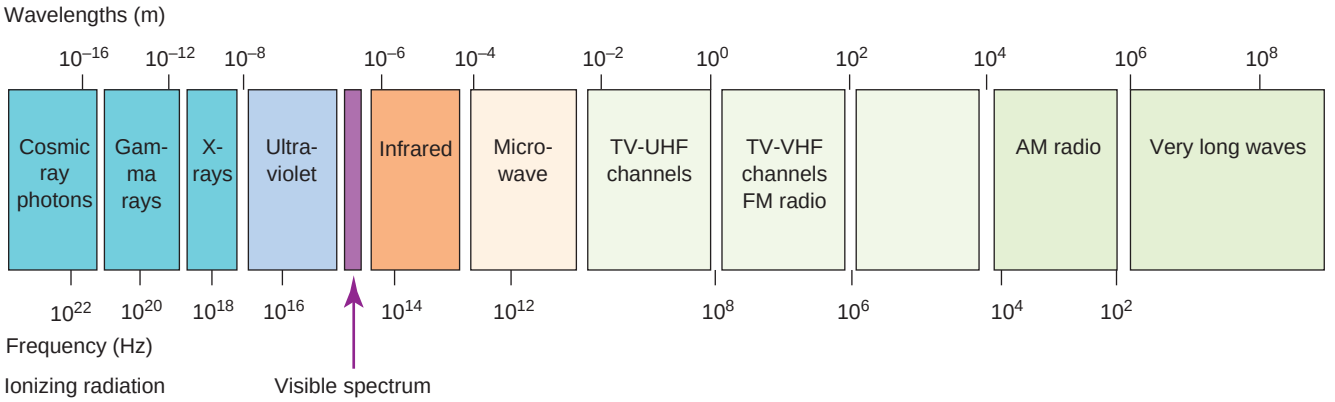


FIGURE 5.6 • Spectrum of electromagnetic radiation.

spectrum of electromagnetic radiation just above the visible range.[15] It contains increasingly energetic rays that are powerful enough to disrupt intracellular bonds and cause sunburn.

Ionizing Radiation. Ionizing radiation impacts cells by causing ionization of molecules and atoms in the cell. This is accomplished by releasing free radicals that destroy cells and by directly hitting the target molecules in the cell.[16] It can immediately kill cells, interrupt cell replication, or cause a variety of genetic mutations, which may or may not be lethal. Most radiation injury is caused by localized irradiation that is used in the treatment of cancer. Except for unusual circumstances such as the use of high-dose irradiation that precedes bone marrow transplantation, exposure to whole-body irradiation is rare.

The injurious effects of ionizing radiation vary with the dose, dose rate (a single dose can cause greater injury than divided or fractionated doses), and the differential sensitivity of the exposed tissue to radiation injury. Because of the effect on deoxyribonucleic acid (DNA) synthesis and interference with mitosis, rapidly dividing cells of the bone marrow and intestine are much more vulnerable to radiation injury than tissues such as bone and skeletal muscle. Over time, occupational and accidental exposure to ionizing radiation can result in increased risk for the development of various types of cancers, including skin cancers, leukemia, osteogenic sarcomas, and lung cancer. This is especially true when the person is exposed to radiation during childhood.[17]

Many of the clinical manifestations of radiation injury result from acute cell injury, dose-dependent changes in the blood vessels that supply the irradiated tissues, and fibrotic tissue replacement. The cell's initial response to radiation injury involves swelling, disruption of the mitochondria and other organelles, alterations in the cell membrane, and marked changes in the nucleus. The endothelial cells in blood vessels are particularly sensitive to irradiation. During the immediate postirradiation period, only vessel dilatation is apparent (*e.g.,* the initial erythema of the skin after radiation therapy). Later or with higher levels of radiation, destructive changes occur in small blood vessels such as the capillaries and venules. Acute reversible necrosis is represented by such disorders as radiation cystitis, dermatitis, and diarrhea from enteritis. More persistent damage can be attributed to acute necrosis of tissue cells that are not capable of regeneration and chronic ischemia. Chronic effects of radiation damage are characterized by fibrosis and scarring of tissues and organs in the irradiated area (*e.g.,* interstitial fibrosis of the heart and lungs after irradiation of the chest). Because the radiation delivered in radiation therapy inevitably travels through the skin, radiation dermatitis is common. There may be necrosis of the skin, impaired wound healing, and chronic radiation dermatitis.

Ultraviolet Radiation. Ultraviolet radiation causes sunburn and increases the risk of skin cancers. The degree of risk depends on the type of UV rays, the intensity of exposure, and the amount of protective melanin pigment in the skin. Skin damage produced by UV radiation is thought to be caused by reactive oxygen species (ROS) and by damage to melanin-producing processes in the skin.[18] UV radiation also damages DNA, resulting in the formation of pyrimidine dimers (*i.e.,* the insertion of two identical pyrimidine bases into replicating DNA instead of one). Other forms of DNA damage include the production of single-stranded breaks and formation of DNA–protein cross-links. Normally errors that occur during DNA replication are repaired by enzymes that remove the faulty section of DNA and repair the damage. The importance of DNA repair in protecting against UV radiation injury is evidenced by the vulnerability of people who lack the enzymes needed to repair UV-induced DNA damage. In a genetic disorder called *xeroderma pigmentosum*, an enzyme needed to repair sunlight-induced DNA damage is lacking. This autosomal recessive disorder is characterized by extreme photosensitivity and an increased risk of skin cancer in sun-exposed skin.[19]

Nonionizing Radiation. Nonionizing radiation includes infrared light, ultrasound, microwaves, and laser energy. Unlike ionizing radiation, which can directly break chemical bonds, nonionizing radiation exerts its effects by causing vibration and rotation of atoms and molecules.[15] All of

this vibrational and rotational energy is eventually converted to thermal energy. Low-frequency nonionizing radiation is used widely in radar, television, industrial operations (*e.g.,* heating, welding, melting of metals, processing of wood and plastic), household appliances (*e.g.,* microwave ovens), and medical applications (*e.g.,* diathermy). Isolated cases of skin burns and thermal injury to deeper tissues have occurred in industrial settings and from improperly used household microwave ovens. Injury from these sources is mainly thermal and, because of the deep penetration of the infrared or microwave rays, tends to involve dermal and subcutaneous tissue injury.

Chemical Injury

Chemicals capable of damaging cells are everywhere around us. Air and water pollution contains chemicals capable of tissue injury, as does tobacco smoke and some processed or preserved foods. Some of the most damaging chemicals exist in our environment, including gases such as carbon monoxide, insecticides, and trace metals such as lead.

Chemical agents can injure the cell membrane and other cell structures, block enzymatic pathways, coagulate cell proteins, and disrupt the osmotic and ionic balance of the cell. Corrosive substances such as strong acids and bases destroy cells as the substances come into contact with the body. Other chemicals may injure cells in the process of metabolism or elimination. Carbon tetrachloride (CCl_4), for example, causes little damage until it is metabolized by liver enzymes to a highly reactive free radical ($CCl_3\bullet$). Carbon tetrachloride is extremely toxic to liver cells.[20]

Drugs. Many drugs—alcohol, prescription drugs, over-the-counter drugs, and street drugs—are capable of directly or indirectly damaging tissues. Ethyl alcohol can harm the gastric mucosa, liver, developing fetus, and other organs. Antineoplastic and immunosuppressant drugs can directly injure cells. Other drugs produce metabolic end products that are toxic to cells. Acetaminophen, a commonly used over-the-counter analgesic drug, is detoxified in the liver, where small amounts of the drug are converted to a highly toxic metabolite. This metabolite is detoxified by a metabolic pathway that uses a substance (*i.e.,* glutathione) normally present in the liver. When large amounts of the drug are ingested, this pathway becomes overwhelmed and toxic metabolites accumulate, causing massive liver necrosis.

Lead Toxicity. Lead is a particularly toxic metal. Small amounts accumulate to reach toxic levels. There are innumerable sources of lead in the environment, including flaking paint, lead-contaminated dust and soil, lead-contaminated root vegetables, lead water pipes or soldered joints, pottery glazes, newsprint, and toys made in foreign countries. Adults often encounter lead through occupational exposure. Children are exposed to lead through ingestion of peeling lead paint, by breathing dust from lead paint, or from playing in contaminated soil. There has been a decline in blood lead levels of both adults and children since the removal of lead from gasoline and from soldered food cans.[21] High lead blood levels continue to be a problem, however, particularly among children. In the United States alone, there are approximately 250,000 children between 1 and 5 years of age who have lead levels greater than 10 μg/mL.[22] The prevalence of elevated blood lead levels was higher for children living in more urbanized areas. By race or ethnicity, non-Hispanic Black children residing in central cities with a population of 1 million or more had the highest proportion of elevated blood lead levels.

Lead is absorbed through the gastrointestinal tract or the lungs into the blood. A deficiency in calcium, iron, or zinc increases lead absorption. In children, most lead is absorbed through the lungs. Although children may have the same or a lower intake of lead, the absorption in infants and children is greater; thus, they are more vulnerable to lead toxicity.[22] Lead crosses the placenta, exposing the fetus to levels of lead that are comparable with those of the mother. Lead is stored in bone and eliminated by the kidneys. Although the half-life of lead is hours to days, bone deposits serve as a repository from which blood levels are maintained. In a sense, bone protects other tissues, but the slow turnover maintains blood levels for months to years.

The toxicity of lead is related to its multiple biochemical effects. It has the ability to inactivate enzymes, compete with calcium for incorporation into bone, and interfere with nerve transmission and brain development. The major targets of lead toxicity are the red blood cells, the gastrointestinal tract, the kidneys, and the nervous system.

Anemia is a cardinal sign of lead toxicity. Lead competes with the enzymes required for hemoglobin synthesis and with the membrane-associated enzymes that prevent hemolysis of red blood cells. The resulting red cells are coarsely stippled and hypochromic, resembling those seen in iron-deficiency anemia. The life span of the red cell is also decreased. The gastrointestinal tract is the main source of symptoms in the adult. This is characterized by "lead colic," a severe and poorly localized form of acute abdominal pain. A lead line formed by precipitated lead sulfite may appear along the gingival margins. The lead line is seldom seen in children. The kidneys are the major route for excretion of lead. Lead can cause diffuse kidney damage, eventually leading to renal failure. Even without overt signs of kidney damage, lead toxicity leads to hypertension.

In the nervous system, lead toxicity is characterized by demyelination of cerebral and cerebellar white matter and death of cortical cells. When this occurs in early childhood, it can affect neurobehavioral development and result in lower IQ levels and poorer classroom performance.[11] Peripheral demyelinating neuropathy may occur in adults. The most serious manifestation of lead poisoning is acute encephalopathy. It is manifested by persistent vomiting, ataxia, seizures, papilledema, impaired consciousness, and coma. Acute encephalopathy may manifest suddenly, or it may be preceded by other signs of lead toxicity such as behavioral changes or abdominal complaints.

Because of the long-term neurobehavioral and cognitive deficits that occur in children with even moderately elevated lead levels, the Centers for Disease Control and Prevention have issued recommendations for childhood lead screening.[22] A safe blood level of lead is still uncertain. At one time, 25 µg/dL was considered safe. Surveys have shown abnormally low IQs in children with lead levels as low as 10 to 15 µg/dL.

Screening for lead toxicity involves use of capillary blood obtained from a finger stick to measure free erythrocyte protoporphyrin (EP). Elevated levels of EP result from the inhibition by lead of the enzymes required for heme synthesis in red blood cells. The EP test is useful in detecting high lead levels but usually does not detect levels below 20 to 25 µg/dL. Thus, capillary screening test values greater than 10 µg/dL should be confirmed with those from a venous blood sample. Because the symptoms of lead toxicity usually are vague, diagnosis is often delayed. Anemia may provide the first clues to the disorder. Laboratory tests are necessary to establish a diagnosis. Treatment involves removal of the lead source and, in cases of severe toxicity, administration of a chelating agent. Asymptomatic children with blood levels of 45 to 69 µg/dL usually are treated. A public health team should evaluate the source of lead because meticulous removal is needed.

Mercury Toxicity. Mercury has been used for industrial and medical purposes for hundreds of years. Mercury is toxic, and the hazards of mercury-associated occupational and accidental exposures are well known. Currently mercury and lead are the most toxic metals. Mercury is toxic in four primary forms: mercury vapor, inorganic divalent mercury, methyl mercury, and ethyl mercury.[23] Depending on the form of mercury exposure, toxicity involving the central nervous system and kidney can occur.[24]

In the case of dental fillings, the concern involves mercury vapor being released into the mouth. However, the amount of mercury vapor released from fillings is very small. The main source of methyl mercury exposure is from consumption of long-lived fish, such as tuna and swordfish. Fish concentrate mercury from sediment in the water. Only certain types of fish pose potential risk, however, and types such as salmon have miniscule amounts or no mercury. Because the developing brain is more susceptible to mercury-induced damage, it is recommended that young children and pregnant and nursing women avoid consumption of fish known to contain high mercury content. Thimerosal is an ethyl mercury–containing preservative that helps prevent microorganism growth in vaccines. Due to the concern of this preservative, it is hardly ever used in the United States.

Injury from Biologic Agents

Biologic agents differ from other injurious agents in that they are able to replicate and can continue to produce their injurious effects. These agents range from submicroscopic viruses to the larger parasites. Biologic agents injure cells by diverse mechanisms. Viruses enter the cell and become incorporated into its DNA synthetic machinery. Certain bacteria elaborate exotoxins that interfere with cellular production of ATP. Other bacteria, such as the gram-negative bacilli, release endotoxins that cause cell injury and increased capillary permeability.

Injury from Nutritional Imbalances

Nutritional excesses and nutritional deficiencies predispose cells to injury. Obesity and diets high in saturated fats are thought to predispose persons to atherosclerosis. The body requires more than 60 organic and inorganic substances in amounts ranging from micrograms to grams. These nutrients include minerals, vitamins, certain fatty acids, and specific amino acids. Dietary deficiencies can occur in the form of starvation, in which there is a deficiency of all nutrients and vitamins, or because of a selective deficiency of a single nutrient or vitamin. Iron-deficiency anemia, scurvy, beriberi, and pellagra are examples of injury caused by the lack of specific vitamins or minerals. The protein and calorie deficiencies that occur with starvation cause widespread tissue damage.

Mechanisms of Cell Injury

The mechanisms by which injurious agents cause cell injury and death are complex. Some agents, such as heat, produce direct cell injury. Other factors, such as genetic derangements, produce their effects indirectly through metabolic disturbances and altered immune responses.[15] There seem to be at least three major mechanisms whereby most injurious agents exert their effects:

- Free radical formation
- Hypoxia and ATP depletion
- Disruption of intracellular calcium homeostasis (Fig. 5.7)

Free Radical Injury

Many injurious agents exert damaging effects through reactive chemical species known as *free radicals*.[23] Free radicals are highly reactive chemical species with an unpaired electron in the outer orbit (valence shell) of the molecule.[15] In the literature, the unpaired electron is denoted by a dot, for example, •NO. The unpaired electron causes free radicals to be unstable and highly reactive, so that they react nonspecifically with molecules in the vicinity. Moreover, free radicals can establish chain reactions consisting of many events that generate new free radicals. In cells and tissues, free radicals react with proteins, lipids, and carbohydrates, thereby damaging cell membranes; inactivate enzymes; and damage nucleic acids that make up DNA. The actions of free radicals may disrupt and damage cells and tissues.

Reactive oxygen species (ROS) are oxygen-containing molecules that include free radicals such as superoxide (O_2^-) and hydroxyl radical (OH•) and nonradicals such as hydrogen peroxide (H_2O_2).[15] These molecules are produced endogenously by normal metabolic processes or cell activities, such as the metabolic burst that accompanies

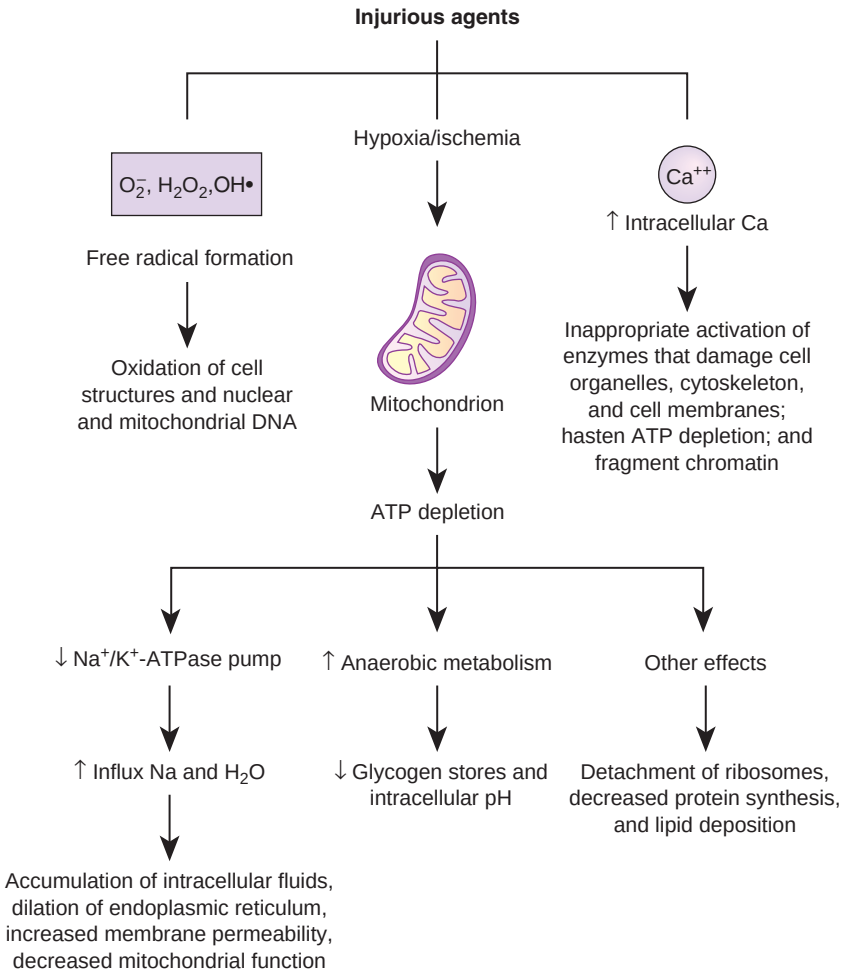


FIGURE 5.7 • Mechanisms of cell injury. The injurious agents tend to cause hypoxia/ischemia (see *middle arrow* that illustrates the manifestations that trigger anaerobic metabolism to develop and cellular injury). Also on the left aspect of the figure, the free radical formation causes oxidation of cell structures leading to decreased ATP, and on the right side, the increased intracellular calcium damages many aspects of the cell that also causes ATP depletion. These three paths illustrate how injurious agents cause cell injury and death.

phagocytosis. However, exogenous causes, including ionizing and UV radiation, can cause ROS production in the body. *Oxidative stress* is a condition that occurs when the generation of ROS exceeds the ability of the body to neutralize and eliminate ROS.[15] Oxidative stress can lead to oxidation of cell components, activation of signal transduction pathways, and changes in gene and protein expression. DNA modification and damage can occur as a result of oxidative stress. Although ROS and oxidative stress are clearly associated with cell and tissue damage, evidence shows that ROS do not always act in a random and damaging manner. Current studies have found that ROS are also important signaling molecules that are used in healthy cells to regulate and maintain normal activities and functions such as vascular tone and insulin and vascular endothelial growth factor signaling.[25] Oxidative damage has been implicated in many diseases. Mutations in the gene for SOD are linked with amyotrophic lateral sclerosis (ALS; so-called Lou Gehrig disease).[26] Oxidative stress is thought to play an important role in the development of cancer.[15] Reestablishment of blood flow after loss of perfusion, as occurs during heart attack and stroke, is associated with oxidative injury to vital organs.[27] The endothelial dysfunction that contributes to the development, progression, and prognosis of cardiovascular disease is thought to be caused in part by oxidative stress.[27] In addition to the many diseases and altered health conditions associated with oxidative damage, oxidative stress has been linked with the age-related functional declines that underlie the process of aging.[28]

Antioxidants are natural and synthetic molecules that inhibit the reactions of ROS with biologic structures or prevent the uncontrolled formation of ROS. Antioxidants include enzymatic and nonenzymatic compounds.[15] Catalase can catalyze the reaction that forms water from hydrogen peroxide. Nonenzymatic antioxidants include carotenes (*e.g.,* vitamin A), tocopherols (*e.g.,* vitamin E), ascorbate (vitamin C), glutathione, flavonoids, selenium, and zinc.[15]

Hypoxic Cell Injury

Hypoxia deprives the cell of oxygen and interrupts oxidative metabolism and the generation of ATP. The actual time necessary to produce irreversible cell damage depends on the degree of oxygen deprivation and the metabolic needs of the cell. Some cells, such as those in the heart, brain, and kidney, require large amounts of oxygen to provide energy to perform their functions. Brain cells, for example, begin to undergo permanent damage after 4 to 6 minutes of oxygen deprivation. A thin margin can exist between the time involved in reversible and irreversible cell damage. During hypoxic conditions,

hypoxia-inducible factors (HIFs) cause the expression of genes that stimulate red blood cell formation, produce ATP in the absence of oxygen, and increase angiogenesis[29] (*i.e.,* the formation of new blood vessels).

Hypoxia can result from an inadequate amount of oxygen in the air, respiratory disease, ischemia (*i.e.,* decreased blood flow due to vasoconstriction or vascular obstruction), anemia, edema, or inability of the cells to use oxygen. Ischemia is characterized by impaired oxygen delivery and impaired removal of metabolic end products such as lactic acid. In contrast to pure hypoxia, which depends on the oxygen content of the blood and affects all cells in the body, ischemia commonly affects blood flow through limited numbers of blood vessels and produces local tissue injury. In some cases of edema, the distance for diffusion of oxygen may become a limiting factor in the delivery of oxygen. In hypermetabolic states, cells may require more oxygen than can be supplied by normal respiratory function and oxygen transport. Hypoxia also serves as the ultimate cause of cell death in other injuries. For example, a physical agent such as cold temperature can cause severe vasoconstriction and impair blood flow.

Hypoxia causes a power failure in the cell, with widespread effects on the cell's structural and functional components. As oxygen tension in the cell falls, oxidative metabolism ceases and the cell reverts to anaerobic metabolism, using its limited glycogen stores in an attempt to maintain vital cell functions. Cellular pH falls as lactic acid accumulates in the cell. This reduction in pH can have adverse effects on intracellular structures and biochemical reactions. Low pH can alter cell membranes and cause chromatin clumping and cell shrinkage.

One important effect of reduced ATP is acute cell swelling caused by failure of the energy-dependent sodium/potassium (Na^+/K^+)–ATPase membrane pump, which extrudes sodium from and returns potassium to the cell. With impaired function of this pump, intracellular potassium levels decrease and sodium and water accumulate in the cell. The movement of water and ions into the cell is associated with multiple changes including widening of the endoplasmic reticulum, membrane permeability, and decreased mitochondrial function.[15] In some instances, the cellular changes due to ischemia are reversible if oxygenation is restored. If the oxygen supply is not restored, however, there is a continued loss of enzymes, proteins, and ribonucleic acid through the hyperpermeable cell membrane. Injury to the lysosomal membranes results in the leakage of destructive lysosomal enzymes into the cytoplasm and enzymatic digestion of cell components. Leakage of intracellular enzymes through the permeable cell membrane into the extracellular fluid provides an important clinical indicator of cell injury and death.

Impaired Calcium Homeostasis

Calcium functions as an important second messenger and cytosolic signal for many cell responses. Various calcium-binding proteins, such as troponin and calmodulin, act as transducers for the cytosolic calcium signal. Calcium/calmodulin–dependent kinases indirectly mediate the effects of calcium on responses such as smooth muscle contraction and glycogen breakdown. Normally, intracellular calcium ion levels are kept extremely low compared with extracellular levels. The low intracellular calcium levels are maintained by membrane-associated calcium/magnesium (Ca^{2+}/Mg^{2+})–ATPase exchange systems. Ischemia and certain toxins lead to an increase in cytosolic calcium because of increased influx across the cell membrane and the release of calcium from intracellular stores. The increased calcium level may inappropriately activate a number of enzymes with potentially damaging effects. These enzymes include the phospholipases, responsible for damaging the cell membrane; proteases that damage the cytoskeleton and membrane proteins; ATPases that break down ATP and hasten its depletion; and endonucleases that fragment chromatin. Although it is known that injured cells accumulate calcium, it is unknown whether this is the ultimate cause of irreversible cell injury.

Reversible Cell Injury and Cell Death

The mechanisms of cell injury can produce sublethal and reversible cellular damage or lead to irreversible injury with

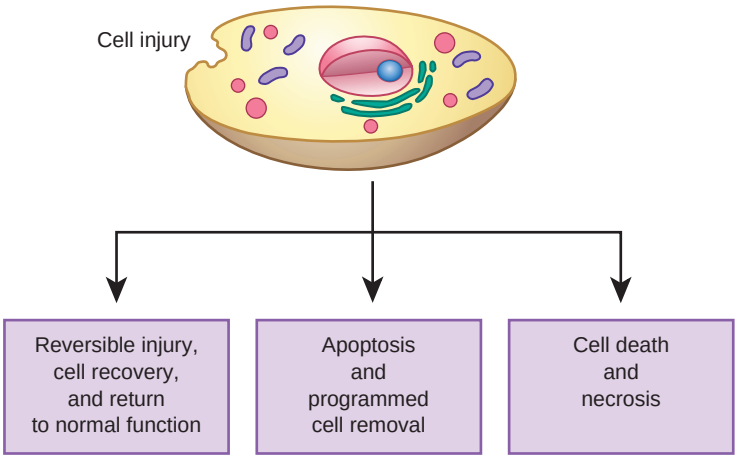


FIGURE 5.8 • Outcomes of cell injury: reversible cell injury, apoptosis and programmed cell removal, and cell death and necrosis.

cell destruction or death (Fig. 5.8). Cell destruction and removal can involve one of two mechanisms:

- Apoptosis, which is designed to remove injured or worn-out cells
- Cell death or necrosis, which occurs in irreversibly damaged cells[1]

Reversible Cell Injury

Reversible cell injury, although impairing cell function, does not result in cell death. Two patterns of reversible cell injury can be observed under the microscope: cellular swelling and fatty change. Cellular swelling occurs with impairment of the energy-dependent Na^+/K^+–ATPase membrane pump, usually as the result of hypoxic cell injury.

Fatty changes are linked to intracellular accumulation of fat. When fatty changes occur, small vacuoles of fat disperse throughout the cytoplasm. The process is usually more ominous than cellular swelling, and although it is reversible, it usually indicates severe injury. These fatty changes may occur because normal cells are presented with an increased fat load or because injured cells are unable to metabolize the fat properly. In obese people, fatty infiltrates often occur within and between the cells of the liver and heart because of an increased fat load. Pathways for fat metabolism may be impaired during cell injury, and fat may accumulate in the cell as production exceeds use and export. The liver, where most fats are synthesized and metabolized, is particularly susceptible to fatty change, but fatty changes may also occur in the kidney, the heart, and other organs.

Programmed Cell Death

In most normal nontumor cells, the number of cells in tissues is regulated by balancing cell proliferation and cell death. Cell death occurs by necrosis or a form of programmed cell death called apoptosis.[1]

Apoptosis is a highly selective process that eliminates injured and aged cells, thereby controlling tissue regeneration. Cells undergoing apoptosis have characteristic morphologic features as well as biochemical changes. As shown in Figure 5.9, shrinking and condensation of the nucleus and cytoplasm occur. The chromatin aggregates at the nuclear envelope, and DNA fragmentation occurs. Then, the cell becomes fragmented into multiple apoptotic bodies in a manner that maintains the integrity of the plasma membrane and does not initiate inflammation. Changes in the plasma membrane induce phagocytosis of the apoptotic bodies by macrophages and other cells, thereby completing the degradation process.

Apoptosis is thought to be responsible for several normal physiologic processes, including the programmed destruction of cells during embryonic development, hormone-dependent involution of tissues, death of immune cells, cell death by cytotoxic T cells, and cell death in proliferating cell populations. During embryogenesis, in the development of a number of organs such as the heart, which begins as a pulsating tube and is gradually modified to become a four-chambered pump, apoptotic cell death allows for the next stage of organ development. It also separates the webbed fingers and toes of the developing

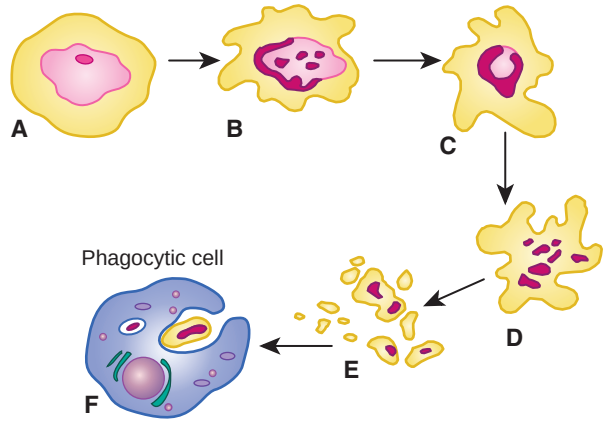


FIGURE 5.9 • Apoptotic cell removal: shrinking of the cell structures (**A**), condensation and fragmentation of the nuclear chromatin (**B** and **C**), separation of nuclear fragments and cytoplasmic organelles into apoptotic bodies (**D** and **E**), and engulfment of apoptotic fragments by phagocytic cell (**F**).

embryo (Fig. 5.10). Apoptotic cell death occurs in the hormone-dependent involution of endometrial cells during the menstrual cycle and in the regression of breast tissue after weaning from breast-feeding. The control of immune cell numbers and destruction of autoreactive T cells in the thymus have been credited to apoptosis. Cytotoxic T cells and natural killer cells are thought to destroy target cells by inducing apoptotic cell death.

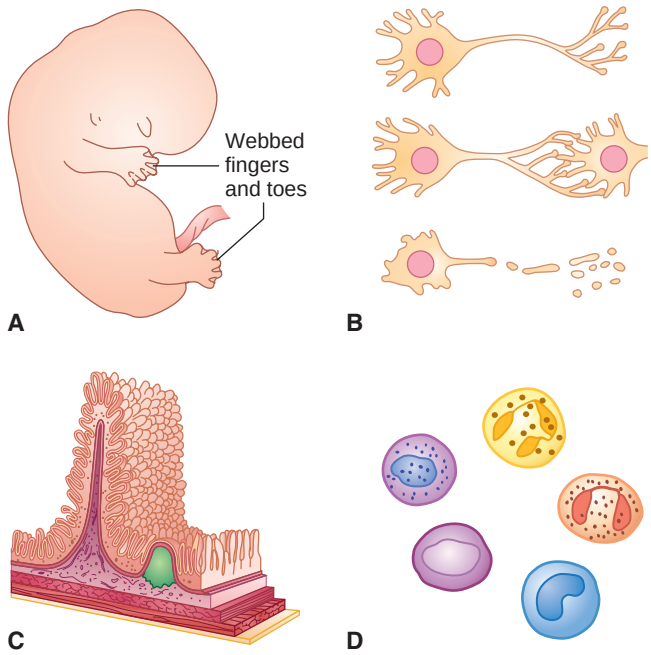


FIGURE 5.10 • Examples of apoptosis: (**A**) separation of webbed fingers and toes in embryo; (**B**) development of neural connections; neurons that do not establish synaptic connections and receive survival factors may be induced to undergo apoptosis; (**C**) removal of cells from intestinal villi; new epithelial cells continuously form in the crypt, migrate to the villus tip as they age, and undergo apoptosis at the tip at the end of their life span; and (**D**) removal of senescent blood cells.

Apoptosis is linked to many pathologic processes and diseases. For example, interference with apoptosis is known to be a mechanism that contributes to carcinogenesis.[30] Apoptosis may also be implicated in neurodegenerative disorders such as Alzheimer disease, Parkinson disease, and ALS. However, the exact mechanisms involved in these diseases remain under investigation.

Two basic pathways for apoptosis have been described (Fig. 5.11). These are the extrinsic pathway, which is death receptor dependent, and the intrinsic pathway, which is death receptor independent. The execution phase of both pathways is carried out by proteolytic enzymes called *caspases*, which are present in the cell as *procaspases* and are activated by cleavage of an inhibitory portion of their polypeptide chain.

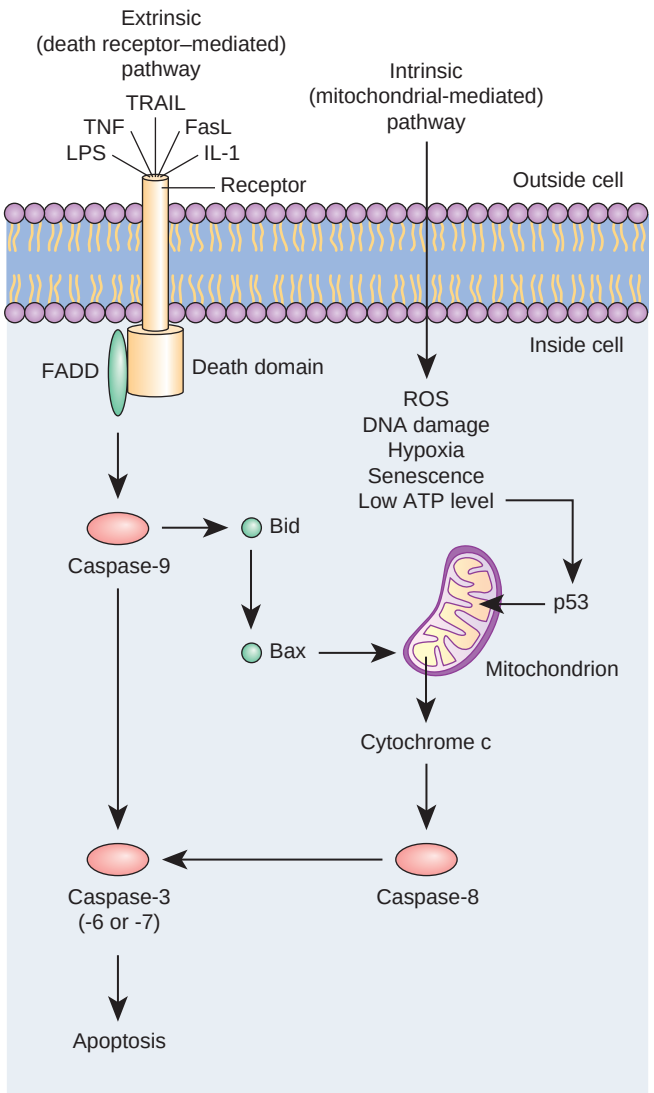


FIGURE 5.11 • Extrinsic and intrinsic pathways of apoptosis. The extrinsic pathway is activated by signals such as Fas ligand (FasL) that, on binding to the Fas receptor, form a death-inducing complex by joining the Fas-associated death domain (FADD) to the death domain of the Fas receptor. The intrinsic pathway is activated by signals, such as reactive oxygen species (ROS) and DNA damage that induce the release of cytochrome c from mitochondria into the cytoplasm. Both pathways activate caspases to execute apoptosis.

The *extrinsic pathway* involves the activation of receptors such as tumor necrosis factor (TNF) receptors and the Fas ligand receptor.[31] Fas ligand may be expressed on the surface of certain cells such as cytotoxic T cells, or appear in a soluble form. When Fas ligand binds to its receptor, proteins congregate at the cytoplasmic end of the Fas receptor to form a death-initiating complex. The complex then converts procaspase-8 to caspase-8. Caspase-8, in turn, activates a cascade of caspases that execute the process of apoptosis.[31] The end result includes activation of endonucleases that cause fragmentation of DNA and cell death. In addition to TNF and Fas ligand, primary signaling molecules known to activate the extrinsic pathway include TNF-related apoptosis-inducing ligand (TRAIL); the cytokine interleukin-1 (IL-1); and lipopolysaccharide (LPS), the endotoxin found in the outer cell membrane of gram-negative bacteria.

The *intrinsic pathway*, or *mitochondrion-induced pathway*, of apoptosis is activated by conditions such as DNA damage, ROS, hypoxia, decreased ATP levels, cellular senescence, and activation of the p53 protein by DNA damage.[32] It involves the opening of mitochondrial membrane permeability pores with release of cytochrome c from the mitochondria into the cytoplasm. Cytoplasmic cytochrome c activates caspases, including caspase-3. Caspase-3 activation is a common step to both the extrinsic and intrinsic pathways. Furthermore, activation or increased levels of proapoptotic proteins, such as Bid and Bax, after caspase-8 activation in the extrinsic pathway can lead to mitochondrial release of cytochrome c, thereby bridging the two pathways for apoptosis. Many inhibitors of apoptosis within cells are known and thought to contribute to cancer and autoimmune diseases.[33] The therapeutic actions of certain drugs may induce or facilitate apoptosis. Apoptosis continues to be an active area of investigation to better understand and treat a variety of diseases.

Necrosis

Necrosis refers to cell death in an organ or tissue that is still part of a living organism.[15] Necrosis differs from apoptosis since it causes loss of cell membrane integrity and enzymatic breakdown of cell parts and triggers the inflammatory process.[1] In contrast to apoptosis, which functions in removing cells so new cells can replace them, necrosis often interferes with cell replacement and tissue regeneration.

With necrotic cell death, there are marked changes in the appearance of the cytoplasmic contents and the nucleus. These changes often are not visible, even under the microscope, for hours after cell death. The dissolution of the necrotic cell or tissue can follow several paths. The cell can undergo liquefaction (*i.e.,* liquefaction necrosis); it can be transformed to a gray, firm mass (*i.e.,* coagulation necrosis); or it can be converted to a cheesy material by infiltration of fatlike substances (*i.e.,* caseous necrosis).[1] *Liquefaction necrosis* occurs when some of the cells die but their catalytic enzymes are not destroyed.[1] An example of liquefaction necrosis is the softening of the center of an abscess with discharge of its contents. During *coagulation necrosis*, acidosis develops and denatures the enzymatic and structural proteins of the cell. This type of necrosis is

characteristic of hypoxic injury and is seen in infarcted areas.[1] Infarction (*i.e.,* tissue death) occurs when an artery supplying an organ or part of the body becomes occluded and no other source of blood supply exists. As a rule, the shape of the infarction is conical and corresponds to the distribution of the artery and its branches. An artery may be occluded by an embolus, a thrombus, disease of the arterial wall, or pressure from outside the vessel.

Caseous necrosis is a distinctive form of coagulation necrosis in which the dead cells persist indefinitely.[1] It is most commonly found in the center of tuberculosis granulomas, or tubercles.[1]

Gangrene. The term *gangrene* is applied when a considerable mass of tissue undergoes necrosis. Gangrene may be classified as dry or moist. In dry gangrene, the part becomes dry and shrinks, the skin wrinkles, and its color changes to dark brown or black. The spread of dry gangrene is slow, and its symptoms are not as marked as those of wet gangrene. The irritation caused by the dead tissue produces a line of inflammatory reaction (*i.e.,* line of demarcation) between the dead tissue of the gangrenous area and the healthy tissue. Dry gangrene usually results from interference with the arterial blood supply to a part without interference with venous return and is a form of coagulation necrosis.

In moist or wet gangrene, the area is cold, swollen, and pulseless. The skin is moist, black, and under tension. Blebs form on the surface, liquefaction occurs, and a foul odor is caused by bacterial action. There is no line of demarcation between the normal and diseased tissues, and the spread of tissue damage is rapid. Systemic symptoms are usually severe, and death may occur unless the condition can be arrested. Moist or wet gangrene primarily results from interference with venous return from the part. Bacterial invasion plays an important role in the development of wet gangrene and is responsible for many of its prominent symptoms. Dry gangrene is confined almost exclusively to the extremities, but moist gangrene may affect the internal organs or the extremities. If bacteria invade the necrotic tissue, dry gangrene may be converted to wet gangrene.

Gas gangrene is a special type of gangrene that results from infection of devitalized tissues by one of several *Clostridium* bacteria, most commonly *Clostridium perfringens*.[1] These anaerobic and spore-forming organisms are widespread in nature, particularly in soil. Gas gangrene is prone to occur in trauma and compound fractures in which dirt and debris are embedded. Some species have been isolated in the stomach, gallbladder, intestine, vagina, and skin of healthy people. Characteristic of this disorder are the bubbles of hydrogen sulfide gas that form in the muscle. Gas gangrene is a serious and potentially fatal disease. Antibiotics are used to treat the infection and surgical methods are used to remove the infected tissue. Amputation may be required to prevent spreading infection involving a limb. Hyperbaric oxygen therapy has been used, but clinical data supporting its efficacy have not been rigorously assessed.

Cellular Aging

Like adaptation and injury, aging is a process that involves the cells and tissues of the body. A number of theories have been proposed to explain the cause of aging. These theories are not mutually exclusive, and aging is most likely complex with multiple causes. The main theories of aging can be categorized based on evolutionary, molecular, cellular, and systems-level explanations.[1]

The *evolutionary theories* focus on genetic variation and reproductive success. After the reproductive years have passed, it is not clear that continued longevity contributes to the fitness of the species. Thus, "antiaging" genes would not necessarily be selected, preserved, and prevalent in the gene pool.

The *molecular theories* of cellular aging focus more on mutations or changes in gene expression. Because the appearance, properties, and function of cells depend on gene expression, this aspect is likely to be involved in aging at some level. Recent attention is being given to the so-called aging genes identified in model systems.

There are a number of *cellular theories of senescence* that are currently under investigation, including those that focus on telomere shortening, free radical injury, and apoptosis. It has been known since the mid-1960s that many cells in culture exhibit a limit in replicative capacity, the so-called Hayflick limit of about 50 population doublings. This limit seems to be related to the length of the telomeres, which are DNA sequences at the ends of chromosomes. Each time a cell divides, the telomeres shorten until a critical minimal length is attained, senescence ensues, and further cell replication does not occur. Some cells have telomerase, an enzyme that "rebuilds" telomeres and lessens or prevents shortening. Cancer cells have high levels of telomerase, which prevents senescence and contributes to the cellular immortality that characterizes cancer. Telomere shortening appears to be related to other theories of cellular causes of aging. For example, free radicals and oxidative damage can kill cells and hasten telomere shortening. Caloric restriction, which appears to increase longevity, may be related to reduced mitochondrial free radical generation owing to reduced methionine or other dietary amino acid intake.[34]

Systems-level theories center on a decline in the integrative functions of organ systems such as the immunologic and neuroendocrine systems, which are necessary for overall control of other body systems. The immune system may decline with age and be less effective in protecting the body from infection or cancer. In addition, mutations and manipulations of genes such as *daf*-2, which is similar to human insulin/IGF-1 receptor genes, in the aging worm model *Caenorhabditis elegans* cause significant changes in longevity.[35] Pathways related to *daf*-2 may be responsible for relationships between caloric restriction and prolonged life span in rodents and other animals. The mechanisms that regulate aging are likely to be complex and multifactorial, as will be any interventions to prolong aging.

Innate and Adaptive Immunity

13

Nancy A. Moriber

THE IMMUNE RESPONSE
- Cytokines and Their Role in Immunity
  - *General Properties of Cytokines*
  - *Chemokines*
  - *Colony-Stimulating Factors*

INNATE IMMUNITY
- Epithelial Barriers
- Cells of Innate Immunity
  - *Neutrophils and Macrophages*
  - *Dendritic Cells*
  - *Natural Killer Cells and Intraepithelial Lymphocytes*
- Pathogen Recognition
  - *Pattern Recognition*
  - *Toll-Like Receptors*
- Soluble Mediators of Innate Immunity
  - *Opsonins*
  - *Inflammatory Cytokines*
  - *Acute-Phase Proteins*
- The Complement System

ADAPTIVE IMMUNITY
- Antigens
- Cells of Adaptive Immunity
  - *Lymphocytes*
  - *Major Histocompatibility Complex Molecules*
  - *Antigen-Presenting Cells*
- B Lymphocytes and Humoral Immunity
  - *Immunoglobulins*
  - *Humoral Immunity*
- T Lymphocytes and Cellular Immunity
  - *Helper T Cells and Cytokines in Adaptive Immunity*
  - *Regulatory T Cells*
  - *Cytotoxic T Cells*
  - *Cell-Mediated Immunity*
- Lymphoid Organs
  - *Thymus*
  - *Lymph Nodes*
  - *Spleen*
  - *Other Secondary Lymphoid Tissues*
- Active versus Passive Immunity
- Regulation of the Adaptive Immune Response

DEVELOPMENTAL ASPECTS OF THE IMMUNE SYSTEM
- Transfer of Immunity from Mother to Infant
- Immune Response in the Older Adult

The human body is constantly exposed to potentially deleterious microorganisms and foreign substances. Therefore, it has evolved a complete system composed of complementary and interrelated mechanisms to defend against invasion by bacteria, viruses, and other foreign substances. Through recognition of molecular patterns, the body's immune system can distinguish itself from these foreign substances and can discriminate potentially harmful from nonharmful agents. In addition, it can defend against abnormal cells and molecules that periodically develop. The skin and its epithelial layers in conjunction with the body's normal inflammatory processes make up the first line of the body's defense and confer *innate or natural immunity* to the host. Once these protective barriers have been crossed, the body relies upon a second line of defense known as the *adaptive immune response* to eradicate infection by invading organisms. The adaptive immune response develops slowly over time but results in the development of antibodies capable of targeting specific microorganisms and foreign substances should a second exposure occur.

This chapter covers immunity and the immune system, including a complete discussion of innate and adaptive immunity. Concepts related to key cellular function, recognition systems, and effector responses integral to the immune system are also presented. In addition, developmental aspects of the immune system are discussed.

THE IMMUNE RESPONSE

After completing this section of the chapter, you should be able to meet the following objectives:

- Discuss the function of the immune system.
- Contrast and compare the general properties of innate and adaptive immunity.
- Characterize the chemical mediators that orchestrate the immune response.

TABLE 13.1 FEATURES OF INNATE AND ADAPTIVE IMMUNITY

FEATURE	INNATE	ADAPTIVE
Time of response	Immediate (minutes/hours)	Dependent upon exposure (first: delayed, second: immediate d/t production antibodies)
Diversity	Limited to classes or groups of microbes	Very large; specific for each unique antigen
Microbe recognition	General patterns on microbes; nonspecific	Specific to individual microbes and antigens (antigen/antibody complexes)
Nonself recognition	Yes	Yes
Response to repeated infection	Similar with each exposure	Immunologic memory; more rapid and efficient with subsequent exposure
Defense	Epithelium (skin, mucous membranes), phagocytes, inflammation, fever	Cell killing; tagging of antigen by antibody for removal
Cellular components	Phagocytes (monocytes/macrophages, neutrophils), NK cells, DCs	T and B lymphocytes, macrophages, DCs, NK cells
Molecular components	Cytokines, complement proteins, acute-phase proteins, soluble mediators	Antibodies, cytokines, complement system

Immunity can be defined as the body's ability to defend against specific pathogens and/or foreign substances in the initiation of disease processes. The multidimensional response initiated by the body's various defense systems is known as the immune response. Some of these responses become active almost immediately, while others develop slowly over time. It is the coordinated interaction of these mechanisms that allows the body to maintain normal internal homeostasis. However, when these mechanisms are either depressed or overactive, they become responsible for many of the pathophysiologic processes encountered in health care.

Innate immunity and adaptive immunity are complementary processes that work to protect the body. *Innate immunity*, the body's first line of defense, occurs early and more rapidly in response to foreign substances, while adaptive immunity is usually delayed unless the host has been exposed before (Table 13.1).

Intact innate immune mechanisms are essential for the initiation of the adaptive immune response and, therefore, a successful immune response dependent upon cooperation between the two systems. Dendritic cells are an essential component of both innate and adaptive immunity and serve as the link between the two immune responses through the release of dendritic cell–derived substances, such as cytokines and chemokines.[1] As a result, innate immune cells are capable of communicating important information regarding key characteristics of the invading microorganism or foreign substance to the B and T lymphocytes involved in adaptive immunity. The adaptive immune response is also capable of increasing its efficiency by recruitment and activation of additional phagocytes and molecules of the innate immune system. Each system is therefore essential for an effective immune response and works in concert in the fight against infection.

Cytokines and Their Role in Immunity

The ability of the cells of both the innate and adaptive immune systems to communicate critical information with each other by cell-to-cell contact and initiate end effector responses is dependent upon the secretion of short-acting, biologically active, soluble molecules called *cytokines*. Cytokines are an essential component of host defense mechanisms and the primary means with which cells of innate and adaptive immunity interact. Chemokines are a subset of cytokines that consist of small protein molecules involved in both immune and inflammatory responses.[2] They are responsible for directing leukocyte migration to areas of injury and to locations where primary immune responses are initiated such as lymph nodes, the spleen, Peyer patches, and the tonsils.[2] The source and function of the main cytokines that participate in innate and adaptive immunity are summarized in Table 13.2.

General Properties of Cytokines

Cytokines are low molecular weight, regulatory, pro- or anti-inflammatory proteins that are produced by cells of the innate and adaptive immune systems and that mediate many of the actions of these cells. The majority of the functionally important cytokines are interleukins (ILs), interferons (IFNs), and tumor necrosis factor alpha (TNF-α). Cytokines generate their responses by binding to specific receptors on their target cells and activating G-protein–coupled receptors.[2,3]

Interleukins (ILs) are produced by macrophages and lymphocytes in response to the presence of an invading microorganism or activation of the inflammatory process. Their primary function is to enhance the acquired immune response through alteration of molecular expression, induction of leukocyte maturation, enhanced leukocyte chemotaxis, and general suppression or enhancement of the inflammatory process.

TABLE 13.2 CYTOKINES OF INNATE AND ADAPTIVE IMMUNITY

CYTOKINES	SOURCE	FUNCTION
Interleukin-1 (IL-1)	Macrophages, endothelial cells, some epithelial cells	Wide variety of biologic effects; activates endothelium in inflammation; induces fever and acute-phase response; stimulates neutrophil production
Interleukin-2 (IL-2)	$CD4^+$, $CD8^+$ T cells	Growth factor for activated T cells; induces synthesis of other cytokines; activates cytotoxic T lymphocytes and NK cells
Interleukin-3 (IL-3)	$CD4^+$ T cells	Growth factor for progenitor hematopoietic cells
Interleukin-4 (IL-4)	$CD4^+$ T_2H cells, mast cells	Promotes growth and survival of T, B, and mast cells; causes T_2H cell differentiation; activates B cells and eosinophils; and induces IgE-type responses
Interleukin-5 (IL-5)	$CD4^+$ T_2H cells	Induces eosinophil growth and development
Interleukin-6 (IL-6)	Macrophages, endothelial cells, T lymphocytes	Stimulates the liver to produce mediators of acute-phase inflammatory response; also induces proliferation of antibody-producing cells by the adaptive immune system
Interleukin-7 (IL-7)	Bone marrow stromal cells	Primary function in adaptive immunity; stimulates pre-B cells and thymocyte development and proliferation
Interleukin-8 (IL-8)	Macrophages, endothelial cells	Primary function in adaptive immunity; chemoattracts neutrophils and T lymphocytes; regulates lymphocyte homing and neutrophil infiltration
Interleukin-10 (IL-10)	Macrophages, some T-helper cells	Inhibitor of activated macrophages and DCs; decreases inflammation by inhibiting T_1H cells and release of IL-12 from macrophages
Interleukin-12 (IL-12)	Macrophages, DCs	Enhances NK cell cytotoxicity in innate immunity; induces T_1H cell differentiation in adaptive immunity
Type I interferons (IFN-α, IFN-β)	Macrophages, fibroblasts	Inhibit viral replication; activate NK cells; and increase expression of MHC-I molecules on virus-infected cells
Interferon-γ (IFN-γ)	NK cells, $CD4^+$ and $CD8^+$ T lymphocytes	Activates macrophages in both innate immune responses and adaptive cell-mediated immune responses; increases expression of MHC-I and MHC-II and antigen processing and presentation
Tumor necrosis factor-α (TNF-α)	Macrophages, T cells	Induces inflammation, fever, and acute-phase response; activates neutrophils and endothelial cells; kills cells through apoptosis
Chemokines	Macrophages, endothelial cells, T lymphocytes	Large family of structurally similar cytokines that stimulate leukocyte movement and regulate the migration of leukocytes from the blood to the tissues
Granulocyte–monocyte CSF (GM-CSF)	T cells, macrophages, endothelial cells, fibroblasts	Promotes neutrophil, eosinophil, and monocyte maturation and growth; activates mature granulocytes
Granulocyte CSF (G-CSF)	Macrophages, fibroblasts, endothelial cells	Promotes growth and maturation of neutrophils consumed in inflammatory reactions
Monocyte CSF (M-CSF)	Macrophages, activated T cells, endothelial cells	Promotes growth and maturation of mononuclear phagocytes

CSF, colony-stimulating factor; NK, natural killer; T_1H, T-helper type 1; T_2H, T-helper type 2; MHC, major histocompatibility complex.

IFNs are cytokines that primarily protect the host against viral infections and play a role in the modulation of the inflammatory response. IFNs are cell-type specific with IFN-α and IFN-β produced primarily by macrophages and IFN-γ produced primarily by T lymphocytes. TNF-α, a cytokine in a class by itself, is one of the most important mediators of the inflammatory response and is produced by macrophages when surface toll-like receptors (TLRs) recognize pathogen-associated molecular patterns (PAMPs) on the surface of microorganisms.[4] TNF-α acts as an endogenous pyogen (fever producer) and induces synthesis of proinflammatory substances in the liver. With prolonged exposure, it has the ability to cause intravascular coagulation and subsequent thrombosis production.

Despite the diverse functions of the cytokines, they all share certain important properties. All cytokines are secreted in a brief, self-limited manner. They are rarely stored as pre-

formed molecules but rather are synthesized through transcription as a result of cellular activation. The actions of cytokines are often pleiotropic, meaning that they have the ability to allow a single cytokine to act on different cell types. For example, IL-17 is produced by the T-helper 17 ($T_{17}H$) cells and acts on several cell types including leukocytes, epithelial cells, mesothelial cells, vascular endothelial cells, and fibroblasts. As a result, $T_{17}H$ cells play a critical role in host defense against pathogens that infiltrate the mucosal barrier.[5] Although pleiotropism allows cytokines to mediate diverse effects, it greatly limits their use for therapeutic purposes because of numerous unwanted side effects. Redundancy refers to the ability of different cytokines to stimulate the same or overlapping biologic functions. Because of this redundancy, antagonists against a single cytokine may not have functional consequences because other cytokines may compensate.

In addition to being pleiotropic and redundant, cytokines can have broad activity. Therefore, several different cell types are capable of producing a single cytokine. For example, IL-1 is a proinflammatory cytokine that is primarily produced by macrophages but can be produced by virtually all leukocytes, endothelial cells, and fibroblasts. Cytokines also function to initiate cascade functions with one cytokine influencing the synthesis and actions of other cytokines. Often the second and third cytokines may mediate the biologic effects of the first cytokine. These effects may be localized, acting on a single cell or group of cells in the area surrounding the effector cell, or the effects can be systemic with the cytokines secreted into the bloodstream and transported to their site of action. TNF-α is an example of a cytokine with wide-reaching systemic effects. Cytokines may also serve as antagonists to inhibit the action of another cytokine and as a result act as anti-inflammatory cytokines. IL-110 is an anti-inflammatory cytokine to down-regulate the inflammatory and adaptive immune responses.

Chemokines

Chemokines are small protein molecules (70 to 130 amino acids) that are involved in immune and inflammatory cellular responses and function to control the migration of leukocytes to their primary site of action in the immune response.[6] There are four distinct classes of chemokines (C, CC, CXC, and CX3C), which are named for the number and location of cysteine residues on the terminal amino acid of the protein.[2] Currently, 48 distinct chemokine molecules have been identified within the four different classes. The vast majority of these are classified as either CC or CXC chemokines. The CC chemokines have the first two cysteine molecules adjacent to each other, while these molecules are separated by an amino acid in the CXC chemokines. The CC chemokines attract monocytes, lymphocytes, and eosinophils to sites of chronic inflammation. The CXC chemokines attract neutrophils to sites of acute inflammation.

Chemokines are named according to structure, followed by "L" and the number of their gene (*e.g.,* CCL1, CXCL1). Likewise, chemokine receptors are named according to the structure, followed by an "R" and a number (*e.g.,* CCR1, CXCR1). Six receptors for CXC (CXCRs) and 10 for CC (CCRs) chemokines have been characterized in terms of their structure and function.[2] Chemokines communicate with their target cells by activating G-protein–coupled receptors that are pertussis toxin sensitive and as a result are capable of activating different populations of leukocytes, thereby controlling the migration of immune cells to their sites of action based upon the needs of the situation.[2] Most receptors recognize more than one chemokine, and most chemokines recognize more than one receptor. Binding of a chemokine to a receptor can result in inhibition or activation with the same chemokine acting as an activator at one type of receptor and as an inhibitor at another. Chemokines are implicated in a number of acute and chronic diseases, including atherosclerosis, rheumatoid arthritis, inflammatory bowel disease (Crohn disease, ulcerative colitis), allergic asthma and chronic bronchitis, multiple sclerosis, systemic lupus erythematosus, and HIV infection. They also play a role in the body's immune response against cancer cells through the up-regulation of CCL21 and other chemokines by activated T cells and other tumor-derived proteins.[7,8]

Colony-Stimulating Factors

Colony-stimulating factors (CSFs) encompass a subset of cytokines that participate in hematopoiesis by stimulating bone marrow pluripotent stem and progenitor or precursor cells to produce large numbers of mature platelets, erythrocytes, lymphocytes, neutrophils, monocytes, eosinophils, basophils, and dendritic cells (DCs). The CSFs were named according to the type of target cell on which they act (see Table 13.2). Macrophages, endothelial cells, and fibroblasts produce granulocyte colony-stimulating factor (G-CSF) during times of stress and inflammation where it promotes growth and maturation of neutrophils. Granulocyte/monocyte colony-stimulating factor (GM-CSF) acts on the granulocyte–monocyte progenitor cells to produce monocytes, neutrophils, and DCs, and monocyte colony-stimulating factor (M-CSF) stimulates the mononuclear phagocyte progenitor. While CSF is necessary for normal blood cell production and maturation, excess CSF production has been implicated in several disease processes and the development of corticosteroid-resistant chronic obstructive pulmonary disease (COPD).[9] Impaired macrophage function and subsequent impairment of G-CSF activity have been associated with the development of neutrophilia in animal studies.[10] In clinical practice, recombinant CSF is being used to increase the success rates of bone marrow transplantations. The availability of recombinant CSFs and cytokines offers the possibility of several clinical therapies where stimulation or inhibition of the immune response or cell production is desirable.

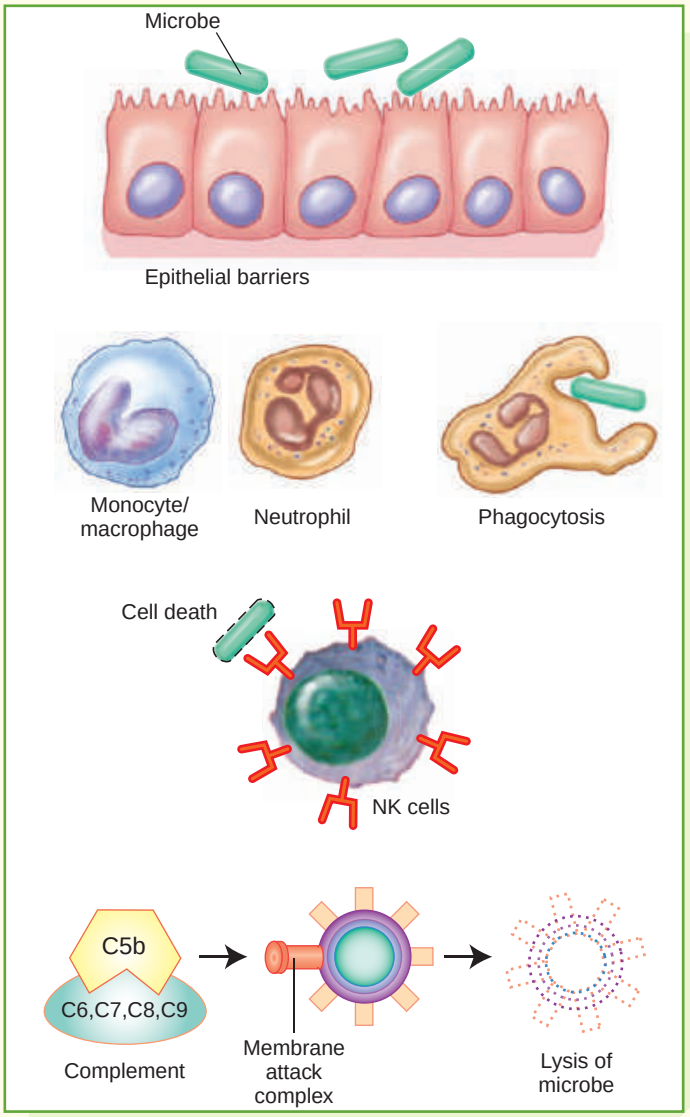
Understanding Innate and Adaptive Immunity

The body's defense against microbes is mediated by two types of immunity: (1) innate immunity and (2) adaptive immunity. Both types of immunity are members of an integrated system in which numerous cells and molecules function cooperatively to protect the body against foreign invaders. The innate immune system stimulates adaptive immunity and influences the nature of the adaptive immune responses to make them more effective. Although they use different mechanisms of pathogen recognition, both types of immunity use many of the same effector mechanisms, including destruction of the pathogen by phagocytosis and the complement system.

Innate Immunity

Innate immunity (also called *natural immunity*) consists of the cellular and biochemical defenses that are in place before an encounter with an infectious agent and provide rapid protection against infection. The major effector components of innate immunity include epithelial cells, which block the entry of infectious agents and secrete antimicrobial enzymes, proteins, and peptides; phagocytic neutrophils and macrophages, which engulf and digest microbes; natural killer (NK) cells, which kill intracellular microbes and foreign agents; and the complement system, which amplifies the inflammatory response and uses the membrane attack response to lyse microbes. The cells of the innate immune system also produce chemical messengers that stimulate and influence the adaptive immune response.

The innate immune system uses pattern recognition receptors that recognize microbial structures (*e.g.,* sugars, lipid molecules, proteins) that are shared by microbes and are often necessary for their survival, but are not present on human cells. Thus, the innate immune system is able to distinguish between self and nonself, but is unable to distinguish between agents.

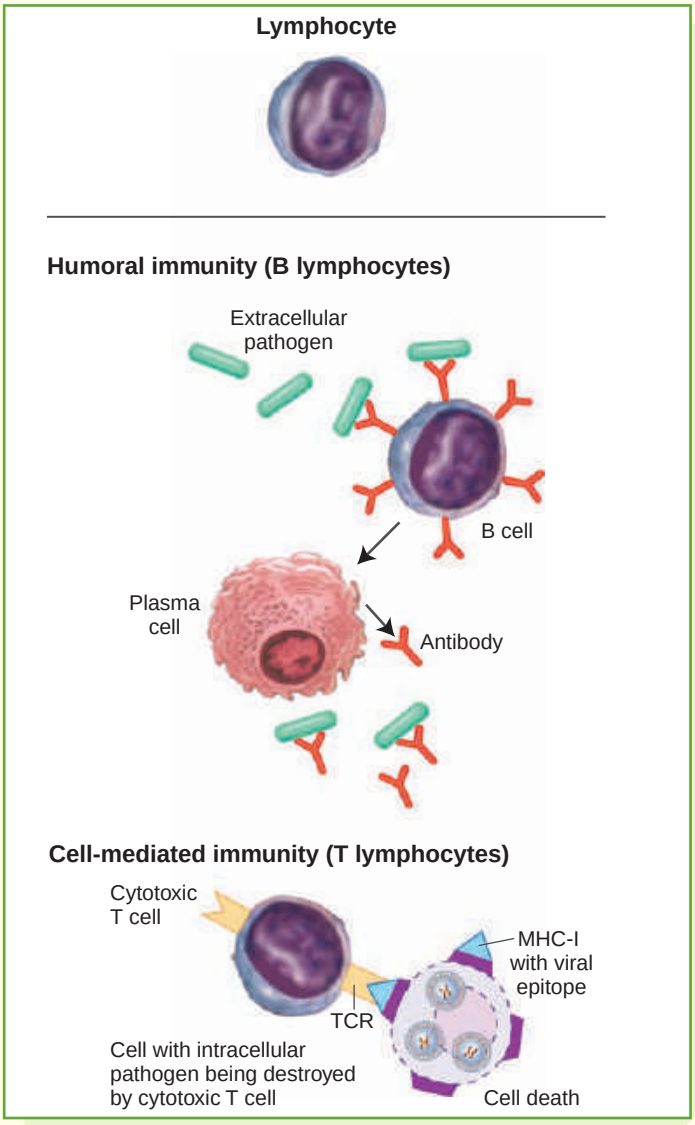


Adaptive Immunity

Adaptive immunity (also called *acquired immunity*) refers to immunity that is acquired through previous exposure to infectious and other foreign agents. A defining characteristic of adaptive immunity is the ability not only to distinguish self from nonself but to recognize and destroy specific foreign agents based on their distinct antigenic properties. The components of the adaptive immune system are the T and B lymphocytes and their products. There are two types of adaptive immune responses, humoral and cell-mediated immunity, that function to eliminate different types of microbes.

Humoral immunity is mediated by the B lymphocytes (B cells) and is the principal defense against extracellular microbes and their toxins. The B cells differentiate into antibody-secreting plasma cells. The circulating antibodies then interact with and destroy the microbes that are present in the blood or mucosal surfaces.

Cell-mediated, or cellular, immunity is mediated by the cytotoxic T lymphocytes (T cells) and functions in the elimination of intracellular pathogens (*e.g.,* viruses). T cells develop receptors that recognize the viral peptides displayed on the surface of infected cells and then signal destruction of the infected cells.



IN SUMMARY

Immunity is the body's defense against disease and invading microorganisms. Immune mechanisms can be divided into two types: innate and adaptive immunity. Innate immunity is the first line of defense and can distinguish between self and nonself through the recognition of cellular patterns on foreign substances and microbes. Adaptive immunity is part of the second line of defense and involves both humoral and cellular mechanisms that respond to cell-specific substances known as antigens. The adaptive immune response is capable of amplifying and sustaining its responses, of distinguishing self from nonself, and finally of memory in that it can recognize the antigen on repeat exposure in order to quickly produce a heightened response on subsequent encounters with the same microorganism. The innate and adaptive immune responses work in concert with one another to ensure that the homeostasis is maintained.

Although cells of both the innate and adaptive immune systems communicate critical information about the invading microbe or pathogen by cell-to-cell contact, many interactions and cellular responses depend on the secretion of chemical mediators in the form of cytokines, chemokines, and CSFs. Cytokines are soluble proteins secreted by cells of both the innate and adaptive immune systems that mediate many of the functions of these cells. Chemokines are cytokines that stimulate the migration and activation of various immune and inflammatory cells. CSFs stimulate the growth and differentiation of bone marrow progenitors of immune cells and play a key role in hematopoiesis.

INNATE IMMUNITY

After completing this section of the chapter, you should be able to meet the following objectives:

- Understand the recognition systems for pathogens in innate immunity.
- Describe the functions of the various cytokines involved in innate immunity.
- Define the role of the complement system in immunity and inflammation.

The innate immune system is comprised of two separate but interrelated lines of defense: the epithelial layer, which acts as a physical barrier to invading substances and organisms, and the inflammatory response. The innate immune response utilizes the body's natural epithelial barriers along with phagocytic cells (mainly neutrophils and macrophages), natural killer (NK) cells, and several plasma proteins, including kinins, clotting factors, and those of the complement system, to maintain internal homeostasis. The innate immune response relies on the body's ability to distinguish evolutionarily conserved structures on pathogens known as PAMPs from structures on human cells.[3] The response of the innate immune system is rapid, usually within minutes to hours, and prevents the establishment of infection and deeper tissue penetration of microorganisms. The innate immune response is usually very effective against most pathogens. However, when the innate response is overwhelmed, adaptive immune responses become activated as the final line of defense against invading organisms. Innate immune mechanisms are always present in the body before an encounter with an infectious agent and are rapidly activated by microorganisms and foreign substances. Therefore, the body's defenses are in full swing before the development of the adaptive immune response. The innate immune system also interacts with and directs adaptive immune responses.

Under normal conditions, the innate immune response is essential to the continued health and well-being of the body. However, during times of hyperresponsiveness or hyporesponsiveness, the innate immune system plays a role in the pathogenesis of disease. One of the main functions of the innate immune system is the initiation of the inflammatory response, which involves the activation of a complex cascade of events and chemical mediators. As part of the innate immune response, inflammation plays a key role in the pathogenesis of many common pathophysiologic states including atherosclerosis and coronary artery disease, bronchial asthma, non–insulin-dependent diabetes mellitus (NIDDM), rheumatoid arthritis, multiple sclerosis, and systemic lupus erythematosus.

KEY POINTS

INNATE IMMUNITY

- Innate immunity consists of physical, chemical, cellular, and molecular defenses that are ready for activation and mediate rapid, initial protection against infection.
- The effector responses of innate immunity involve the inflammatory process and phagocytosis by cells that express pattern recognition receptors (PRRs) that bind with broad patterns shared by groups of microbes but not present on mammalian cells. Toll-like receptors, a major type of PRR, are expressed on phagocytes and are potent activators of innate immune system cells and molecules.

Epithelial Barriers

Physical, mechanical, and biochemical barriers against microbial invasion are found in all common portals of entry into the body, including the skin and respiratory, gastrointestinal, and urogenital tracts. The intact skin is by far the most formidable physical barrier available to infection because of its design. It is comprised of closely packed cells that are organized in multiple layers that are continuously shed. In addition, a protective layer of protein, known as keratin, covers the skin. The skin has simple chemicals that create a nonspecific, salty, acidic environment and antibacterial proteins, such as the enzyme lysozyme, that inhibit the colonization of microorganisms and aid in their destruction. The complexity of the skin becomes evident in cases of contact dermatitis where increased susceptibility to cutaneous infection occurs as the result of abnormalities of the innate immune response including defects in the epithelial layer itself and defects in both signaling and or expression of innate responses.[11]

Sheets of tightly packed epithelial cells line and protect the gastrointestinal, respiratory, and urogenital tracts and physically prevent microorganisms from entering the body. These cells destroy the invading organisms by secreting antimicrobial enzymes, proteins, and peptides. Specialized cells in these linings, such as the goblet cells in the gastrointestinal tract, secrete a viscous material comprised of high molecular weight glycoproteins known as mucins, which when hydrated form *mucus*. The mucins bind to pathogens, thereby trapping them and washing away potential invaders. In the lower respiratory tract, hairlike, mobile structures called **cilia** protrude through the epithelial cells and move microbes trapped in the mucus up the tracheobronchial tree and toward the throat. The physiologic responses of coughing and sneezing further aid in their removal from the body.

Microorganisms that are trapped by mucus are then subjected to various chemical defenses present throughout the body. Lysozyme is a hydrolytic enzyme found in tears,

saliva, and human milk, which is capable of cleaving the walls of bacterial cells by hydrolyzing the 1,4 beta-linkages between residues in peptidoglycan. The complement system is found in the blood and is essential for the activity of antibodies. It is comprised of 20 different proteins, many of which act as precursors of enzymes. An antigen–antibody complex initiates this system. Activation of the complement system increases bacteria aggregation, which renders them more susceptible to phagocytosis through activation of mast cells and basophils and through the direct release of lytic complexes that rupture cell membranes of invading organisms (Fig. 13.1). In addition, recent research has shown that complement plays a key role in bridging the innate–adaptive immune responses through the release of C3 and C5 from DCs.[12] In the stomach and intestines, death of microbes results from the action of digestive enzymes, acidic conditions, and secretions of *defensins*, small cationic peptides that kill within minutes both gram-positive and gram-negative microorganisms by disrupting the microbial membrane.

When pathogens overcome the epithelial defenses, the innate immune response is initiated by the body's leukocytes by the recognition of common surface receptors present on the invading microorganisms.

Cells of Innate Immunity

The cells of the innate immune response are capable of recognizing microbes that share common surface receptor characteristics and in response initiate a broad spectrum of responses that target the invading microorganisms. The key cells of innate immunity include neutrophils, macrophages, DCs, NK cells, and intraepithelial lymphocytes.

Neutrophils and Macrophages

The leukocytes involved in the innate immune response are derived from myeloid stem cells and subdivided into two distinct groups based upon the presence or absence of specific staining granules in their cytoplasm. Leukocytes that contain granules are classified as granulocytes and include neutrophils, eosinophils, and basophils. Cells that lack granules are classified as agranulocytes and include lymphocytes, monocytes, and macrophages.

Neutrophils, which are named for their neutral-staining granules, are the most abundant granulocytes found in the body and make up approximately 55% of all white blood cells. They are also known as polymorphonuclear neutrophils (PMNs). They are phagocytic cells and are capable of ameboid-like movement. They function as early responder cells in innate immunity. They are rare in the tissues and in body cavities and lay predominantly dormant in the blood and bone marrow until they are needed in the immune response.[13] Eosinophils have large coarse granules and normally comprise only 1% to 4% of the total white cell count. In contrast to neutrophils, these cells do not ingest cellular debris but rather antigen–antibody complexes and viruses. They frequently become active in parasitic infections and allergic responses. Basophils make up less than 1% of the total white cell count and contain granules that release a multitude of substances including histamine and proteolytic enzymes. There function is not completely understood, but they are believed to play a role in allergy and parasitic infection as well.

The agranulocytes involved in innate immunity are part of the mononuclear phagocyte system (MPS) and include the monocytes and macrophages. Monocytes are the largest in size of all the white blood cells but make up only 3% to 7% of the total leukocyte count. They are released from the bone

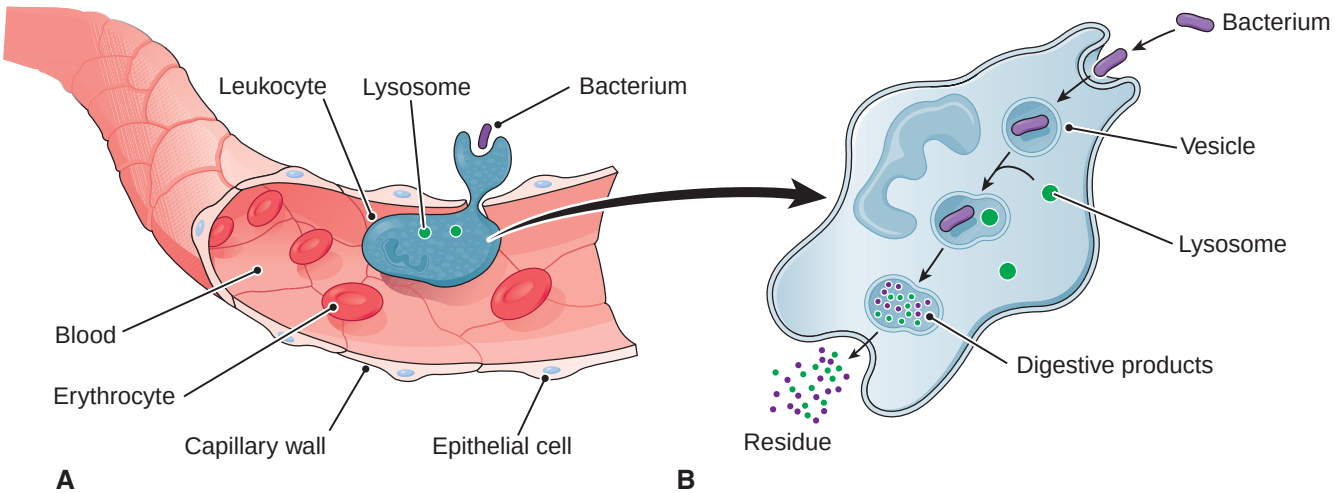


FIGURE 13.1 • Phagocytosis. (**A**) A phagocytic white blood cell moves through a capillary that is in an infected area and engulfs the bacteria. (**B**) The lysosome digests the bacteria that was in a vesicle. (From Cohen B. J. (2013). *Memmler's the human body in health and disease* (12th ed.) Philadelphia, PA: Lippincott Williams & Wilkins.)

marrow into the bloodstream where they migrate into tissues and mature into macrophages and dendritic cells where they participate in the inflammatory response and phagocytize foreign substances and cellular debris. Macrophages have a long life span, reside in the tissues, and act as the first phagocyte that invading organisms encounter upon entering the host.[13] Neutrophils and macrophages work in concert with each other and are crucial to the host's defense against all intracellular and extracellular pathogens.[13]

Macrophages are essential for the clearance of bacteria that breach the epithelial barrier in the intestine and other organ systems.[14] They also have remarkable plasticity that allows them to efficiently respond to environmental signals and change their functional characteristics.[14] This makes them more efficient phagocytic cells than the more abundant neutrophils. Once activated, these cells engulf and digest microbes that attach to their cell membrane. The ability of these phagocytic cells to initiate this response is dependent upon the recognition of pathogenic surface structures known as PAMPs or PRRs of which the TLRs have been the most extensively studied.[3] Phagocytosis of invading microorganisms helps to limit the spread of infection until adaptive immune responses can become fully activated.

In addition to phagocytosis, macrophages and dendritic cells process and present antigens in the initiation of the immune response acting as a major initiator of the adaptive immune response.[1] These cells secrete substances that initiate and coordinate the inflammatory response or activate lymphocytes. Macrophages can also remove antigen–antibody aggregates or, under the influence of T cells, they can destroy malignant host or virus-infected cells.

Dendritic Cells

Dendritic cells (DCs) are specialized, bone marrow–derived leukocytes found in lymphoid tissue and are the bridge between the innate and adaptive immune systems. DCs take their name from the dendrites within the central nervous system because they have surface projections that give them a similar appearance. DCs are relatively rare cells that are found mainly in tissues exposed to external environments such as the respiratory and gastrointestinal systems.[1] They are present primarily in an immature form that is available to directly sense pathogens, capture foreign agents, and transport them to secondary lymphoid tissues.[15] Once activated DCs undergo a complex maturation process in order to function as key antigen-presenting cells (APCs) capable of initiating adaptive immunity.[15] They are responsible for the processing and presentation of foreign antigens to the lymphocytes. DCs, like macrophages, also release several communication molecules that direct the nature of adaptive immune responses.

Natural Killer Cells and Intraepithelial Lymphocytes

NK cells and intraepithelial cells (IELs) are other cell types involved in the innate immune response. NK cells are so named because of their ability to spontaneously kill target organisms. Both types of cells rely on the recognition of specific PAMPs associated with the microorganism cell type.

NK cells are a heterogeneous population of innate lymphocytes that mediate spontaneous cytotoxicity against infected cells.[16] They resemble large granular lymphocytes and are capable of killing some types of tumor and/or infected cells without previous exposure to surface antigens. NK cells were given their name because of their ability to mediate spontaneous cytotoxicity during both innate immune responses. However, they have been shown to play an equally important role in limiting the spread of infection and assisting in the development of adaptive immune responses through the production of cytokines.[16] NK cells assist in dendritic cell maturation and innate immune control of viral infections. These cells are capable of directly killing host cell infected with intracellular (viral) or bacterial pathogenic organisms. They comprise approximately 10% to 15% of peripheral blood lymphocytes but do not bear T-cell receptors (TCR) or cell surface immunoglobulins (Igs). Two-cell surface molecules have been identified, CD16 and CD56, which are widely used to identify NK cell activity. CD16 serves as a receptor for the IgG molecule, which provides NK cells with the ability to lyse IgG-coated target cells.

NK cells can be divided into two main subsets based upon their ability to excrete proinflammatory cytokines. In addition, they differ in their expression of inhibitory versus activating receptors. Cells that express activating receptors (*i.e.,* NKG2D) are induced in response to pathogen-infected or stressed cells, whereas the inhibitory receptors on NK cells recognize patterns (major histocompatibility complex [MHC]-I, lectins) on normal host cells and function to inhibit the action of the NK cells.[16] This assures that only "foreign" cells are destroyed (see Fig. 13.2). In addition to their role as phagocytes, NK cells assist in T-cell polarization, DC maturation, and innate immune control of viral infection through the secretion of immune modulators and antiviral cytokines. [16] Current research is investigating the utilization of these properties of NK cells for the development of vaccines that can modulate and direct the immune response through enhanced cytokine activity.

Pathogen Recognition

The innate immune response plays a crucial role in the proinflammatory response to infection and relies upon the ability of host defenses to differentiate self from nonself so that only invading organisms are targeted. The leukocytes involved in this response recognize certain evolutionarily retained patterns present on the surface of pathogens and in response bind to the membrane and destroy the invading organism through the process of phagocytosis (Fig. 13.3).

Pattern Recognition

Invading pathogens contain conserved structures in their cell membranes termed *pathogen-associated molecular patterns (PAMPs*), which are recognized by the cells of the innate immune system because they possess a limited number of

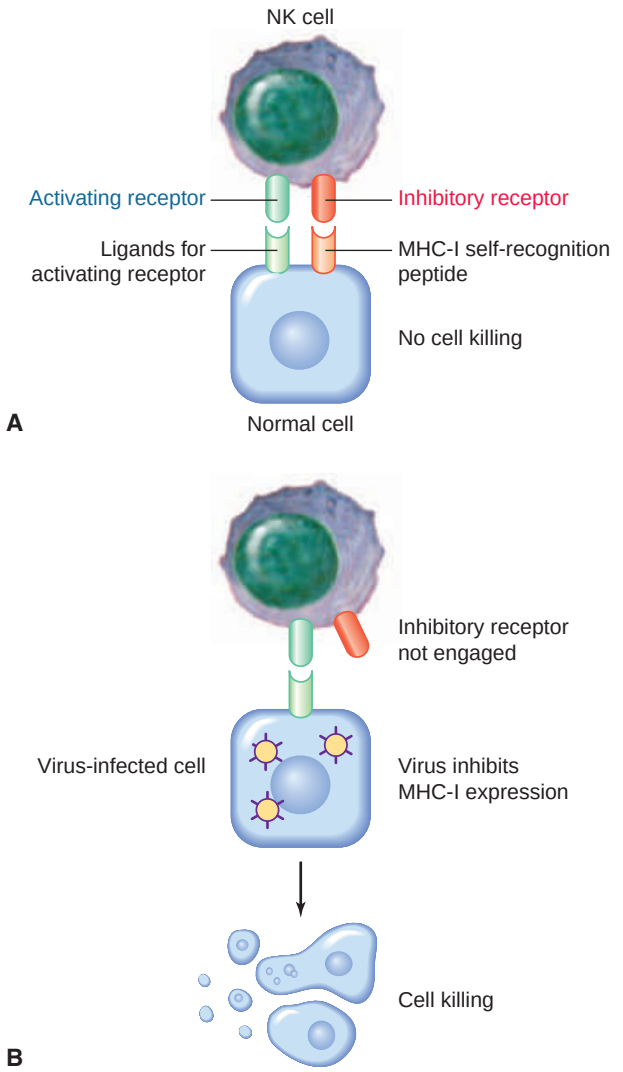


FIGURE 13.2 • Natural killer (NK) cell receptors. (**A**) NK cells express activating receptors that respond to ligands from virus-infected or injured cells and inhibiting receptors that bind to the class I major histocompatibility complex (MHC-I) self-recognition molecules expressed by normal cells. Normal cells are not killed because inhibitory signals from normal MHC-I molecules override activating signals. (**B**) In virus-infected or tumor cells, increased expression of ligands for activating receptors and reduced expression or alteration of MHC molecules interrupts the inhibitory signals, allowing activation of NK cells and lysis of target cells.

germline-encoded *pattern recognition receptors* (*PRRs*). Upon PAMP recognition, PRRs come in contact with the cell surface and/or send intracellular signals to the host that trigger proinflammatory and antimicrobial responses including the synthesis and release of cytokines, chemokines, and cell adhesion molecules.[3] The PAMPs recognized by the host PRRs are made up of a combination of sugars, lipid molecules, proteins, or patterns of modified nucleic acids and are essential to the functioning and infectivity of the pathogen. Because the PAMPs are essential for the functioning of the microorganism, mutation cannot help it avoid immune recognition. The human complement of PRRs is very extensive (approximately 1000) so the classes of pathogens recognized by them are very diverse. Therefore, pathogens of very different biochemical composition are recognized by relatively similar mechanisms by host PRRs, and no single class of pathogens is sensed by only one type of PRR. Therefore, the host genetic code allows for the unique receptors involved in both innate and adaptive immunity to recognize fine details of molecular structure.

The ability of the innate immune response to limit microbes early in the infectious process results from the binding of pathogens to the PRRs on leukocytes, which in turn initiates the signaling events that lead to complement activation, phagocytosis, and autophagy. Once initiated, white blood cells, neutrophils, and monocytes migrate from the blood to the tissues, along with other body fluids causing peripheral edema. Blood monocytes mature into macrophages as they traverse the tissues and join the macrophages and DCs already present in the tissues. PRRs present on these cells become activated, which amplifies the inflammatory response through enhanced secretion of all chemical mediators including cytokines and complement.

Toll-Like Receptors

The most studied PRRs associated with the innate immune response are the *Toll-like receptors* (*TLRs*). TLRs derive their name from the study of the *Drosophila melanogaster* toll protein, which is responsible for the resistance of *Drosophila* to bacterial and fungal infections.[3,4] Structurally, TLRs are integral glycoproteins that possess an extracellular or luminal ligand-binding site containing leucine-rich repeats and a cytoplasmic signaling toll/interleukin-1 (IL-1) domain.[17] Binding of PAMP to a TLR induces a conformational change in the receptor, which subsequently triggers intracellular signal transduction and activation of cellular processes, such as activation of transcription factors such as nuclear factor κβ (NF-κβ). NF-κβ regulates the production of a number of proteins that are important components of innate immunity. TLRs can be found in most of the bone marrow cells including the macrophages, DCs, neutrophils, T cells, B cells, and non–bone marrow cells including epithelial and fibrocytes. Eleven different TLRs have been identified in humans, and they each recognize distinct PAMPs derived from various microorganisms including bacteria, viruses, fungi, and protozoa.[18]

Human TLRs can be divided into subfamilies that primarily recognize related PAMPs. TLR1, TLR2, TLR4, and TLR6 recognize lipids and lipopolysaccharides (LPS), whereas TLR3, TLR7, TLR8, and TLR9 recognize nucleic acids.[18] TLRs can also be classified according to their cellular distribution such that TLR1, TLR2, TLR4, TLR5, TLR6, TLR10, and TLR11 are expressed extracellularly and THR3, TLR7, TLR8, and TLR9 are mainly expressed in intracellular compartments.[19,20] These receptors are involved in responses to widely divergent types of molecules that are commonly expressed by microbial, but not mammalian, cell types. For example, TLR4 is essential for phagocytic recognition and response to the LPS present in

Characteristics	Innate immunity	Adaptive immunity
Recognition	Molecular patterns common to microbes Different microbes Identical mannose receptor	Specific microbial molecules Different microbes Distinct antibodies
Receptors	Limited diversity expressed by germline genes Toll-like receptor Mannose receptor	Great diversity expressed through recombination of somatic genes B-cell receptor B cell Plasma cell Antibody
Cellular expression	Effector cell types express identical receptors (*e.g.,* neutrophils express Toll-like receptors).	Each clone of lymphocytes expresses unique receptors.
Self–nonself discrimination	Yes, by recognizing molecules unique to pathogen, NK cells recognize MHC-I self-recognizing molecules.	Yes, lymphocytes use MHC-I and -II and foreign peptides (*e.g.,* microbial peptides in recognition).

FIGURE 13.3 • Recognition systems of innate and adaptive immunity.

gram-negative bacteria. TLR2 binds to peptidoglycan, which is an essential component of the cell wall of gram-positive bacteria. Finally, TLR5 can recognize the protein flagellin found in flagellated bacteria. In addition to their role in the immune response, TLRs have been shown to have a pathologic role in disorders such as atherosclerosis, allergies, and certain autoimmune diseases.[21,22]

Soluble Mediators of Innate Immunity

While cells of the innate immune system communicate critical information about invading microorganisms and self–nonself recognition through cell-to-cell contact, soluble mediators are also essential for many other aspects of the innate immune response. Development of innate immune response is very much dependent upon the secretion of soluble molecules such as opsonins, cytokines, and acute-phase proteins.

Opsonins

Opsonins are molecules that coat negatively charged particles on cell membranes and as a result enhance the recognition and binding of phagocytic cells to microorganisms. The process by which the cellular particles on microbes are coated is called opsonization. Once the opsonin binds to the microbe, it is able to activate the phagocyte after attachment to a PRR on the phagocytic cell. There are several opsonins important in innate immunity and the acute inflammatory process including acute-phase proteins, lectins, and complement. Components of the adaptive immune response can also act as opsonins. For example, when the humoral response is activated, IgG and IgM antibodies can coat cellular particles on pathogens and bind to Fc receptors on neutrophils and macrophages, enhancing the phagocytic function of innate cells.

Inflammatory Cytokines

Cytokines are low molecular weight proteins that serve as soluble chemical messengers and which mediate the interaction between immune and tissue cells. They are part of an integrated signaling network with extensive functions in both the innate (nonspecific) and adaptive immune defenses. The cytokines involved in innate immunity include TNF-α and lymphotoxin; interferons (IFN-γ, IFN-α, IFN-β); the interleukins IL-1, IL-6, and IL-12; and chemokines (see Table 13.2). These substances modulate innate immunity by stimulating the development of cells involved in both innate and adaptive immunity, producing chemotaxis within leukocytes, stimulating acute-phase protein production, and inhibiting viral replication. Once an innate immune phagocyte is activated via PRR–PAMP with a pathogen, cytokines are released into the surrounding tissues where they exert their effect. If large numbers of cells are activated, then cytokines may be able to stimulate inflammatory processes in tissues far from the initial site of infection. Under normal circumstances, the duration of activity of cytokines is relatively short so that a prolonged immune response does not occur.

TNF-α and lymphotoxins are cytokines that are structurally related and that have similar cytotoxic activities.[23] The two cytokines differ in that TNF-α can be secreted by a variety of immune cells, but the lymphotoxins are predominantly secreted by activated lymphocytes and NK cells. These cytokines regulate development of the lymphoid tissues and the inflammatory process through induction of adhesion molecules and other cytokines/chemokines.[23] The IFNs are another family of cytokines that are critically involved in initiating and enhancing the cellular immune response to viral infection of host cells. In addition, they play a key role in amplifying the presentation of antigens to specific T cells. Type I interferon (IFN-α and IFN-β) is secreted by virus-infected cells, while type II, immune or gamma interferon (IFN-γ), is mainly secreted by T cells, NK cells, and macrophages.[23,24] When activated IFNs interact with specific cellular receptors, causing the expression of antiviral and immune modulatory genes. IFNs activate macrophages, induce B cells to switch Ig type, alter T-helper response, inhibit cell growth, promote apoptosis, and induce an antiviral state in uninfected cells. Finally, ILs help to regulate the immune response by increasing the expression of adhesion molecules on endothelial cells, stimulating migration of leukocytes into infected tissues, and by stimulating the production of antibodies by the cells of the adaptive immune response.

Acute-Phase Proteins

Two acute-phase proteins that are involved in the defense against infections are the mannose-binding ligand (MBL) and C-reactive protein (CRP). MBL and CRP are produced in the liver in response to activation of proinflammatory cytokines. MBL binds specifically to mannose residues, and CRP binds to both phospholipids and sugars that are found on the surface of microbes. These substances act as "costimulatory" opsonins and enhance the binding of phagocytic cells to suboptimally opsonized invading microorganisms.[25] They also act as activators of the alternative complement pathway.

The Complement System

The complement system is a powerful effector mechanism of both innate and adaptive immunity that allows the body to localize infection and destroy invading microorganisms. The complement system is composed of group of proteins found in the circulation and in various extracellular fluids. The proteins of the complement system normally circulate as inactive precursors. When activated a series of proteolytic and protein–protein interactions is initiated that ultimately culminates in opsonization of invading pathogens, migration of leukocytes to the site of invasion, initiation of a localized inflammatory reaction, and ultimate lysis of the pathogen.[25] The proteins of the complement system are mainly proteolytic enzymes and make up approximately 10% to 15% of the plasma proteins. For a complement reaction to occur, the complement components must be activated in the proper sequence. Inhibitor proteins and the instability of the activated complement proteins at each step of the process prevent uncontrolled activation of the complement system.

There are three parallel but independent pathways that result in activation of the complement system during the innate immune response: the classical, the lectin, and the alternative pathways. The reactions of the complement systems can be divided into three phases:

1. Initiation or activation
2. Amplification of inflammation
3. Membrane attack response

The three pathways differ in the proteins used in the early stage of activation, but all ultimately converge on the key complement protein C3, which is essential for the amplification stage. Activated C3 then activates all subsequent complement molecules (C5 through C9) resulting in the ultimate lysis of cells.

The classic pathway is initiated by an antigen–antibody complex (either IgG or IgM mediated), which causes a specific reactive site on the antibody to be "uncovered" so that it can bind directly to the C1 molecule in the complement system. Once C1 is activated, a "cascade" of sequential reactions is set in motion. Initially a small amount of enzyme is produced, but with activation of successive complement proteins successively increasing, concentrations of proteolytic enzymes are produced. This process is known as *amplification*. In the lectin or alternative complement pathway, inactive circulating complement proteins are activated when they are exposed to microbial surface polysaccharides, MBL, CRP, and other soluble mediators that are integral to innate immunity. Like the classic pathway, the lectin and alternative pathways create a series of enzymatic reactions that cleave successive complement proteins in the pathway.

During the activation phase of the complement cascade, cleavage of C3 produces C3a and C3b. C3b is a key opsonin that coats bacteria and allows them to be phagocytized after binding to type I complement receptor on leukocytes. The presence of C3a triggers the migration of neutrophils into the tissues to enhance the inflammatory response. Production of C3a, C4a, and C5a also leads to activation of mast cells and basophils causing them to release histamine, heparin, and other substances. These mediators of the inflammatory response increase tissue blood flow and increase localized capillary permeability allowing increased leakage of fluids and protein into the area. In addition, they stimulate changes in the endothelial cells in order to stimulate chemotaxis of neutrophils and macrophages to the site of inflammation. During the late phase of the complement cascade, cleavage of C5 triggers the assembly of a membrane attack complex from the C5 to C9 proteins. The resulting complex creates a tubelike structure, which penetrates the microbial cell membrane allowing the passage of ions, small molecules, and water into the cell, causing the cell to ultimately burst. The multiple and complementary functions of the complement system make it an integral component of innate immunity and inflammation. It also serves as an essential bridge between the innate and humoral responses. Pathophysiological manifestations associated with deficiencies of complement range from increased susceptibility to infection to inflammatory tissue and autoimmune disorders that are the result of impaired activated complement clearance.

IN SUMMARY

The innate immune system is a complex system that works in an organized, rapid, yet nonspecific fashion as the body's first line of defense against invasion. It is comprised of the epithelial cells of the skin and mucus membranes; phagocytic cells such as the neutrophils, macrophages, and NK cells; and a series of plasma proteins including cytokines, chemokines, and the proteins of the complement system. These defenses exist before the body encounters an invading microorganism and are activated independent of the adaptive humoral response. The epithelial cells of the skin and mucous membranes block the entry of infectious agents and secrete antimicrobial enzymes, proteins, and peptides in an attempt to prevent microorganisms from invading the internal environment.

The phagocytes of the innate immune response engulf and digest infectious agents. They utilize PRRs, which are present on their membranes to recognize and bind broad patterns of molecules (PAMPs) shared by microbes and that are essential for their survival. TLRs are the most studied of all PRRs and are expressed on many of the cells of the innate immune system. TLRs are involved in responses to widely divergent types of molecules that are commonly expressed by microbial but not mammalian cell types.

Development of a healthy innate immune response is dependent not only upon the coordinated activity of the leukocytes but on the secretion of chemical mediators and soluble molecules, such as opsonins, cytokines, acute-phase proteins, and complement. Opsonins bind to and tag microorganisms for more efficient recognition by phagocytes. Activated leukocytes release cytokines that stimulate the migration of leukocytes to the site of inflammation, stimulate production of acute-phase proteins, and enhance phagocytosis.

The complement system is a primary effector system that functions as part of both the innate and adaptive immune responses. It is comprised of a group of proteins that are activated by three distinct but convergent pathways: the classical, the lecithin, and the alternative pathways. The primary function of the complement system is the promotion of inflammation and destruction of the microbes.

ADAPTIVE IMMUNITY

After completing this section of the chapter, you should be able to meet the following objectives:

- Characterize the significance and function of major histocompatibility complex molecules.
- Compare and contrast the development and function of the T and B lymphocytes.
- Describe the function of cytokines involved in the adaptive immune response.

The adaptive immune system is the final line of defense against infection and is activated once the innate immune response initiates the inflammatory process. In contrast to innate immunity, the adaptive immune response is capable of targeting specific cells or organisms that it recognizes as foreign to the body through activation of various lymphocytes and their products, including antibodies. The lymphocytes involved in adaptive immunity have the unique ability to remember specific pathogen and mount a heightened immune response during repeat exposures. Each exposure results in a more rapid and aggressive response. Substances present on the surface of pathogens or other foreign substances that elicit adaptive immune responses are called *antigens*. Adaptive immunity involves two distinct but interconnected mechanisms: humoral and cell-mediated responses. *Humoral immunity* is mediated by *B-lymphocyte* activation and subsequent antibody production. It is the primary defense against extracellular microbes and toxins. In contrast, *cell-mediated immunity* involves the activation of specific *T lymphocytes* (T-helper and T-cytotoxic lymphocytes), which are responsible for the body's defense against intracellular microbes such as viruses.

KEY POINTS

ADAPTIVE IMMUNITY

- The adaptive immune response involves a complex series of interactions between components of the immune system and the antigens of a foreign pathogen. It is able to distinguish between self and nonself, recognize and specifically react to large numbers of different microbes and pathogens, and remember the specific agents.
- Humoral immunity consists of protection provided by the B lymphocyte–derived plasma cells, which produce antibodies that travel in the blood and interact with circulating and cell surface antigens, whereas cell-mediated immunity provides protection through cytotoxic T lymphocytes, which protect against virus-infected or cancer cells.

Antigens

Antigens, or *immunogens*, are substances or molecules that are foreign to the body but when introduced trigger the production of antibodies by B lymphocytes leading to the ultimate destruction of the invader. They are usually large macromolecules (>10,000 Da) such as proteins, polysaccharides, lipids, and free nucleic acids. Antigens are recognized by specific receptors present on the surface of lymphocytes and by the *antibodies* or *immunoglobulins* secreted in response to the antigen. Antigens can take the form of any foreign substance including bacteria, fungi, viruses, protozoa, parasites, and nonmicrobial agents such as plant pollens, insect venom, and transplanted organs.

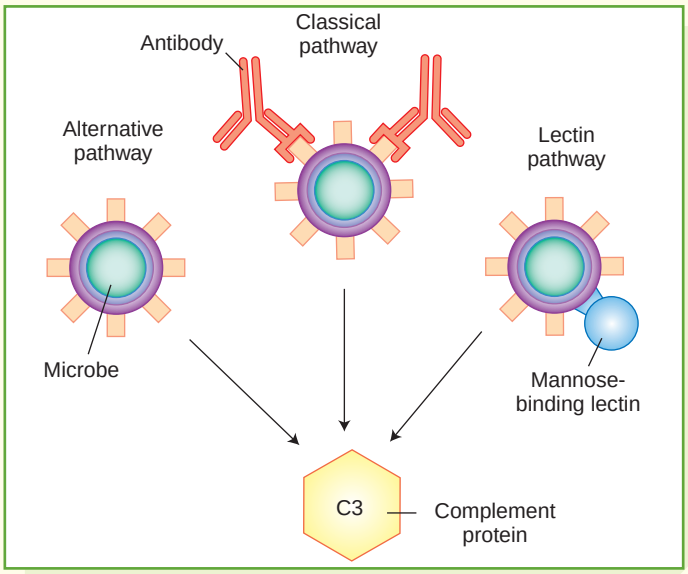
Understanding The Complement System

The complement system provides one of the major effector mechanisms of both humoral and innate immunity. The system consists of a group of proteins (complement proteins C1 through C9) that are normally present in the plasma in an inactive form. Activation of the complement system is a highly regulated process, involving the sequential breakdown of the complement proteins to generate a cascade of cleavage products capable of proteolytic enzyme activity. This allows for tremendous amplification because each enzyme molecule activated by one step can generate multiple activated enzyme molecules at the next step. Complement activation is inhibited by proteins that are present on normal host cells; thus, its actions are limited to microbes and other antigens that lack these inhibitory proteins.

The reactions of the complement system can be divided into three phases: (1) the initial activation phase, (2) the early-step inflammatory responses, and (3) the late-step membrane attack responses.

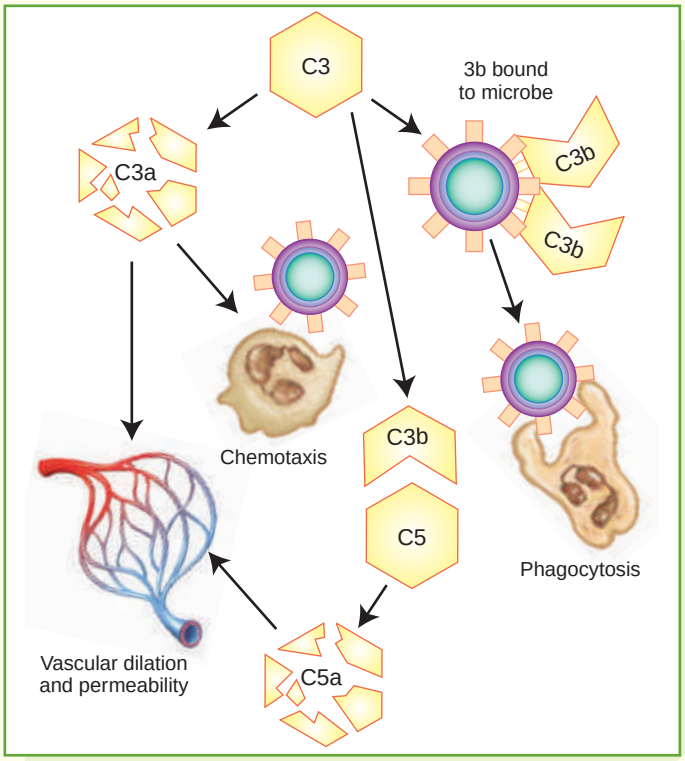
Initial Activation Phase

There are three pathways for recognizing microbes and activating the complement system: (1) the alternative pathway, which is activated on microbial cell surfaces in the absence of antibody and is a component of innate immunity; (2) the classical pathway, which is activated by certain types of antibodies bound to antigen and is part of humoral immunity; and (3) the lectin pathway, which is activated by a plasma lectin that binds to mannose on microbes and activates the classical system pathway in the absence of antibody.



Early-Step Inflammatory Responses

The central component of complement for all three pathways is the activation of the complement protein C3 and its enzymatic cleavage into a larger C3b fragment and a smaller C3a fragment. The smaller 3a fragment stimulates inflammation by acting as a chemoattractant for neutrophils. The larger 3b fragment becomes attached to the microbe and acts as an opsonin for phagocytosis. It also acts as an enzyme to cleave C5 into two components: a C5a fragment, which produces vasodilation and increases vascular permeability, and a C5b fragment, which leads to the late-step membrane attack responses.

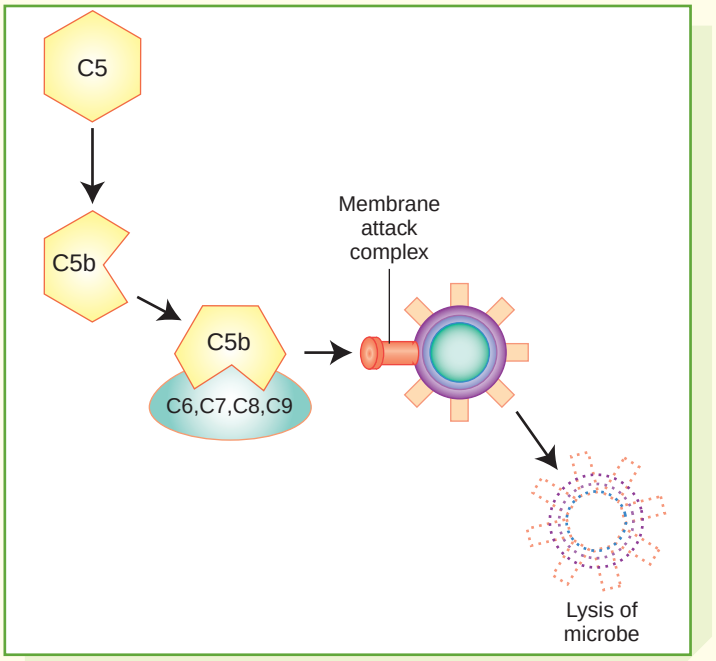


Continued

Understanding The Complement System *(Continued)*

Late-Step Membrane Attack

In the late-step responses, C3b binds to other complement proteins to form an enzyme that cleaves C5, generating C5a and C5b fragments. C5a stimulates the influx of neutrophils and the vascular phase of acute inflammation. The C5b fragment, which remains attached to the microbe, initiates the formation of a complex of complement proteins C6, C7, C8, and C9 into a membrane attack complex protein, or pore, that allows fluids and ions to enter and cause cell lysis.



Antigens possess immunologically active sites called *antigenic determinants*, or epitopes. These are smaller, discrete components of the antigen that have a unique molecular shape, which can be recognized by and bound to a specific Ig receptor found on the surface of the lymphocyte or by an antigen-binding site of a secreted antibody (Fig. 13.4). It is not unusual for a single antigen to possess several antigenic determinants and, therefore, be capable of stimulating several different T and B lymphocytes. For example, different proteins that comprise the influenza virus may function as unique antigens (A, B, C, H, and N antigens), each of which contains several antigenic determinants. Hundreds of antigenic determinants are found on structures such as the bacterial cell wall.

Low molecular weight molecules (<10,000 Da) may contain antigenic determinants but alone are usually unable to stimulate an immune response. These molecules are known as haptens. When they are complexed with an immunogenic carrier (usually a protein), they function as antigens. Many haptens exist in nature and frequently create problems for humans. Urushiol is a toxin found in the oils on poison ivy that

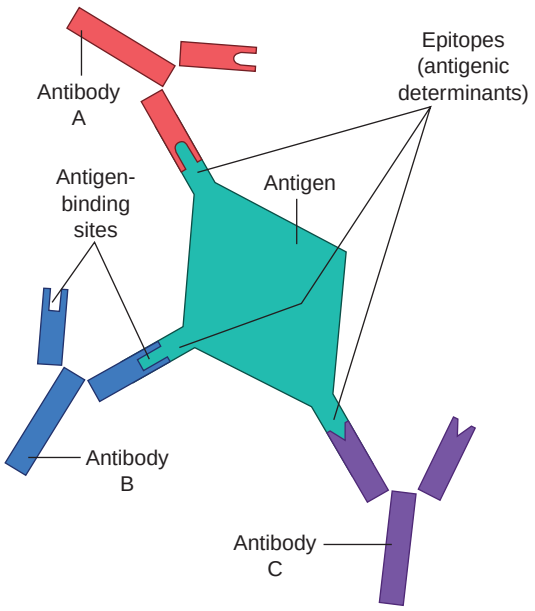


FIGURE 13.4 • Multiple epitopes on a complex antigen being recognized by their respective (*A*, *B*, *C*) antibodies.

is responsible for initiating an allergic reaction. An allergic response to the antibiotic penicillin is an example of a medically important reaction due to hapten–carrier complexes. The penicillin molecule is very small (350 Da) and usually nonantigenic. However, in susceptible people it can complex with carrier proteins in the body, which are then recognized as "foreign" and capable of initiating an antigen–antibody reaction.

Cells of Adaptive Immunity

The principal cells of the adaptive immune system are the lymphocytes, APCs, and effector cells.

Lymphocytes

Lymphocytes make up approximately 36% of the total white cell count and are the primary cells of the adaptive immune response. They arise from the lymphoid stem cell line in the bone marrow and differentiate into two distinct but interrelated cell types: the B lymphocytes and T lymphocytes. B lymphocytes are responsible for forming the antibodies that provide humoral immunity, whereas T lymphocytes provide cell-mediated immunity. T and B lymphocytes are unique in that they are the only cells in the body capable of recognizing specific antigens present on the surfaces of microbial agents and other pathogens. As a result, adaptive immune processes are organism specific and possess the capacity for memory.

The recognition of specific surface antigens by lymphocytes is made possible because of the presence of specific receptors or antibodies on the surface of B and T lymphocytes. Scientists have been able to identify these specific proteins and correlate them with a specific cellular function. This has lead to the development of a classification system for these surface molecules known as the "cluster of differentiation" (CD). The nomenclature for the surface proteins utilizes the letters "CD" followed by a number that specifies the surface proteins that define a particular cell type or stage of cell differentiation and are recognized by a cluster or group of antibodies. The utilization of this nomenclature has spread to other immune cells and cytokines all of which contribute to the acquired immune response.

Leukocytes involved in the innate immune response, such as macrophages and DCs, also play a key role in adaptive immunity because they function as APCs. They are capable of processing complex antigens into epitopes, which are then displayed on their cell membranes in order to activate the appropriate lymphocytes. Functionally, there are two types of immune cells: regulatory cells and effector cells. The *regulatory cells* assist in orchestrating and controlling the immune response, while effector cells carry out the elimination of the antigen (microbial, nonmicrobial, or toxin). In the body, helper T lymphocytes activate other lymphocytes and phagocytes, while regulatory T cells keep these cells in check so that an exaggerated immune response does not occur. Cytotoxic T lymphocytes, macrophages, and other leukocytes function as effector cells in different immune responses.

While T and B lymphocytes are generated from lymphoid stem cells in the bone marrow, they do not stay there to mature.

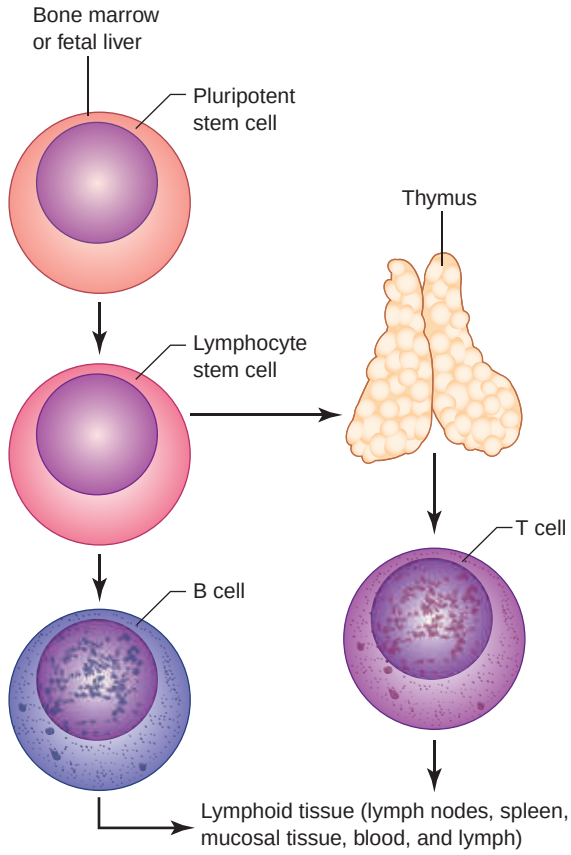


FIGURE 13.5 • Pathway for T- and B-cell differentiation.

Undifferentiated, immature lymphocytes migrate to lymphoid tissues, where they develop into distinct types of mature lymphocytes (Fig. 13.5). The T lymphocytes first migrate to the thymus gland where they divide rapidly and develop extensive diversity in their ability to react against different antigens.[26] Each T lymphocyte develops specificity against a specific antigen. Once this differentiation occurs, the lymphocytes leave the thymus gland and migrate via the bloodstream to peripheral lymphoid tissue. At this time, they have been preprogrammed not to attack the body's own issues. Unfortunately, in many autoimmune diseases it is believed that this process goes astray. The *B lymphocytes* mature primarily in the bone marrow and are essential for humoral, or *antibody-mediated, immunity*. Unlike the T lymphocytes, where the entire cell is involved in the immune response, B lymphocytes secrete antibodies, which then act as the reactive agent in the immune process. Therefore, the lymphocytes are distinguished by their function and response to antigen, their cell membrane molecules and receptors, their types of secreted proteins, and their tissue location. High concentrations of mature lymphocytes are found in the lymph tissue throughout the body including the lymph nodes, spleen, skin, and mucosal tissues.

T and B lymphocytes possess all of the processes necessary for the adaptive immune response—specificity, diversity, memory, and self–nonself recognition. When antigens come in contact with the lymphocytes in the lymphoid tissues of the

body, specific T cells become activated and specific B cells are stimulated to produce antibodies. Once the first encounter occurs, these cells can exactly recognize a particular microorganism or foreign molecule because each lymphocyte is capable of targeting a specific antigen and differentiating the invader from self or from other substances that may be similar to it. Cell-mediated and humoral immunity is capable of responding to millions of antigens each day because there is an enormous variety of lymphocytes that have been programmed and selected during cellular development. Once the invading substance or organism has been removed from the body, the lymphocytes "remember" the presenting antigen and can respond rapidly during the next encounter. These lymphocytes are called "memory" T and B lymphocytes. They remain in the body for a longer period of time than their predecessors and as a result can respond more rapidly on repeat exposure. The immune system usually can respond to commonly encountered microorganisms so efficiently that we are unaware of the response.

B and T lymphocyte activation is triggered by antigen presentation to unique surface receptors (Fig. 13.6). The antigen receptor present on the B lymphocyte consists of membrane-bound Ig molecules that can bind a specific epitope. However, in order for B lymphocytes to produce antibodies, they require the help of specific T lymphocytes, called *helper T cells*. While the B lymphocytes bind to one determinant (or hapten) on an antigen molecule, the antigen-specific helper T cell recognizes and binds to another determinant known as the "carrier." The carrier is an APC, which has previously picked

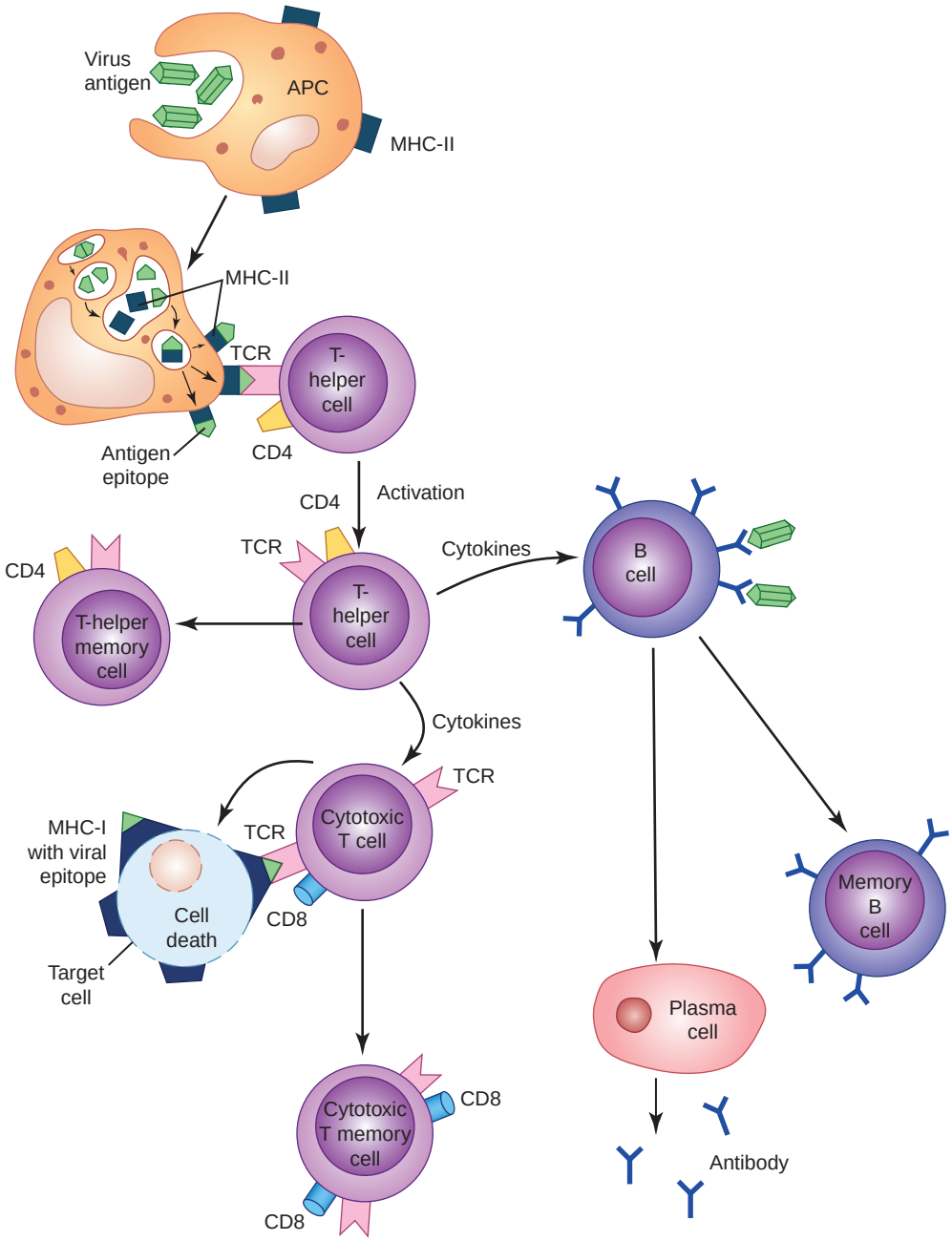


FIGURE 13.6 • Pathway for immune cell participation in an immune response. (APC, antigen-presenting cell; MHC, major histocompatibility complex; TCR, T-cell receptor.)

up the specified antigen. This interaction (B cell–T cell–APC) is restricted by the presence of cellular products genetically encoded by a self-recognition protein, called a *major histocompatibility complex* (MHC) molecule. This allows the lymphocyte to differentiate between self and foreign peptides.

Once the B and T lymphocytes are activated and amplified by cytokines released as part of the innate response, the lymphocytes divide several times to form populations or clones of cells that continue to differentiate into several types of effector and memory cells. In the adaptive immune response, the effector cells destroy the antigens and the memory cells retain the ability to target antigen during future encounters.

Major Histocompatibility Complex Molecules

In order for the adaptive immune response to function properly, it must be able to discriminate between molecules that are native to the body and those that are foreign or harmful to the body. The T lymphocytes are designed to respond to a limitless number of antigens, but at the same time they need to be able to ignore self-antigens expressed on tissues. The MHC molecules enable the lymphocytes to do just this. The MHC is a large cluster of genes located on the short arm of chromosome 6. The complex occupies approximately 4 million base pairs and contains 128 different genes, only some of which play a role in the immune response. The MHC genes are divided in three classes: I, II, and III, based upon their underlying function (Fig. 13.7).

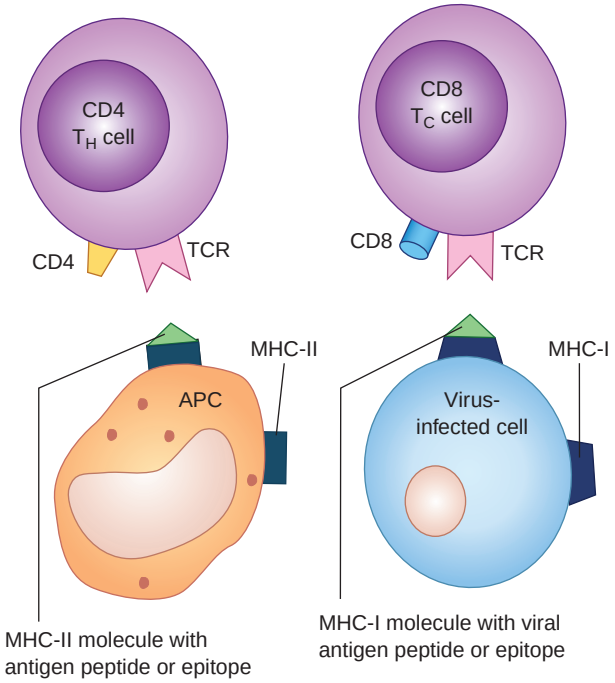


FIGURE 13.7 • Recognition by a T-cell receptor (TCR) on a $CD4^+$ helper T (T_H) cell of an epitope associated with a class II major histocompatibility complex (MHC) molecule on an antigen-presenting cell (APC) and by a TCR on a $CD8^+$ cytotoxic T (T_C) cell of an epitope associated with a class I MHC molecule on a virus-infected cell.

The class I and II MHC genes are responsible for encoding human leukocyte antigens (HLAs), which are proteins found on cell surfaces and define the individual's tissue type. These molecules are present on the cell surface glycoproteins that form the basis for human tissue typing. Each individual has a unique collection of MHC proteins representing a unique set of polymorphisms. MHC polymorphisms affect immune responses as well as susceptibility to a number of diseases. Because of the number of MHC genes and the possibility of several alleles for each gene, it is almost impossible for any two individuals to have an identical MHC profile.

The class I and II MHC genes also encode proteins that play an important role in antigen presentation. Protein fragments from inside the cell are displayed by MHC complex on the cell surface, allowing the immune system to differentiate between the body's own tissues and foreign substances. Cells, which present unfamiliar peptide fragments on the cell surface, are attacked and destroyed by the B and T lymphocytes. Class III MHC genes encode for many of the components of the complement system and play an important role in the innate immune process.

The MHC-I complexes contain a groove that accommodates a peptide fragment. T-cytotoxic cells can only become activated if they are presented with a foreign antigen peptide. MHC-1 complexes may present degraded viral protein fragments from infected cells. *Class II MHC* (MHC-II) molecules are found only on phagocytic APCs, immune cells that engulf foreign particles including bacteria and other microbes. This includes the macrophages, DCs, and B lymphocytes, which communicate with the antigen receptor and CD4 molecule on T-helper lymphocytes.

Like class I MHC proteins, class II MHC proteins have a groove or cleft that binds a fragment of antigen. However, these bind fragments from pathogens that have been engulfed and digested during the process of phagocytosis. The engulfed pathogen is degraded into free peptide fragments within cytoplasmic vesicles and then complexed with the MHC-II molecules on the surface of the cells.[26,27] T-helper cells recognize these complexes on the surface of APCs and become activated.

The first human MHC proteins discovered are called *human leukocyte antigens* (*HLAs*) and are so named because they were identified on the surface of white blood cells. HLAs are the major target involved in organ transplant rejection and as a result are the focus of a great deal of research in immunology. Recent analysis of the genes for the HLA molecules has allowed for better understanding of the proteins involved in this response. The classic human MHC-I molecules are divided into types called HLA-A, HLA-B, and HLA-C, and the MHC-II molecules are identified as HLA-DR, HLA-DP, and HLA-DQ (Table 13.3). Multiple alleles or alternative genes can occupy each of the gene loci that encode for HLA molecules. More than 350 possible alleles for the A locus, 650 alleles for the B locus, and 180 alleles for the C locus have been identified. These genes and their expressed MHC molecules are designated by a letter and numbers (*i.e.*, HLA-B27).

TABLE 13.3 PROPERTIES OF CLASS I AND II MHC MOLECULES

PROPERTIES	HLA ANTIGENS	DISTRIBUTION	FUNCTIONS
Class I MHC	HLA-A, HLA-B, HLA-C	Virtually all nucleated cells	Present processed antigen to cytotoxic $CD8^+$ T cells; restrict cytolysis to virus-infected cells, tumor cells, and transplanted cells
Class II MHC	HLA-DR, HLA-DP, HLA-DQ	Immune cells, antigen-presenting cells, B cells, and macrophages	Present processed antigenic fragments to $CD4^+$ T cells; necessary for effective interaction among immune cells

HLA, human leukocyte antigen; MHC, major histocompatibility complex.

HLA genes are inherited as a unit, called a *haplotype*, because the class I and II MHC genes are closely linked on one chromosome. Since each person inherits one chromosome from each parent, each person has two HLA haplotypes. Tissue typing in forensics and organ transplantation involves the identification of these haplotypes. In organ or tissue transplantation, the closer the matching of HLA types, the greater is the probability of identical antigens and the lower the chance of rejection. However, not all people that develop organ rejection after transplantation develop anti-HLA antibodies. Non-HLA target antigens exist including the MHC class I chain-related antigens A (MICA).[28] These antigens are expressed on epithelial cells, monocytes, fibroblasts, and endothelial cells. Therefore, donor-specific antibodies are not detected prior to organ tissue typing prior to transplantation because they are not expressed on the leukocytes tested.[28]

Antigen-Presenting Cells

During the adaptive immune response, activation of a T lymphocyte requires the recognition of a foreign peptide (antigen) bound to a self-MHC molecule. This process requires that stimulatory signals be delivered simultaneously to the T lymphocyte by another specialized cell known as an *antigen- presenting cell (APC)*. Therefore, APCs play a key role in bridging the innate and adaptive immune systems through cytokine-driven up-regulation of MHC-II molecules. Cells that function as APCs must be able to express both classes of MHC molecules and include DCs, monocytes, macrophages, and B lymphocytes residing in lymphoid follicles. Under certain conditions, endothelial cells are also able to function as APCs. APCs have been shown to play a key role in the development of autoimmune diseases and atherosclerosis. Activated T lymphocytes appear to be proatherogenic, and in experimental models, APC and T-cell deficiency have been associated with up to an 80% reduction in atherosclerosis.[29]

Macrophages function as a principal APC. They are key cells of the mononuclear phagocytic system and engulf and digest microbes and other foreign substances that gain access to the body. Since macrophages arise from monocytes in the blood, they can move freely throughout the body to the appropriate site of action. Tissue macrophages are scattered in connective tissue or clustered in organs such as the lung (*i.e.,* alveolar macrophages), liver (*i.e.,* Kupffer cells), spleen, lymph nodes, peritoneum, central nervous system (*i.e.,* microglial cells), and other areas. Macrophages are activated during the innate immune response where they engulf and break down complex antigens into peptide fragments. These fragments can then be associated with MHC-II molecules for presentation to cells of the "cell-mediated" response so that self–nonself recognition and activation of the immune response can occur.

DCs are also responsible for presenting processed antigen to activated T lymphocytes. The starlike structure of the DCs provides an extensive surface area rich in MHC-II molecules and other non-HLA molecules important for initiation of adaptive immunity. DCs are found throughout the body in tissues where antigen enters the body and in the peripheral lymphoid tissues. Both DCs and macrophages are capable of "specialization" depending upon their location in the body. For example, Langerhans cells are specialized DCs in the skin, whereas follicular DCs are found in the lymph nodes. Langerhans cells transport antigens found on the skin to nearby lymph nodes for destruction. They are also involved in the development of cell-mediated immune reactions such as allergic type IV contact dermatitis. Finally, DCs are found in the mucosal lining of the bowel and have been implicated in the development of inflammatory bowel diseases such as Crohn disease and ulcerative colitis, where they present antigens to the B and T lymphocytes through the production of proinflammatory cytokines.[22]

B Lymphocytes and Humoral Immunity

The humoral immune response is mediated by antibodies, which are produced by the B lymphocytes. The primary functions of the B lymphocytes are the elimination of extracellular microbes and toxins and subsequent "memory" for a heightened response during future encounters. Humoral immunity is more important than cellular immunity in defending against microbes with capsules rich in polysaccharides and lipid toxins because only the B lymphocytes are capable of responding to and producing antibodies specific for many types of these molecules. The T cells, which are the mediators of cellular immunity, respond primarily to surface protein antigens.

B lymphocytes are produced in the bone borrow and are classified according to the MHC-II proteins, Ig, and

complement receptors expressed on the cell membrane. During development Ig gene rearrangement takes place to insure that only B lymphocytes are capable of producing antibodies (Ig). At each stage of development, a cell-specific pattern of Ig gene is expressed, which then serves as a phenotypic marker of these maturational stages. The B lymphocyte progenitors are known as pro-B and pre-B cells and develop into both mature and naive B lymphocytes in the bone marrow. Naïve (or immature) B lymphocytes display IgM on the cell surface. These immature cells respond to antigen differently from a mature B cell. They can be functionally removed from the body as a result of interaction with a self-antigen, by undergoing programmed cell death (apoptosis) or by the process of anergy where they become nonresponsive in the presence of the antigen. Naïve B lymphocytes can leave the bone marrow and migrate to peripheral or secondary lymphoid tissues such as the spleen and lymph nodes where they complete the maturation process. Once B lymphocytes become fully mature, they become capable of expressing IgD, in addition to the IgM on the cell membrane surface. Mature B lymphocytes are fully responsive to antigens and are capable of interacting with T cells.

The commitment of a B-cell line to a specific antigen is evidenced by the expression of the membrane-bound Ig receptors that recognize the specific antigen. Initially, when mature B lymphocytes encounter antigens that are complementary to their encoded surface Ig receptor and in the presence of T lymphocyte antigen presentation, they undergo a series of conformational changes that transform them into antibody-secreting plasma cells or into memory B cells (Fig. 13.8). Both cell types are necessary for the ultimate success of the humoral response. The antibodies produced by the plasma cells are released into the lymph and blood, where they can then bind and remove their specific antigen with the help of other immune effector cells and molecules. The memory B lymphocytes have a longer life span and are distributed to the peripheral tissues in preparation for subsequent antigen exposure.

Immunoglobulins

Antibodies are protein molecules also known as *immunoglobulins*. Igs are classified into five different categories based upon their role in the humoral defense mechanisms. The five classes include IgG, IgA, IgM, IgD, and IgE (Table 13.4). The classic structure of Igs is comprised of four-polypeptide chains with at least two identical antigen-binding sites (Fig. 13.9). Each Ig is composed of two identical light (L) chains and two identical heavy (H) chains that form a characteristic "Y"-shaped molecule. The "Y" ends of the Ig molecule carry the antigen-binding sites and are called *Fab* (*i.e.,* antigen-binding) fragments. The tail end of the molecule, which is called the *Fc* fragment, determines the biologic and functional characteristics of the class of Igs.

The heavy and light chains of the Ig have certain amino acid sequences, which show constant (C) regions and variable (V) regions. The *constant regions* have sequences of amino acids that vary little among the antibodies of a particular class of Ig and determine the classification of the particular Ig (*e.g.,* IgG, IgE). The constant regions, therefore, determine the effector function of the particular antibody. For example, IgG can tag an antigen for recognition and destruction by phagocytes. In contrast, the amino acid sequences of the *variable regions* differ from antibody to antibody. They also contain the antigen-binding sites of the particular molecule. The different amino acid sequences found in these binding sites allow this region of the antibody to recognize its complementary epitope (antigen). The variable amino acid sequence determines the shape of the binding site, forming a three-dimensional pocket that is complementary to the specific antigen. When B lymphocytes divide, they form clones that produce antibodies with identical antigen-binding regions. During the course of the immune response, class switching (*e.g.,* from IgM to IgG) can occur, causing the B cell clone to produce one of the different Ig types.

IgG (gamma globulin) is the most abundant of the Igs making up 75% of the total circulating antibodies. It is a large molecule with a molecular weight of approximately 150 kDa

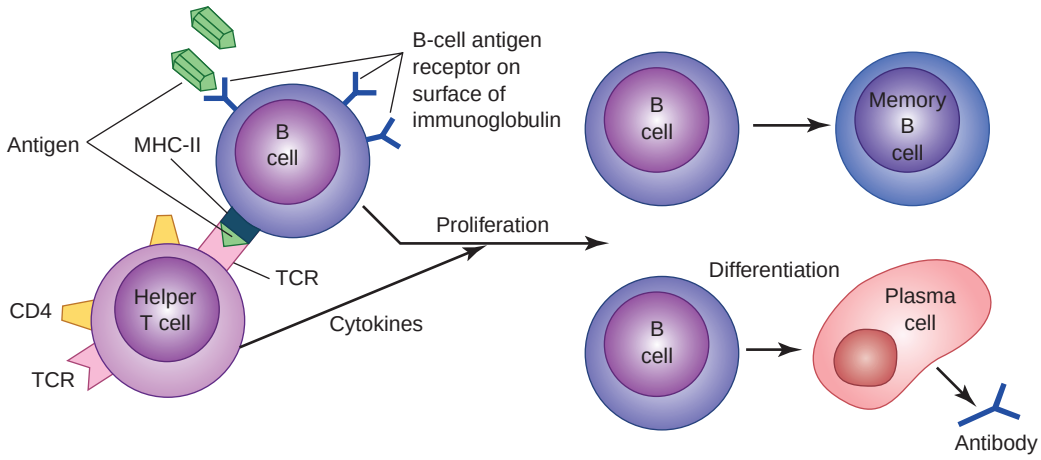
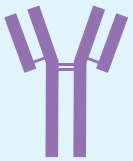
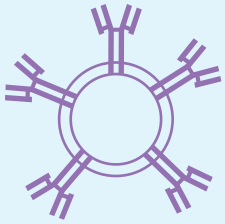
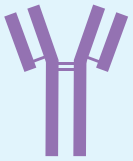
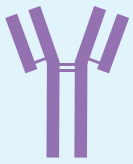


FIGURE 13.8 • Pathway for B-cell differentiation.

TABLE 13.4 CLASSES AND CHARACTERISTICS OF Igs

FIGURE	CLASS	PERCENTAGE OF TOTAL	CHARACTERISTICS
	IgG	75.0	Displays antiviral, antitoxin, and antibacterial properties; only Ig that crosses the placenta; responsible for protection of newborn; activates complement and binds to macrophages
	IgA	15.0	Predominant Ig in body secretions, such as saliva, nasal and respiratory secretions, and breast milk; protects mucous membranes
	IgM	10.0	Forms the natural antibodies such as those for ABO blood antigens; prominent in early immune responses; activates complement
	IgD	0.2	Found on B lymphocytes; needed for maturation of B cells
	IgE	0.004	Binds to mast cells and basophils; involved in parasitic infections, allergic and hypersensitivity reactions

and is composed of two different kinds of polypeptide chain. IgG possesses antiviral, antibacterial, and antitoxin properties. It is present in all body fluids, readily enters the tissues, and is capable of crossing the placenta where it confers immunity upon the fetus. Intact IgG functioning requires the help of APCs. It binds to target cells as well as Fc receptors on NK cells and macrophages, leading to lysis of the target cell. There are four subclasses of IgG (*i.e.,* IgG1, IgG2, IgG3, and IgG4) with specificity for certain types of antigens. For example, IgG2 appears to be responsive to bacteria that are encapsulated with a lipopolysaccharide layer, such as *Streptococcus pneumoniae*, *Neisseria gonorrhoeae*, and several strains of *Salmonella*.[30]

IgA possesses a dimeric structure and is the second most common Ig found in serum accounting for approximately 15% of all antibodies. It is primarily a secretory Ig that is found in saliva, tears, colostrum (*i.e.,* first milk of a nursing mother), and bronchial, gastrointestinal, prostatic, and vaginal secretions. Because it is primarily found in secretions, its primary function is in local immunity on mucosal surfaces. IgA prevents the attachment of viruses and bacteria to epithelial cells.

IgM accounts for approximately 10% of all circulating antibodies. It normally exists as a pentamer with identical heavy chains and identical light chains. Because of its structure, it is an efficient complement fixing Ig and is instrumental in the ultimate lysis of microorganisms. It also functions as an effective agglutinating antibody, capable of clumping organisms for eventual lysis and elimination. IgM is the first antibody to be produced by the developing fetus and by immature B lymphocytes.

IgD is a monomer found primarily on the cell membranes of B lymphocytes where it functions as a receptor for antigen. It circulates in the serum in extremely low levels where its function is essentially unknown. IgD on the surface of B lymphocytes contains extra amino acids at C-terminal so that it can successfully anchor to the membrane. It also associates with the Ig-alpha and Ig-beta chains.

IgE is the least common serum IgE because it binds very tightly to the Fc receptors on basophils and mast cells. It is involved in inflammation and allergic responses by causing mast cell degranulation and release of chemical mediators including histamine. IgE is also essential for combating parasitic infections.

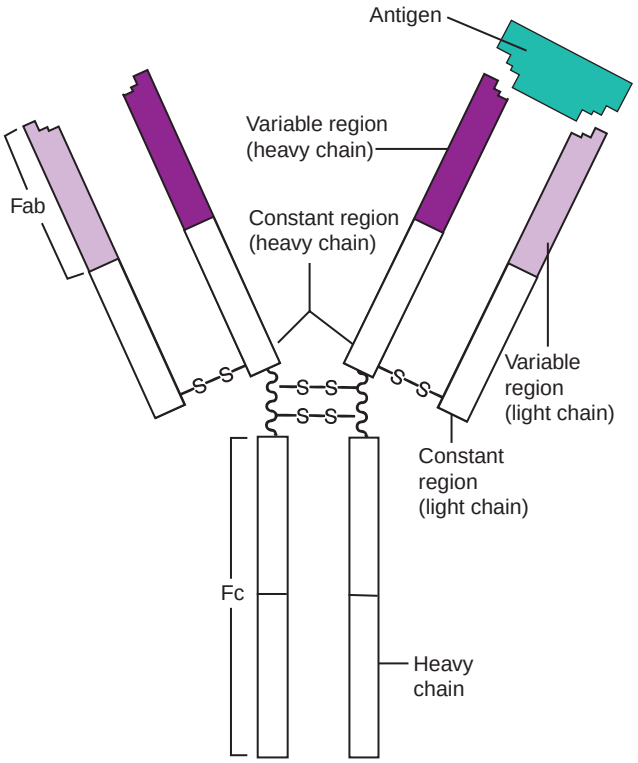


FIGURE 13.9 • Schematic model of an immunoglobulin G (IgG) molecule showing the constant and variable regions of the light and heavy chains.

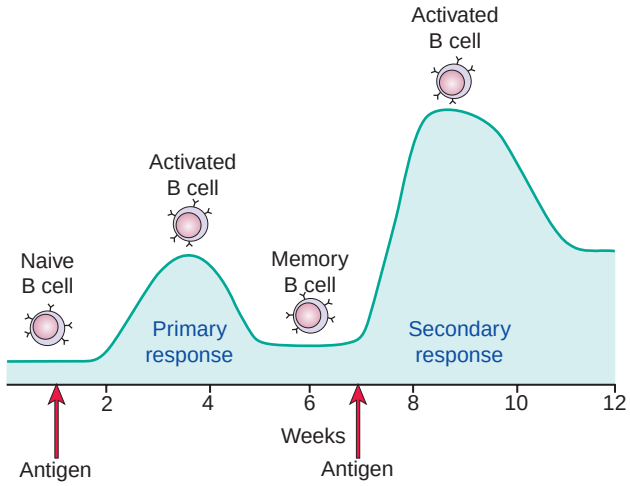


FIGURE 13.10 • Primary and secondary or memory phases of the humoral immune response to the same antigen.

Humoral Immunity

Humoral immunity requires the presence of mature B lymphocytes capable of recognizing antigen and which can ultimately mature into antibody-secreting plasma cells. The ultimate response of the antigen–antibody complex formation can take several forms including antigen–antibody complex precipitation, agglutination of pathogens, neutralization of toxins, phagocytosis or lysis of invading organisms, immune cell activation, and complement activation.

Two separate but interrelated responses occur in the development of humoral immunity: a primary and a secondary response (Fig. 13.10). A *primary immune response* develops when the body encounters the antigen for the first time. The antigen comes in contact with various APCs including macrophages, DCs, and B lymphocytes. The antigen is processed by these cells in association with the MHC-II molecules on the cells surface and then presented to the lymphocytes (*i.e.,* $CD4^+$ T-helper cells) to initiate the immune process. APCs such as macrophages also secrete ILs, which are essential for $CD4^+$ helper T cell activation.[31] The activated $CD4^+$ helper T cells trigger B cells to proliferate and differentiate into clone plasma cells that produce antibody. The primary immune response takes 1 to 2 weeks, but once generated, detectable antibody continues to rise for several more weeks even though the infectious process has resolved. The *memory phase or secondary immune response* occurs on subsequent exposure to the antigen. During the secondary response, the rise in antibody occurs sooner and reaches a higher level because of available memory cells.

During the primary response, B lymphocytes proliferate and differentiate into antibody-secreting plasma cells. A fraction of the activated B cells do not undergo differentiation but rather remain intact to form a pool of memory B lymphocytes that then become available to efficiently respond to invasion during subsequent exposure. Activated T cells can also generate primary and secondary cell-mediated immune responses and the concurrent development of T memory cells.

The immunization process makes use of the primary and secondary immune responses. The initial vaccination causes production of both plasma cells and memory cells. The plasma cells destroy the invading organism or toxin, and the memory cells provide defense against future exposure. "Booster" immunizations produce an immediate antigen–antibody response that simulates an immediate rise in antibody levels. Current phase I clinical immunization trials for cancer treatment show dense concentrations of $CD4^+$ and $CD8^+$ T lymphocytes and plasma cells in preexisting tumors after vaccination with irradiated malignant cells.[32]

T Lymphocytes and Cellular Immunity

T lymphocytes serve many functions in the immune system including the activation of other T cells and B cells, control of intracellular viral infections, rejection of foreign tissue grafts, activation of autoimmune processes, and activation of delayed hypersensitivity reactions. These processes make up the body's *cell-mediated* or *cellular immunity*. The effector phase of cell-mediated immunity is carried out by T lymphocytes and macrophages.

T lymphocytes arise from lymphoid stem cells in the bone marrow, but unlike B lymphocytes, they migrate to the thymus gland to undergo the process of maturation. The thymus gland is richly innervated and produces several peptide

hormones such as *thymulin* and *thymopoietin*, which are believed to be involved in T-cell maturation. T-cell precursors are attracted to the thymus by thymotaxin, a chemotactic factor secreted by thymic epithelial cells. Once the prothymocyte enters the cortex of the thymus, *terminal deoxynucleotidyl transferase* (TdT) is expressed causing gene rearrangement and increased TCR diversity. The pre–T lymphocytes are designated $CD3^+$, $CD4^-$, $CD8^-$, and double negative cells. The majority of these cells go on to rearrange their alpha and beta chain gene segments. The beta segment is expressed first resulting in the formation of a pre-TCR. This halts further gene rearrangement, enhances alpha chain gene rearrangement, and causes full maturation and expression of $CD4^+$ (helper) and $CD8^+$ (cytotoxic) lymphocytes. These are the predominant lymphocytes in the body. Mature T lymphocytes leave the thymus and migrate to peripheral lymphoid tissues, where they multiple and differentiate into memory T cells and various other mature lymphocytes upon encountering an antigen.

The TCR on the mature lymphocyte is composed of two polypeptides that fold to form a groove that recognizes processed antigen peptide–MHC complexes. It consists of two transmembrane molecules, the TCR-α and the TCR-β, that are the result of rearrangement of first the TCR-β and then the TCR-α gene.[33] The majority of TCRs recognize antigenic peptides that are bound to MHC-derived molecules. The TCR is associated with several surface molecules such as CD4 and CD8. CD4 is associated with the helper T cell, and CD8 is associated with the cytotoxic T cell. CD4 and CD8 help stabilize the TCR–antigen–MHC complex during T-cell activation. The TCR is also associated with other surface molecules known as the *CD3 complex*, which also aid in cell signaling.

Helper T Cells and Cytokines in Adaptive Immunity

The activation of helper T cells is the central event in the initiation of the humoral and cell-mediated immune response. $CD4^+$ helper T cells (T_H) serve as master regulators for the immune system. They become activated when their TCRs interact with antigens that are complexed with class II MHC on the surface of APCs. Once $CD4^+$ cells are activated, the cytokines they secreted in response influence the function of nearly all other cells of the immune system. Depending upon the specific cytokine that is released by the $CD4^+$ T cell the subsequent immunologic response will be activated. These cytokines are able to activate and regulate B cells, cytotoxic T lymphocytes, NK cells, macrophages, and other immune cells. The first cytokine to be produced by $CD4^+$ T cells after activation is IL-2. IL-2 is necessary for the proliferation and function of helper T cells, cytotoxic T cells, B cells, and NK cells. IL-2 interacts with T lymphocytes by binding to specific membrane receptors that are present on activated T cells but not on resting T cells. T-cell amplification relies on the presence of both IL-2 and IL-2 receptors; if either is missing, cell proliferation ceases. There are other cytokines that are not produced by $CD4^+$, but are essential for its function. IL-1 is produced by inflammatory cells and is responsible for increasing the expression of adhesion molecules on endothelial cells, enabling transmigration of leukocytes, and by stimulating antibody production.[23] Another cytokine essential for $CD4^+$ function is IL-6. IL-6 influences T cell effector functions by promoting helper T cell (T_2H) differentiation through up-regulation of NFATc2 and c-maf.[31]

The activated $CD4^+$ helper T cell can differentiate into two distinct subpopulations of helper T cells (*i.e.,* T_1H or T_2H) based on the cytokines secreted by the APCs at the site of activation (Table 13.5). Macrophages and DCs produce IL-12, which directs the maturation of $CD4^+$ helper T cells toward the T_1H subtype, whereas mast cells and T cells produce IL-4, which induces differentiation toward the T_2H subtype. The T_2H cells direct B lymphocytes to switch class and produce the IgE antibodies necessary for an allergic or hypersensitivity response. The distinct pattern of cytokine secreted by mature T_1H and T_2H cells further defines these subpopulations of T_H cells and determines whether a humoral or cell-mediated response will ultimately occur. Activated T_1H cells produce the cytokines IL-2 and IFN-γ, whereas T_2H cells produce IL-4 and IL-5. IL-5 is an activator of eosinophils that, along with IgE, functions in the control of helminth (intestinal parasite) infections. Some of the cytokines (*e.g.,* IL-10) made by T_2H cells are anti-inflammatory and inhibit macrophage activation and suppress other T_1H responses.

TABLE 13.5 COMPARISON OF PROPERTIES OF HELPER T-CELL SUBTYPES 1 (T_1H) AND 2 (T_2H)

	T_1H	T_2H
Stimulus for differentiation to T_H subtype	Microbes	Allergens and parasitic worms
Cells and cytokines influencing T_H subtype maturation	IL-12 produced by macrophages and DCs	IL-4 produced by mast cells and T cells
Cytokines secreted by T_H subtype	IFN-γ, IL-2	IL-4, IL-5
Effector functions	Phagocyte-mediated defense against infections, especially intracellular microbes; stimulates production of IgG	IgE- and eosinophil/mast cell–mediated immune reactions; stimulates production of IgE

NK, natural killer; IL, interleukin; IFN, interferon; Ig, immunoglobulin.

Regulatory T Cells

Regulatory T cells (T_R) are a subset of T lymphocytes that function to control immune system responses. Different populations of T_R cells produced in the thymus have been identified including those that express CD4 and CD25 on their surface. These cells represent a subset of $CD4^+$ cells that act as "negative regulators" of the immune process[34]. They suppress immune responses by inhibiting the proliferation of other potentially harmful self-reactive lymphocytes. Production of regulatory T cells is highly dependent upon the presence of antigen, activation of a TCR by the antigen, and the release of the cytokines IL-10 and transforming growth factor-β (TGF-β).[34] These cytokines inhibit the proliferation and activation of lymphocytes and macrophages. There is also recent evidence of the existence of regulatory $CD8^+$ T cells that can selectively down-regulate T cells activated by either self or foreign antigens. These cells differentiate into regulatory cells during the primary immune response and function to suppress the secondary immune response. The $CD8^+$ regulators are, therefore, primarily involved in self–nonself discrimination. The ability of the regulatory T cells to control many aspects of the immune response has significant implications for clinical practice. Promise has been shown in the control of inflammatory bowel disease, experimental allergic encephalitis, and autoimmune diabetes.

Cytotoxic T Cells

The primary function of cytotoxic T ($CD8^+$) cells is to monitor the activity of all cells in the body and destroy any that threaten the integrity of the body. $CD8^+$ T cells recognize antigens that are presented on the cell surface by MHC class I–derived molecules that sample peptides from protein degradation productions from inside cells infected by viruses or transformed by cancer[33] (Fig. 13.11). The ability of $CD8^+$ cells to recognize the class I MHC–antigen complexes on infected target cells ensures that neighboring uninfected host cells, which express class I MHC molecules alone or with self-peptide, are not indiscriminately destroyed. The $CD8^+$ cytotoxic T lymphocytes destroy target cells by a variety of mechanisms including the release of cytolytic enzymes, toxic cytokines, and pore-forming molecules (*i.e.,* perforins) or by triggering membrane molecules and intracellular apoptosis. Apoptosis is a normal biological process that eliminates excessive, dangerous, or damaged cells from the body. The $CD8^+$ T cells play a large role in controlling replicating viruses and intracellular bacteria because antibody cannot readily penetrate the membrane of living cells.

Cell-Mediated Immunity

In order for the cell-mediated immune response to carry out its function, healthy $CD4^+$ and $CD8^+$ T lymphocytes are required. Activated $CD4^+$ helper T cells release various cytokines (*i.e.,* IFN-γ) that recruit and activate other $CD8^+$ cytotoxic T cells, macrophages, and inflammatory cells. Cytokines (*e.g.,* chemokines) stimulate migration of several types of inflammatory cells, including macrophages, neutrophils, and basophils,

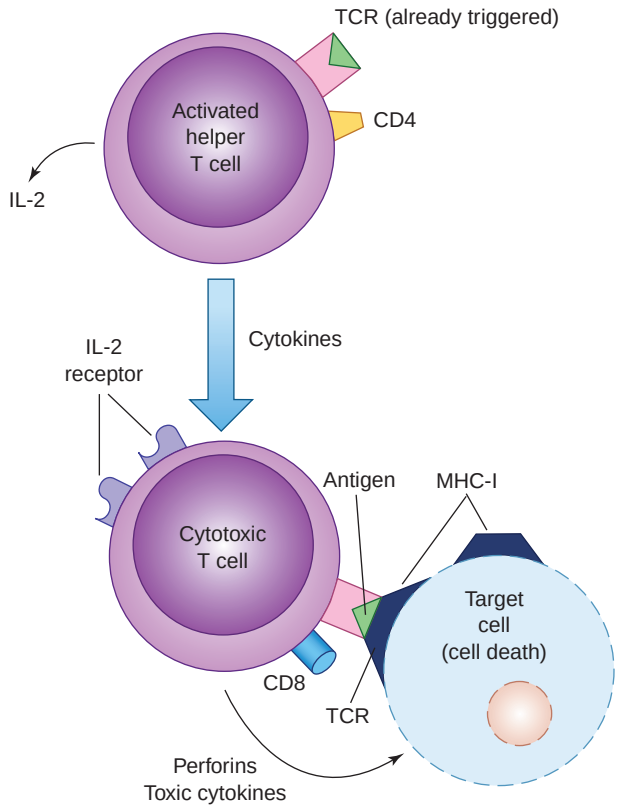


FIGURE 13.11 • Destruction of target cell by cytotoxic T cell. Cytokines released from the activated helper T cell enhance the potential of the cytotoxic T cell in destruction of the target cell.

which further enhances the phagocytic, metabolic, and enzymatic functions of the cell-mediated immune response. This results in a more rapid and more efficient destruction of infected cells. This type of defense is important against many intracellular pathogens such as *Mycobacterium* species and *Listeria monocytogenes* but unfortunately plays a role in delayed hypersensitivity reactions. Allergic contact dermatitis (delayed hypersensitivity type IV) results from the activation of both $CD4^+$ and $CD8^+$ T-cell precursors in the lymph nodes draining the site of antigen presentation. These "haptenated peptides" stimulate the recruitment of T cells at the site of antigen presentation, inducing inflammatory signals and apoptosis of epidermal cells, leading to the development of inflammation, to the release of chemical mediators, and to clinical symptoms.

In cell-mediated immune responses, the actions of T lymphocytes and effector macrophages predominate. The most aggressive and abundant phagocyte, the macrophage, becomes activated after exposure to T-cell cytokines, especially IFN-γ.[23] The initial stages of cell-mediated immunity are initiated when an APC displays an antigen peptide–class I or II MHC complex to the $CD4^+$ helper T cell and activates it. The activated helper T cell then synthesizes IL-2, IL-4, and other cytokines, which stimulate increased production of $CD4^+$ helper T cells and then amplify the response. Additional cytokine release enhances the activity of cytotoxic T cells and effector macrophages.

Lymphoid Organs

The central and peripheral lymphoid organs are responsible for the production, maturation, and storage of large numbers of immune system cells including the B and T lymphocytes. These organs and tissues are widely distributed throughout the body and provide different, but often overlapping, functions (Fig. 13.12). The central lymphoid organs are comprised of the bone marrow and the thymus and are responsible for immune cell production and maturation. The tissues and cells of the peripheral lymphoid system store the cells of the immune system where they function to concentrate and process antigen as well as support cellular processes necessary for development of fully functioning, adaptive immune responses. The peripheral lymphoid tissues are comprised of the lymph nodes, spleen, tonsils, appendix, Peyer patches in the intestine, and mucosa-associated lymphoid tissues in the respiratory, gastrointestinal, and reproductive systems. Networks of lymph channels, blood vessels, and capillaries connect the lymphoid organs and transport immune cells, antigens, and cellular debris throughout the body.

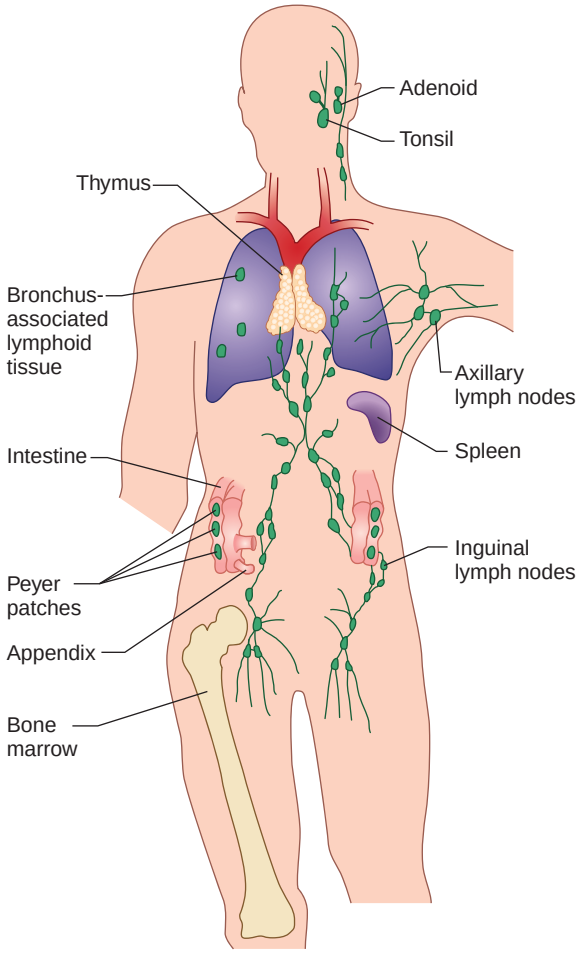


FIGURE 13.12 • Central and peripheral lymphoid organs and tissues.

Thymus

The thymus is an elongated, bilobed structure located in the mediastinum above the heart and serves as a specialized immune system organ. Each lobe is surrounded by a connective tissue capsule layer and is divided into lobules. The lobules can be divided into an outer cortex and a central medulla, which play different roles in the process of T-lymphocyte maturation. The outer cortex contains densely packed immature T lymphocytes (thymocytes). The inner medulla is a less dense area of tissue that contains fewer but more histologically mature lymphocytes. The medulla is comprised of Hassall corpuscles but also stores DCs and macrophages (Fig. 13.13).

The thymus is essential to the development of the immune system because it is responsible for the production of mature, immunocompetent T lymphocytes. The thymus is a fully developed organ at birth, weighing approximately 15 to 20 g. It is most active in the neonatal and preadolescent periods. At puberty, when the immune cells are well established in peripheral lymphoid tissues, the thymus begins to atrophy and is replaced by adipose tissue. Nevertheless, residual T-lymphocyte production continues throughout adult life. Precursor T (pre-T) cells enter the thymus as functionally and phenotypically immature T cells. They then mature during different cycles and then move from the cortex to the medulla until they are released into the peripheral lymphoid tissues. Rapid cell division, maturation, and selection occur

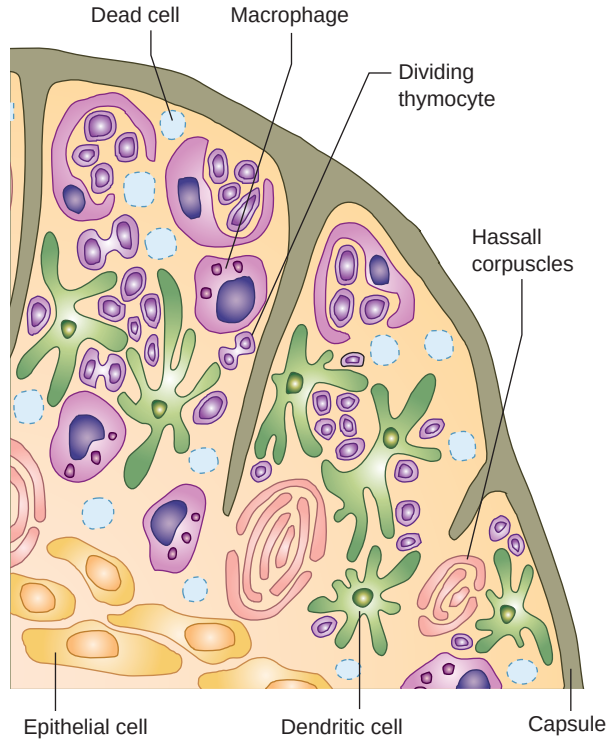


FIGURE 13.13 • Structural features of the thymus gland. The thymus gland is divided into lobules containing an outer cortex densely packed with dividing thymocytes or premature T cells and an inner medulla that contains mature T lymphocytes, macrophages, dendritic cells, and Hassall corpuscles.

in the cortex under the influence of thymic hormones and cytokines. As the T cells mature, they develop the TCRs that differentiate them from other types of T cells. The majority of the thymocytes die in the cortex during the process of gene rearrangement and maturation because they fail to develop the appropriate receptor types on their cell membranes. Only those T cells capable of recognizing foreign antigen displayed by self-MHC are allowed to mature. This process is called *thymic selection*. Mature, immunocompetent T-helper and T-cytotoxic cells leave the thymus in 2 to 3 days and enter the peripheral lymphoid tissues through the bloodstream.

Lymph Nodes

Lymph nodes are small aggregates of lymphoid tissue located along lymphatic vessels throughout the body. The lymphatic vessels carry lymph, which is a clear sometimes yellow-tinged fluid that contains a variety of white blood cells (predominantly lymphocytes) and transports cellular debris and organisms to the lymph modes to be removed from the body. Each lymph node processes lymph from a discrete, adjacent anatomic site. Lymph nodes are congregated in the axillae and groin and along the great vessels of the neck, thorax, and abdomen. The lymph nodes receive lymph from the collecting ducts, which ultimately drain into the thoracic duct located in the left side of the chest at the level of the subclavian vein. Lymph nodes have two functions: removal of foreign material from lymph before it enters the bloodstream and serving as centers for proliferation and response of immune cells.

Lymph nodes are bean-shaped, encapsulated tissues, approximately 0.5 to 1 cm in diameter. Lymph enters the node through afferent lymph channels and leaves through the efferent lymph vessels located in the deep indentation of the hilus. Lymphocytes and macrophages move slowly through the lymph nodes so that they have adequate time to engulf microorganisms and interact with circulating antigen. The lymphatic system provides a large surface upon which macrophages and DCs can more easily present antigens to T lymphocytes.

Lymph nodes are divided into three distinct and specialized areas—an outer cortex, a paracortex, and an inner medulla (Fig. 13.14). The T lymphocytes predominate in the paracortex and the B lymphocytes predominate in the follicles and germinal centers of the outer cortex. The T lymphocytes proliferate when antigens enter the paracortex of the lymph node. They then migrate to the outer cortex so that they can interact with B lymphocytes that are stored there. Within the follicles the lymphocytes continue to mature, replicate, and interact with the PACs present in the nodes (macrophages and follicular DCs). Activated B cells then migrate to the medulla of the lymph node, where they complete their maturation into plasma cells. Large quantities of antibodies are then released into the systemic circulation.

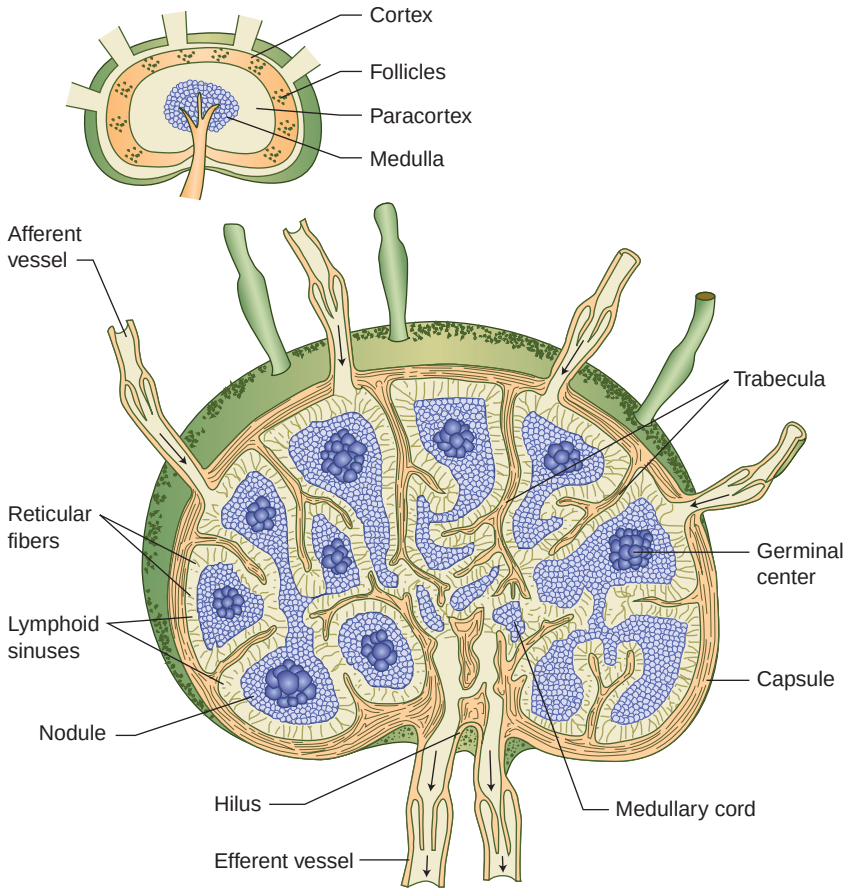


FIGURE 13.14 • Structural features of a lymph node. Bacteria that gain entry to the body are filtered out of the lymph as it flows through the node.

Spleen

The spleen is a large, ovoid secondary lymphoid organ located high in the left upper quadrant of the abdominal cavity between the diaphragm and the stomach. The spleen filters antigens from the blood and is important in the response to systemic infections. It is divided into two systems: the white pulp and the red pulp. The red pulp is well supplied with arteries and venous sinusoids and is the area where senescent and injured red blood cells are removed. The white pulp contains lymphatic nodules and diffuse lymphoid tissue where concentrated areas of B and T lymphocytes permeated by macrophages and DCs exist. The lymphocytes (primarily T cells) that surround the central arterioles form the area called the *periarterial lymphoid sheath*. There is also a diffuse marginal zone that contains the follicles and germinal centers and is rich in B cells. This separates the white pulp from the red pulp and allows lymphocytes to move easily between the blood and the lymphatic tissue. A sequence of activation events similar to that seen in the lymph nodes occurs in the spleen.

Other Secondary Lymphoid Tissues

Other secondary lymphoid tissues include the *mucosa-associated lymphoid tissues*, which are nonencapsulated clusters of lymphoid tissues located around membranes lining the respiratory, digestive, and urogenital tracts. These organ systems constantly came in contact with pathogens and toxins and, therefore, require the presence of immune cells in order to respond to the potential invasion by pathogens and harmful substances. In some tissues, the lymphocytes are organized in loose, nondescript clusters, but in other tissues such as the tonsils, Peyer patches in the intestine, and the appendix, their structure is better organized. These tissues contain all the cellular components (*i.e.,* T cells, B cells, macrophages, and DCs) required to mount an immune response. Immunity at the mucosal layers helps to exclude many pathogens from the body and, as a result, protects the more vital internal structures.

Active versus Passive Immunity

The goal of the immune system is to protect the host against invasion by potentially dangerous pathogens, foreign substances, and other sources of harmful antigens. Adaptive immune responses accomplish this goal through the activation of cell-mediated and humoral responses. This type of protection can be induced in one of two ways:

1. After exposure to the offending substance and activation of B and T lymphocytes (active immunity)
2. Through the transfer of antibodies against an antigen directly to the host (passive immunity)

Active immunity is acquired when the host mounts an immune response to an antigen either through the process of vaccination or from environmental exposure. It is called *active immunity* because it requires the host's own immune system to develop an immunological response including the development of memory. Active immunity is usually long lasting but requires a few days to weeks after a first exposure to sufficiently develop an appropriate immunological response that culminates in the destruction of the presenting antigen. However, with subsequent exposure the immune system rapidly becomes fully activated because of the presence of memory B and T lymphocytes and circulating antibodies. The process by which active immunity is acquired through the administration of a vaccine is termed *immunization*. An acquired immune response can improve on repeated exposures to an injected antigen (booster vaccines) or a natural infection.

Passive immunity is immunity transferred from another source. The most common form of passive immunity is that conferred from mother to fetus. During fetal development, maternal IgG antibodies are transferred to the fetus via the placenta. After birth, the neonate also receives IgG antibodies from the mother in breast milk or colostrum. Therefore, infants are provided with some degree of protection from infection for approximately 3 to 6 months, giving their own immune systems time to mature. Some protection against infectious disease can also be provided by the administration of Igs pooled from human or animal sources. Passive immunity produces only short-term protection that lasts weeks to months.

Regulation of the Adaptive Immune Response

In order for a host organism to remain healthy, the immune system must function properly. A weakened immune response may lead to immunodeficiency, but an inappropriate or excessive response can cause allergic reactions and autoimmune diseases. Therefore, the immune system must be capable of regulating itself. The process by which the body regulates itself is poorly understood but must involve all aspects of the innate and adaptive immune responses.

Each exposure to an antigen elicits a predictable response from the immune system. Once the immune system is activated, the response is amplified until it peaks and eventually subsides. This occurs because the body's normal immune responses are self-limiting. Once the antigen is destroyed and the action of chemical mediators terminated, the immune response ceases. It is believed that anti-inflammatory cytokines and regulatory T lymphocytes play a role in this process.[34]

Tolerance also plays a role in the self-regulation of the immune response. Tolerance is the ability of the immune system to react to foreign antigens but remain nonreactive to self-antigens. Tolerance to self-antigens protects the body from harmful autoimmune responses. This is exquisitely important in vital organs such as the brain, testes, ovaries, and eyes where immunological damage could be lethal to the organism.

Many autoimmune diseases such as Hashimoto thyroiditis and insulin-dependent diabetes mellitus are caused by impairment in both B and T lymphocyte (specifically cytotoxic lymphocytes) functions resulting in direct cellular damage because the body immune system is no longer capable of distinguishing "self" from "nonself."[35,36]

Inflammation

Sheila Grossman

THE INFLAMMATORY RESPONSE

Acute Inflammation

Cells of Inflammation
Vascular Stage
Cellular Stage
Inflammatory Mediators
Local Manifestations

Chronic Inflammation

Nonspecific Chronic Inflammation
Granulomatous Inflammation

Systemic Manifestations of Inflammation

Acute-Phase Response
White Blood Cell Response
Lymphadenitis

Inflammation is a response intended to eliminate the initial cause of cell injury, remove the damaged tissue, and generate new tissue. It accomplishes this by destroying, enzymatically digesting, walling off, or otherwise neutralizing the harmful agents such as toxins, foreign agents, or infectious organisms.[1] These processes set the stage for the events that will eventually heal the damaged tissue. Thus, inflammation is intimately interwoven with the repair processes that replace damaged tissue or fill in the residual defects with fibrous scar tissue.

Although first described over 2000 years ago, the inflammatory response has evoked renewed interest during the past several years. As a result, the pathogeneses of multiple diseases are known to be linked to the inflammatory response.[1] In these cases, the inflammatory cascade is triggered to be overly zealous to the point of damaging multiple types of human tissue with autoimmune disorders, such as in rheumatoid arthritis.[2] This chapter focuses on the morphologic and functional manifestations of acute and chronic inflammation, tissue repair, and wound healing.

THE INFLAMMATORY RESPONSE

After completing this section of the chapter, you should be able to meet the following objectives:

- Describe the vascular changes in an acute inflammatory response.
- Characterize the interaction of adhesion molecules, chemokines, and cytokines in leukocyte adhesion, migration, and phagocytosis, which are part of the cellular phase of inflammation.
- List four types of inflammatory mediators and state their function.

Inflammation is the reaction of vascularized tissues to injury. It is characterized by inflammatory mediators, such as complement, tumor necrosis factor alpha, vascular endothelial growth factor (VEGF), neutrophils, and serum amyloid, and the movement of fluid. Inflammation generally localizes and eliminates microbes, foreign particles, and abnormal cells and paves the way for repair of the injured tissue. Inflammatory conditions are commonly named by adding the suffix *-itis* to the affected organ or system. For example, *appendicitis* refers to inflammation of the appendix, *pericarditis* to inflammation of the pericardium, and *neuritis* to inflammation of a nerve. More descriptive expressions of the inflammatory process might indicate whether the process was acute or chronic and what type of exudate was formed.

The classic description of inflammation has been handed down through the ages. In the first century AD, the local reaction of injury was described in terms that are now known as the *cardinal signs* of inflammation.[1] These signs are *rubor* (redness), *tumor* (swelling), *calor* (heat), and *dolor* (pain). In the second century AD, a fifth cardinal sign, *functio laesa* (loss of function) was added. In addition to the cardinal signs that appear at the site of injury, systemic or constitutional manifestations (*e.g.*, fever) may occur as chemical mediators (*e.g.*, cytokines) produced at the site of inflammation gain entrance to the circulatory system. The constellation of systemic manifestations that may occur during acute inflammation is known as the *acute-phase response*.

The degree of the inflammatory response is impacted by multiple factors, such as the duration of the insult, the type of foreign agent, the degree of injury, and the microenvironment.[1] Inflammation can be divided into acute and chronic types.[1] *Acute inflammation* is of relatively short duration, lasting from a few minutes to several days, and is characterized by the exudation of fluid and plasma components and emigration of leukocytes, predominantly neutrophils, into the extravascular tissues. *Chronic inflammation* is of a longer duration, lasting for days to years, and is associated with the presence of lymphocytes and macrophages, proliferation of blood vessels, fibrosis, and tissue necrosis. These basic forms of inflammation often overlap, and many factors may influence their course.

Acute Inflammation

Acute inflammation is the early (almost immediate) reaction of local tissues and their blood vessels to injury. It typically occurs before adaptive immunity becomes established and is aimed primarily at removing the injurious agent and limiting the extent of tissue damage. Acute inflammation can be triggered by a variety of stimuli, including infections, immune reactions, blunt and penetrating trauma, physical or chemical agents (*e.g.,* burns, frostbite, irradiation, caustic chemicals), and tissue necrosis from any cause.

Cells of Inflammation

Acute inflammation involves two major components: the vascular and cellular stages.[1–3] Many tissues and cells are involved in these reactions, including the endothelial cells that line blood vessels, circulating white blood cells, connective tissue cells (mast cells, fibroblasts, tissue macrophages, and lymphocytes), and components of the extracellular matrix (ECM) (Fig. 14.1). The ECM consists of fibrous proteins (collagen and elastin), adhesive glycoproteins, and proteoglycans. At the biochemical level, the inflammatory mediators, acting together or in sequence, amplify the initial response and influence its evolution by regulating the subsequent vascular and cellular responses.

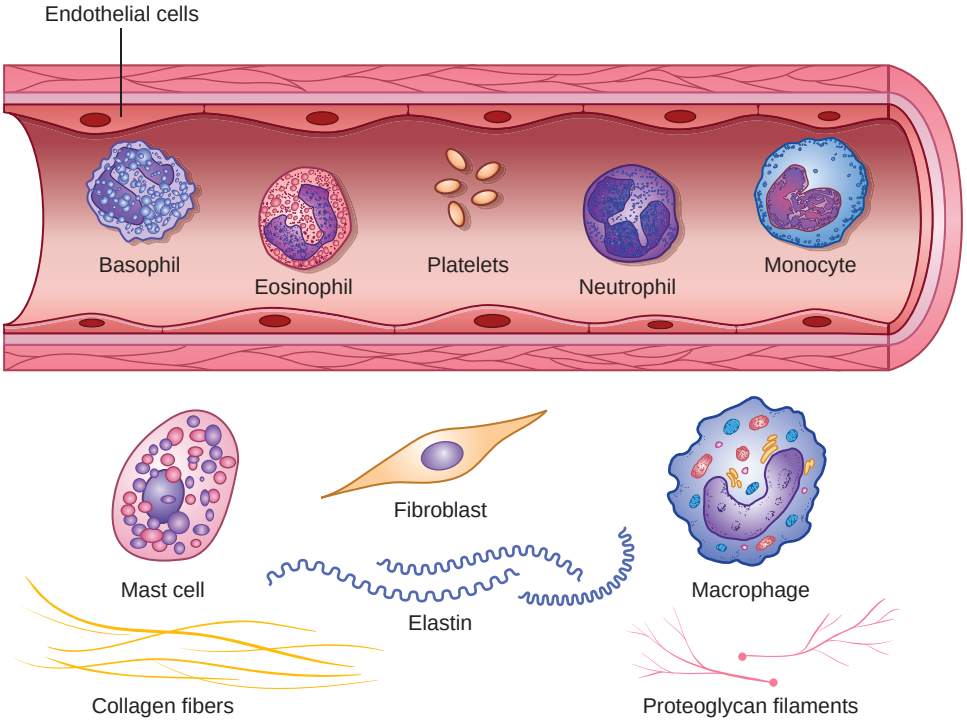


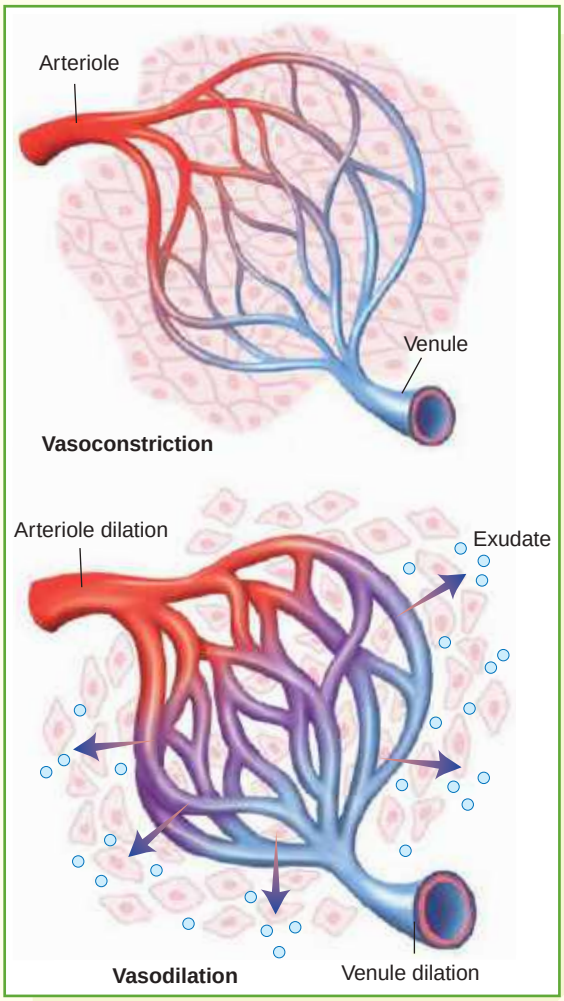
FIGURE 14.1 • Cells of acute inflammation.

Understanding Acute Inflammation

Acute inflammation is the immediate and early response to an injurious agent. The response, which serves to control and eliminate altered cells, microorganisms, and antigens, occurs in two phases: (1) the vascular phase, which leads to an increase in blood flow and changes in the small blood vessels of the microcirculation, and (2) the cellular phase, which leads to the migration of leukocytes from the circulation and their activation to eliminate the injurious agent. The primary function of inflammatory response is to limit the injurious effect of the pathologic agent and remove the injured tissue components, thereby allowing tissue repair to take place.

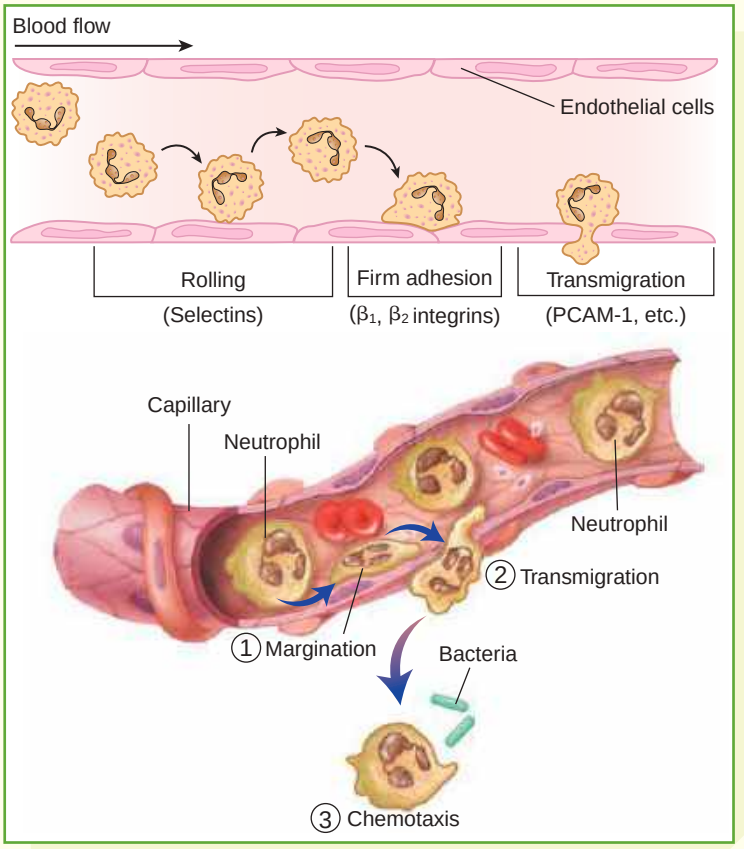
Vascular Phase

The vascular phase of acute inflammation is characterized by changes in the small blood vessels at the site of injury. It begins with momentary vasoconstriction followed rapidly by vasodilation. Vasodilation involves the arterioles and venules with a resultant increase in capillary blood flow, causing heat and redness, which are two of the cardinal signs of inflammation. This is accompanied by an increase in vascular permeability with outpouring of protein-rich fluid (exudate) into the extravascular spaces. The loss of proteins reduces the capillary osmotic pressure and increases the interstitial osmotic pressure. This, coupled with an increase in capillary pressure, causes a marked outflow of fluid and its accumulation in the tissue spaces, producing the swelling, pain, and impaired function that represent the other cardinal signs of acute inflammation. As fluid moves out of the vessels, stagnation of flow and clotting of blood occur. This aids in localizing the spread of infectious microorganisms.



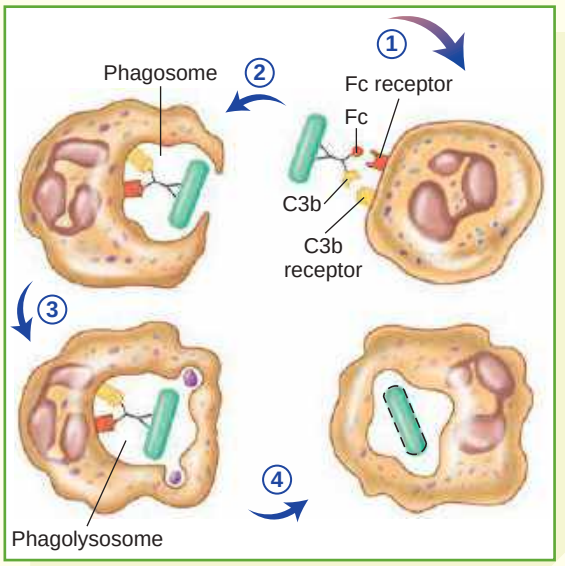
Cellular Phase: Leukocyte Margination, Adhesion, and Transmigration

The cellular phase of acute inflammation involves the delivery of leukocytes, mainly neutrophils, to the site of injury so they can perform their normal functions of host defense. The delivery and activation of leukocytes can be divided into the following steps: adhesion and margination, transmigration, and chemotaxis. The recruitment of leukocytes to the precapillary venules, where they exit the circulation, is facilitated by the slowing of blood flow and margination along the vessel surface. Leukocyte adhesion and transmigration from the vascular space into the extravascular tissue is facilitated by complementary adhesion molecules (*e.g.,* selectins, integrins) on the leukocyte and endothelial surfaces. After extravasation, leukocytes migrate in the tissues toward the site of injury by chemotaxis or locomotion oriented along a chemical gradient.



Leukocyte Activation and Phagocytosis

Once at the sight of injury, the products generated by tissue injury trigger a number of leukocyte responses, including phagocytosis and cell killing. Opsonization of microbes (1) by complement factor C3b and antibody facilitates recognition by neutrophil C3b and the antibody Fc receptor. Receptor activation (2) triggers intracellular signaling and actin assembly in the neutrophil, leading to formation of pseudopods that enclose the microbe within a phagosome. The phagosome (3) then fuses with an intracellular lysosome to form a phagolysosome into which lysosomal enzymes and oxygen radicals (4) are released to kill and degrade the microbe.



Endothelial Cells. Endothelial cells constitute the single cell–thick epithelial lining of blood vessels.[1,4,5] They produce antiplatelet and antithrombotic agents that maintain vessel patency and vasodilators and vasoconstrictors that regulate blood flow. Endothelial cells are also key players in the inflammatory response and experience significant pathology in people with inflammatory disorders. Functioning endothelial cells provide a selective permeability barrier to exogenous (microbial) and endogenous inflammatory stimuli, regulate leukocyte extravasation by expression of cell adhesion molecules and receptors, contribute to the regulation and modulation of immune responses through synthesis and release of inflammatory mediators, and regulate immune cell proliferation through secretion of hematopoietic colony-stimulating factors (CSFs). Endothelial cells also participate in the repair process that accompanies inflammation through the production of growth factors that stimulate angiogenesis (formation of new blood vessels) and ECM synthesis.[3–5] Circulating endothelial cells can be used as a trend indicator of vascular dysfunction in people who have systemic lupus erythematosus (SLE), even in people with SLE who have no diagnosed cardiovascular disease.[6]

Platelets. Platelets or thrombocytes are the cell fragments circulating in the blood that are involved in the cellular mechanisms of primary hemostasis. Activated platelets also release a number of potent inflammatory mediators, thereby increasing vascular permeability and altering the chemotactic, adhesive, and proteolytic properties of the endothelial cells.[7,8] When a platelet undergoes activation, over 300 proteins are released. Although only a relatively small proportion of these have been identified, it appears that a significant number are inflammatory mediators.[7] The association between platelets and inflammatory diseases is highlighted by the number of inflammatory disease processes (*e.g.,* atherosclerosis, migraine headache, SLE) shown to be associated with platelet activation.[7]

Neutrophils and Monocytes/Macrophages. The neutrophils and macrophages are phagocytic leukocytes that are present in large numbers and are evident within hours at the site of inflammation. Both types of leukocytes express a number of surface receptors and molecules that are involved in their activation. They include mannose receptors that bind glycoproteins of bacteria; toll-like receptors that respond to different types and components of microbes; cell communication receptors that recognize specific cytokines and chemokines produced in response to infections and tissue injury; cell adhesion molecules that affect leukocyte adhesion; and complement receptors that recognize degraded fragments of complement deposited on the microbial surface (Fig. 14.2).

The *neutrophil* is the primary phagocyte that arrives early at the site of inflammation, usually within 90 minutes of injury.[1] These leukocytes have nuclei that are divided into three to five lobes. Therefore, they often are referred to as *polymorphonuclear neutrophils (PMNs)* or *segmented neutrophils (segs)*. A white blood cell identified by distinctive cytoplasmic granules is called a granulocyte. The cytoplasmic granules of the granulocytes, which resist staining and remain a neutral color, contain enzymes and antibacterial material that are used in destroying engulfed microbes and dead tissue.[9] Neutrophils are able to generate oxygen (hydrogen peroxide) and nitrogen products (nitric oxide [NO]) that assist in destroying the engulfed debris.[9]

The neutrophil count in the blood often increases greatly during an inflammatory process, especially with bacterial infections. After being released from the bone marrow, circulating neutrophils have a life span of only approximately 10 hours and therefore must be constantly replaced if their numbers are to remain adequate. This requires an increase in circulating white blood cells, a condition called *leukocytosis*, which is frequently elevated with bacterial infections and tissue injury.[1] With excessive demand for phagocytes, immature forms of neutrophils are released from the bone marrow. These immature cells often are called *bands* because of the horseshoe shape of their nuclei.

Circulating monocytes, which have a single kidney-shaped nucleus and are the largest of the circulating leukocytes, constitute 3% to 8% of the white blood cell count. The monocytes are released from the bone marrow to act as macrophages.[1,9] Mononuclear cells arrive at the inflammatory site shortly after the neutrophils and perform their phagocytic functions for several days.[1]

Monocytes and macrophages produce potent vasoactive mediators, including prostaglandins and leukotrienes, platelet-activating factor (PAF), inflammatory cytokines, and growth factors that promote regeneration of tissues. The macrophages engulf larger and greater quantities of foreign material than the neutrophils. These longer-lived phagocytes help to destroy the causative agent, aid in the signaling processes of immunity, serve to resolve the inflammatory process, and contribute to initiation of the healing processes. They also play an important role in chronic inflammation, where they can surround and wall off foreign material that cannot be digested.

Eosinophils, Basophils, and Mast Cells. Eosinophils, basophils, and mast cells produce lipid mediators and cytokines that induce inflammation. Although all three-cell types have unique characteristics, they all contain cytoplasmic granules that induce inflammation. They are particularly important in inflammation associated with immediate hypersensitivity reactions and allergic disorders.

Eosinophils circulate in the blood and are recruited to tissues, similar to neutrophils. These granulocytes increase in the blood during allergic reactions and parasitic infections. The granules of eosinophils, which stain red with the acid dye eosin, contain a protein that is highly toxic to large parasitic worms that cannot be phagocytized. They also play an important role in allergic reactions by controlling the release of specific chemical mediators.

Basophils are blood granulocytes with structural and functional similarities to mast cells of the connective tissue.

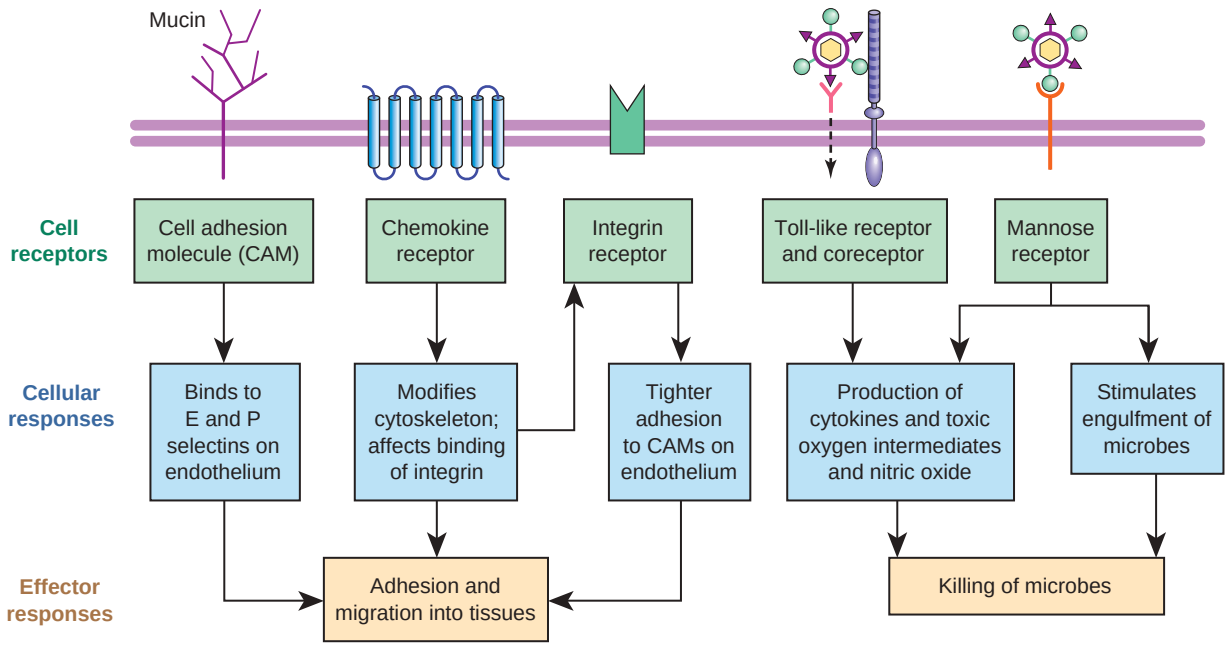


FIGURE 14.2 • Leukocyte activation. Different classes of leukocyte cell surface receptors recognize different stimuli. The receptors initiate responses that mediate the functions of the leukocytes.

They are derived from bone marrow progenitors and circulate in blood. The granules of the *basophils*, which stain blue with a basic dye, contain histamine and other bioactive mediators of inflammation. Both basophils and mast cells bind an antibody, immunoglobulin E (IgE), secreted by plasma cells through receptors on their cell surface.[10] Binding of IgE triggers release of histamine and vasoactive agents from the basophil granules.

Mast cells derive from the same hematopoietic stem cells as basophils but do not develop until they leave the circulation and lodge in tissue sites. Activation of mast cells results in release of preformed contents of their granules (histamine, proteoglycans, proteases, and cytokines such as tumor necrosis factor-α [TNF-α] and interleukin [IL]-16), synthesis of lipid mediators derived from cell membrane precursors (arachidonic acid metabolites, such as prostaglandins, and PAF), and stimulation of cytokine and chemokine synthesis by other inflammatory cells such as monocytes and macrophages. Mast cells are involved in IgE-triggered reactions and with helminth infections.[11]

Vascular Stage

The vascular changes that occur with inflammation involve the arterioles, capillaries, and venules of the microcirculation. These changes begin soon after injury and are characterized by vasodilation, changes in blood flow, increased vascular permeability, and leakage of fluid into the extravascular tissues.[1]

Vasodilation, which is one of the earliest manifestations of inflammation, follows a transient constriction of the arterioles, lasting a few seconds. Vasodilation first involves the arterioles and then results in opening of capillary beds in the area. As a result, the area becomes congested, causing the redness (erythema) and warmth associated with acute inflammation. Vasodilation is induced by the action of several mediators, such as histamine and NO.

Vasodilation is quickly followed by increased permeability of the microvasculature, with the outpouring of a protein-rich fluid (exudate) into the extravascular spaces. The loss of fluid results in an increased concentration of blood constituents (red cells, leukocytes, platelets, and clotting factors), stagnation of flow, and clotting of blood at the site of injury. This aids in localizing the spread of infectious microorganisms. The loss of plasma proteins reduces the intracapillary osmotic pressure and increases the osmotic pressure of the interstitial fluid, causing fluid to move into the tissues and produce the swelling (*i.e.,* edema), pain, and impaired function that are the cardinal signs of acute inflammation. The exudation of fluid into the tissue spaces also serves to dilute the offending agent.

The increased permeability characteristic of acute inflammation results from formation of endothelial gaps in the venules of the microcirculation. Binding of the chemical mediators to endothelial receptors causes contraction of endothelial cells and separation of intercellular junctions. This is the most common mechanism of vascular leakage and is elicited by histamine, bradykinin, leukotrienes, and many other classes of chemical mediators.

Vascular Response Patterns. Depending on the severity of injury, the vascular changes that occur with inflammation follow one of three patterns of responses.[2] The first pattern is an *immediate transient response*, which occurs with minor injury. It develops rapidly after injury and is usually reversible and of short duration (15 to 30 minutes). Typically, this type of leakage affects venules 20 to 60 μm in diameter, leaving capillaries

and arterioles unaffected.[2] Although the precise mechanism for restriction of this effect to the venules is unknown, it may reflect the greater density of receptors in the endothelium of the venules. It has also been suggested that the later leukocyte events of inflammation (*i.e.,* adhesion and emigration) also occur predominantly in the venules of most organs.

The second pattern is *an immediate sustained response*, which occurs with more serious types of injury and continues for several days. It affects arterioles, capillaries, and venules and is generally due to direct damage of the endothelium. Neutrophils that adhere to the endothelium may also injure endothelial cells.

The third pattern is a *delayed hemodynamic response*, in which the increased permeability occurs in the venules and capillaries. A delayed response often accompanies injuries due to radiation, such as sunburn. The mechanism of the leakage is unknown, but it may result from the direct effect of the injurious agent, leading to delayed endothelial cell damage.

Cellular Stage

The cellular stage of acute inflammation is marked by changes in the endothelial cells lining the vasculature and movement of phagocytic leukocytes into the area of injury or infection. Although attention has been focused on the recruitment of leukocytes from the blood, a rapid response also requires the release of chemical mediators from tissue cells (mast cells and macrophages) that are prepositioned in the tissues. The sequence of events in the cellular response to inflammation includes leukocyte

1. Margination and adhesion to the endothelium
2. Transmigration across the endothelium
3. Chemotaxis
4. Activation and phagocytosis[1,3]

Margination, Adhesion, and Transmigration. During the early stages of the inflammatory response, the leukocytes are concentrated along the endothelium wall. Cross-talk between the blood leukocytes and the vascular endothelium defines a definite inflammatory event and ensures secure adhesion and arrest of the leukocytes along the endothelium.[12] As a consequence, the leukocytes slow their migration, adhere tightly to the endothelium, and begin to move along the periphery of the blood vessels. This process of leukocyte accumulation is called *margination*. The subsequent release of cell communication molecules called **cytokines** causes the endothelial cells lining the vessels to express cell adhesion molecules, such as *selectins*, that bind to carbohydrates on the leukocytes.[10] This interaction slows their flow and causes the leukocytes to move along the endothelial cell surface with a rolling movement, finally coming to rest and adhering strongly to intercellular adhesion molecules (ICAMs), thus, attaching on the endothelium.[1,3,10] The adhesion causes the endothelial cells to separate, allowing the leukocytes to extend pseudopodia and *transmigrate* through the vessel wall and then, under the influence of chemotactic factors, migrate into the tissue spaces.

Several families of adhesion molecules, including selectins, integrins (VLA-5), and the immunoglobulin superfamily, are involved in leukocyte recruitment.[12–14] The selectins are a family of three closely related proteins (P-selectin, E-selectin, and L-selectin) that differ in their cellular distribution but all function in adhesion of leukocytes to endothelial cells. The integrin superfamily consists of 30 structurally similar proteins that promote cell-to-cell and cell-to-ECM interactions. The name *integrin* derives from the hypothesis that they coordinate (integrate) signals of extracellular ligands with cytoskeleton-dependent motility, shape change, and phagocytic responses of immune cells. Adhesion molecules of the immunoglobulin superfamily include ICAM-1, ICAM-2, and vascular adhesion molecule (VCAM)-1, all of which interact with integrins on leukocytes to mediate their recruitment.

Chemotaxis. Chemotaxis is the dynamic and energy-directed process of directed cell migration.[1] Once leukocytes exit the capillary, they wander through the tissue guided by a gradient of secreted chemoattractants, such as chemokines, bacterial and cellular debris, and protein fragments generated from activation of the complement system (*e.g.,* C3a, C5a). Chemokines, an important subgroup of chemotactic cytokines, are small proteins that direct the trafficking of leukocytes during the early stages of inflammation or injury.[15] Several immune (*e.g.,* macrophages) and nonimmune cells secrete these chemoattractants to ensure the directed movement of leukocytes to the site of infection.

Leukocyte Activation and Phagocytosis. During the final stage of the cellular response, monocytes, neutrophils, and tissue macrophages are activated to engulf and degrade the bacteria and cellular debris in a process called phagocytosis.[1] Phagocytosis involves three distinct steps: (1) recognition and adherence, (2) engulfment, and (3) intracellular killing. Phagocytosis is initiated by the recognition and binding of particles by specific receptors on the surface of phagocytic cells. This binding is essential for trapping the agent, which triggers engulfment and activates the killing potential of the cell. Microbes can be bound directly to the membrane of phagocytic cells by several types of pattern recognition receptors (*e.g.,* toll-like and mannose receptors) or indirectly by receptors that recognize microbes coated with carbohydrate-binding lectins, antibody, or complement. The coating of an antigen with antibody or complement to enhance binding is called opsonization. Receptor-mediated endocytosis is triggered by opsonization and binding of the agent to phagocyte cell surface receptors. Endocytosis is accomplished through cytoplasmic extensions (pseudopods) that surround and enclose the particle in a membrane-bounded phagocytic vesicle or *phagosome*. Once inside the cell cytoplasm, the phagosome merges with a cytoplasmic lysosome containing antibacterial molecules and enzymes that can kill and digest the microbe.

Intracellular killing of pathogens is accomplished through several mechanisms, including toxic oxygen and

nitrogen products, lysozymes, proteases, and defensins. The metabolic burst pathways that generate toxic oxygen and nitrogen products (*i.e.,* NO, hydrogen peroxide, and hypochlorous acid) require oxygen and metabolic enzymes such as myeloperoxidase, nicotinamide adenine dinucleotide phosphate (NADPH) oxidase, and NO synthetase. Oxygen-independent pathways generate several types of digestive enzymes and antimicrobial molecules (*e.g.,* defensins). Individuals born with genetic defects in some of these enzymes have immunodeficiency conditions that make them susceptible to repeated bacterial infection.

Inflammatory Mediators

Although inflammation is precipitated by infection and injury, its signs and symptoms are produced by chemical mediators. Mediators can originate either from the plasma or from cells (Fig. 14.3). The plasma-derived mediators, which are synthesized in the liver, include the coagulation factors and the complement proteins. These mediators are present in the plasma in a precursor form that must be activated by a series of proteolytic processes to acquire their biologic properties. Cell-derived mediators are normally sequestered in intracellular granules that need to be secreted (*e.g.,* histamine from mast cells) or are newly synthesized (*e.g.,* cytokines) in response to a stimulus. Although the major sources of these mediators are platelets, neutrophils, monocytes/macrophages, and mast cells, endothelial cells, smooth muscle, fibroblasts, and most epithelial cells can be induced to produce some of the mediators.

The production of active mediators is triggered by microbes or host proteins, such as those of the complement, kinin, or coagulation systems, that are themselves activated by microbes or damaged tissues. Mediators can act on one or a few target cells, have diverse targets, or have differing effects on different types of cells. Once activated and released from the cell, most mediators are short-lived. They may be transformed into inactive metabolites, inactivated by enzymes, or otherwise scavenged or degraded.

Inflammatory mediators can be classified by function: (1) those with vasoactive and smooth muscle–constricting properties such as histamine, arachidonic acid metabolites (prostaglandins and leukotrienes), and PAF; (2) plasma proteases that activate members of the complement system, coagulation factors of the clotting cascade, and vasoactive peptides of the kinin system; (3) chemotactic factors such as complement fragments and chemokines; and (4) reactive molecules and cytokines liberated from leukocytes, which when released into the extracellular environment can affect the surrounding tissue and cells.

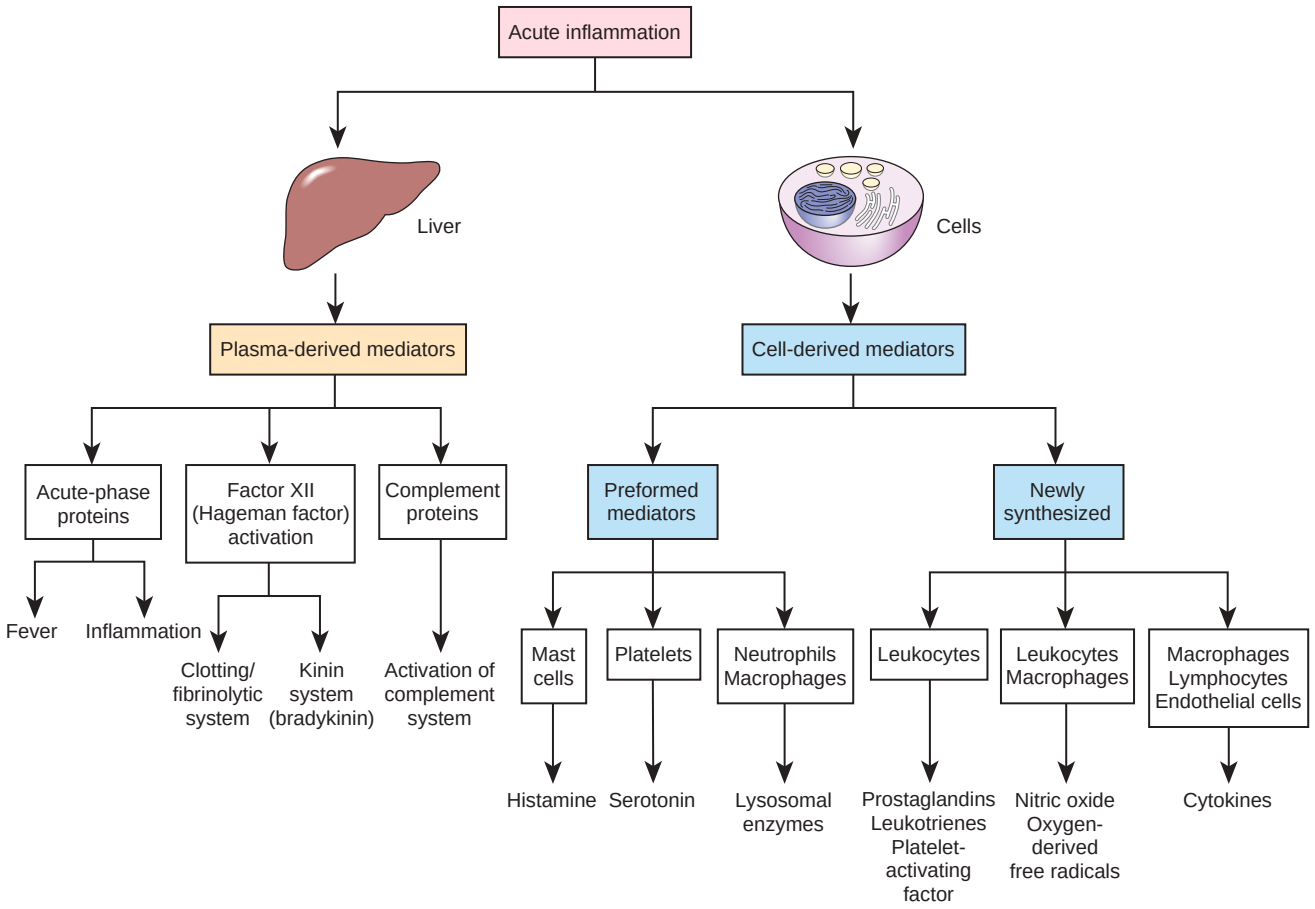


FIGURE 14.3 • Plasma- and cell-derived mediators of acute inflammation.

Histamine. Histamine is present in preformed stores in cells and is therefore among the first mediators to be released during an acute inflammatory reaction. Preformed histamine is widely distributed in tissues, the highest concentrations being found in the connective tissues adjacent to blood vessels. It is also found in circulating blood platelets and basophils. Preformed histamine is found in mast cell granules and is released in response to a variety of stimuli, including trauma and immune reactions involving binding of IgE antibodies. Histamine causes dilation of arterioles and increases the permeability of venules. It acts at the level of the microcirculation by binding to histamine type 1 (H_1) receptors on endothelial cells and is considered the principal mediator of the immediate transient phase of increased vascular permeability in the acute inflammatory response. Antihistamine drugs (H_1 receptor antagonists), which bind to the H_1 receptors, act competitively to antagonize many of the effects of the immediate inflammatory response.

Arachidonic Acid Metabolites. Arachidonic acid is a 20-carbon unsaturated fatty acid found in phospholipids of cell membranes. Release of arachidonic acid by phospholipases initiates a series of complex reactions that lead to the production of the *eicosanoid* family inflammatory mediators (prostaglandins, leukotrienes, and related metabolites). Eicosanoid synthesis follows one of two pathways: the cyclooxygenase pathway, which culminates in the synthesis of prostaglandins, and the lipoxygenase pathway, which culminates in the synthesis of the leukotrienes (Fig. 14.4).[16]

Through the cyclooxygenase metabolic pathway, many prostaglandins are synthesized from arachidonic acid.[10] The prostaglandins (*e.g.,* PGD_2, PGE_2, $PGF_{2\alpha}$, and PGI_2) induce inflammation and potentiate the effects of histamine and other inflammatory mediators. The prostaglandin thromboxane A_2 promotes platelet aggregation and vasoconstriction. Aspirin and the nonsteroidal anti-inflammatory drugs (NSAIDs) reduce inflammation by inactivating the first enzyme in the cyclooxygenase pathway for prostaglandin synthesis.

Like the prostaglandins, the leukotrienes are formed from arachidonic acid, but through the lipoxygenase pathway. Histamine and leukotrienes are complementary in action in that they have similar functions. Histamine is produced rapidly and transiently while the more potent leukotrienes are being synthesized. The leukotrienes also have been reported to affect the permeability of the postcapillary venules, the adhesion properties of endothelial cells, and the extravasation and chemotaxis of neutrophils, eosinophils, and monocytes. Leukotrienes (LT) C_4, LTD_4, and LTE_4, collectively known as the *slow-reacting substance of anaphylaxis* (SRS-A), cause slow and sustained constriction of the bronchioles and are important inflammatory mediators in bronchial asthma and anaphylaxis.

Dietary modification of the inflammatory response through the use of omega-3 polyunsaturated fatty acids, specifically eicosapentaenoic acid and docosahexaenoic acid, which are present in oily fish and fish oil may be effective in preventing some negative manifestations of inflammation.[16–18] Alpha-linolenic acid, which is present in flaxseed, canola oil, green leafy vegetables, walnuts, and soybeans, is another source of

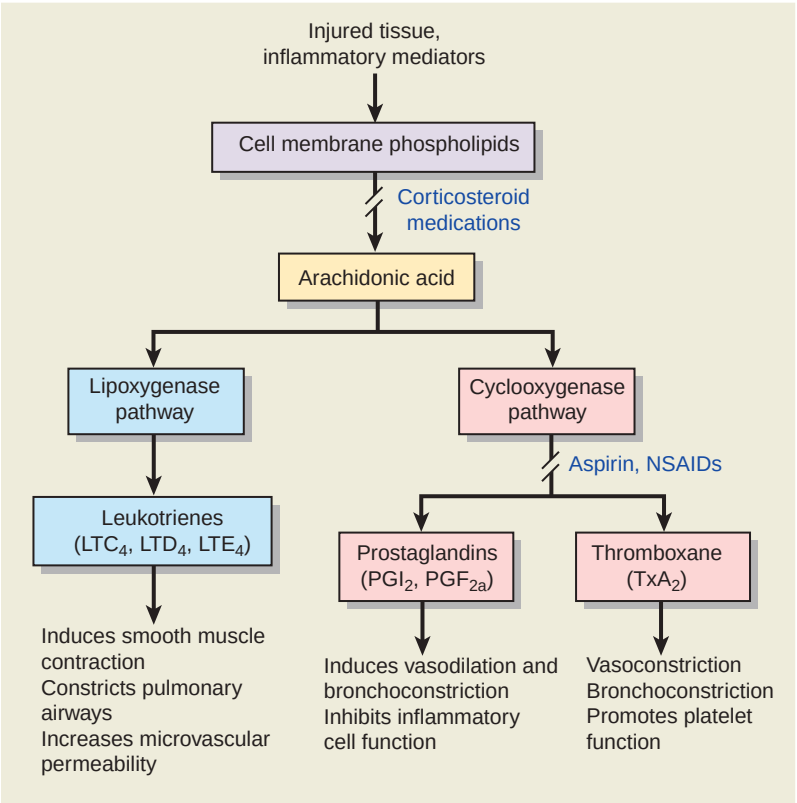


FIGURE 14.4 • The cyclooxygenase and lipoxygenase pathways and sites where the corticosteroids and nonsteroidal anti-inflammatory drugs (NSAIDs) exert their action.

omega-3 fatty acid. The omega-3 polyunsaturated fatty acids, which are considered antithrombotic and anti-inflammatory, are structurally different from the prothrombotic and proinflammatory omega-6 polyunsaturated fatty acids, which are present in most seeds, vegetable oils, and meats. Typically, the cell membranes of inflammatory cells contain high proportions of omega-6 arachidonic acid, which is the source of prostaglandin and leukotriene inflammatory mediators. Eating oily fish and other foods that are high in omega-3 fatty acids results in partial replacement of arachidonic acid in inflammatory cell membranes by eicosapentaenoic acid, a change that leads to decreased production of arachidonic acid–derived inflammatory mediators. This response alone is a potentially beneficial effect of omega-3 fatty acids. However, omega-3 fatty acids have a number of other effects that might occur downstream of altered eicosanoid production or might be independent of this function. For example, animal and human research has shown that dietary fish oil results in suppressed production of proinflammatory cytokines and decreased expression of adhesion molecules that participate in the inflammatory response.

Platelet-Activating Factor. PAF, which is generated from a complex lipid stored in cell membranes, affects a variety of cell types and induces platelet aggregation. It activates neutrophils and is a potent eosinophil chemoattractant. When injected into the skin, PAF causes a wheal-and-flare reaction and the leukocyte infiltrate characteristic of immediate hypersensitivity reactions. When inhaled, PAF causes bronchospasm, eosinophil infiltration, and nonspecific bronchial hyperreactivity.

Plasma Proteins. A number of phenomena in the inflammatory response are mediated by plasma proteins that belong to three interrelated systems, the clotting, complement, and kinin systems.

The clotting system contributes to the vascular phase of inflammation, mainly through fibrinopeptides that are formed during the final steps of the clotting process. The protease thrombin, which binds to receptors called *protease-activated receptors* (PARs), provides the final link between the coagulation system and inflammation.[19] Engagement of the so-called type 1 receptor (PAR-1) by proteases, particularly thrombin, triggers several responses that induce inflammation, including production of chemokines, expression of endothelial adhesion molecules, induction of prostaglandin synthesis, and production of PAF.

The complement system consists of 20 component proteins (and their cleavage products) that are found in greatest concentration in the plasma. The complement proteins are present in inactive forms in the plasma. Many of them are activated to become proteolytic enzymes that degrade other complement proteins, thus forming a cascade that plays an important role in both immunity and inflammation.[20–22] The complement proteins assist the inflammatory cascade by increasing vascular permeability, improving phagocytosis, and causing vasodilation.

The kinin system generates vasoactive peptides from plasma proteins called *kininogens*, by the action of proteases called *kallikreins*.[10] Activation of the kinin system results in release of bradykinin, which increases vascular permeability and causes contraction of smooth muscle, dilation of blood vessels, and pain when injected into the skin. These effects are similar to those of histamine. The action of bradykinin is short-lived, because it is quickly inactivated by an enzyme called *kininase*. Any bradykinin that escapes inactivation by the kininase enzyme is degraded by the angiotensin-converting enzyme in the lung.[10]

Cytokines and Chemokines. Cytokines are proteins produced by many cell types (principally activated macrophages and lymphocytes but also endothelium, epithelium, and connective tissue types) that modulate the function of other cells.[1,12,23] Although well known for their role in immune responses, these products also play important roles in both acute and chronic inflammation.

Tumor necrosis factor-α and IL-1 are two of the major cytokines that mediate inflammation. The major cellular source of TNF-α and IL-1 is activated macrophages (Fig. 14.5). IL-1 is also produced by many cell types other than macrophages, including neutrophils, endothelial cells, and epithelial cells

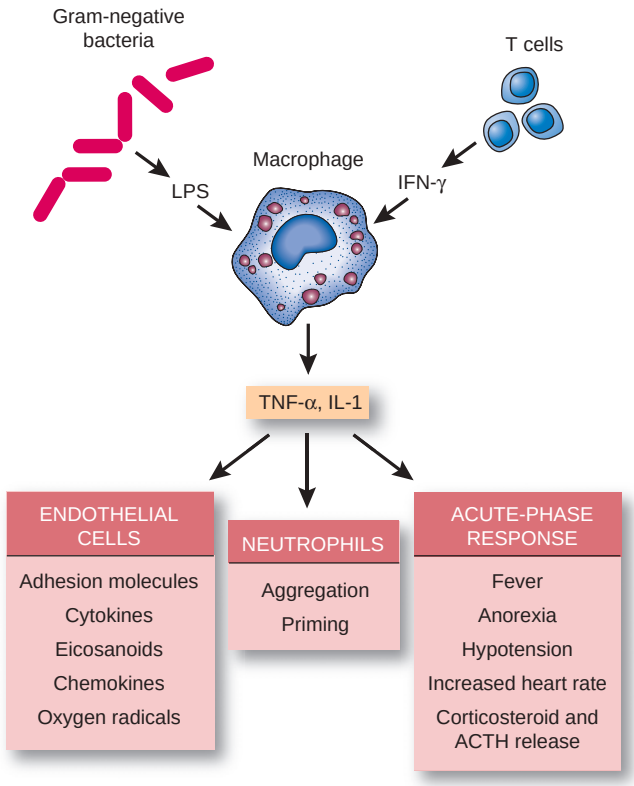


FIGURE 14.5 • Central role of interleukin (IL)-1 and tumor necrosis factor (TNF)-α in the acute inflammatory response. Lipopolysaccharide (LPS) and interferon (IFN)-γ activate macrophages to release inflammatory cytokines, principally IL-1 and TNF-α, responsible for directing both local and systemic inflammatory responses. (ACTH, adrenocorticotropic hormone.) (From Rubin R., Strayer D. E. (Eds.) (2012). *Rubin's pathology: Clinicopathologic foundations of medicine* (6th ed., p. 60). Philadelphia, PA: Lippincott Williams & Wilkins.)

(*e.g.,* keratinocytes). The secretion of TNF-α and IL-1 can be stimulated by endotoxin and other microbial products, immune cells, injury, and a variety of inflammatory stimuli. TNF-α and IL-1 induce endothelial cells to express adhesion molecules and release cytokines, chemokines, and reactive oxygen species. TNF-α induces priming and aggregation of neutrophils, leading to augmented responses of these cells to other mediators. IL-1 and TNF-α are also mediators of the acute-phase responses associated with infection or injury. Features of these systemic responses include fever, hypotension and increased heart rate, anorexia, release of neutrophils into the circulation, and increased levels of corticosteroid hormones.

Chemotactic cytokines, or *chemokines*, are a family of small proteins that act primarily as chemoattractants to recruit and direct the migration of immune and inflammatory cells.[24] Chemokines generate a chemotactic gradient by binding to proteoglycans on the surface of endothelial cells or in the ECM. As a result, high concentrations of chemokines persist at sites of tissue injury or infection. Two classes of chemokines have been identified: inflammatory chemokines and homing chemokines. Inflammatory chemokines are produced in response to bacterial toxins and inflammatory cytokines (*i.e.,* IL-1, TNF-α). These chemokines recruit leukocytes during an inflammatory response. Homing chemokines are constitutively expressed and are up-regulated during inflammatory reactions and immune responses.

Nitric Oxide- and Oxygen-Derived Free Radicals. NO- and oxygen-derived free radicals play an important role in the inflammatory response. NO, which is produced by a variety of cells, plays multiple roles in inflammation, including smooth muscle relaxation and antagonism of platelet adhesion, aggregation, and degranulation, and it serves as an endogenous regulator of leukocyte recruitment. Blocking of NO production can increase leukocyte adhesion, and delivery of exogenous NO reduces the number of leukocytes. Thus, production of NO appears to be a compensatory mechanism that reduces the cellular phase of inflammation. Impaired production of NO by vascular endothelial cells is implicated in the inflammatory changes that occur with atherosclerosis. NO and its derivatives also have antimicrobial actions, and thus, NO is also a host mediator against infection.

Oxygen free radicals may be released extracellularly from leukocytes after exposure to microbes, cytokines, and immune complexes or during the phagocytic process that occurs during the cellular phase of the inflammatory process. The superoxide radical, hydrogen peroxide, and the hydroxyl radical are the major species produced in the cell. These species can combine with NO to form other reactive nitrogen intermediates, which can increase the inflammatory process and cause more tissue injury.

Local Manifestations

Although all acute inflammatory reactions are characterized by vascular changes and leukocyte infiltration, the severity of the reaction, its specific cause, and the site of involvement introduce variations in its manifestations and clinical correlates. These manifestations can range from swelling and the formation of exudates to abscess formation or ulceration.

Characteristically, the acute inflammatory response involves the production of exudates. These exudates vary in terms of fluid type, plasma protein content, and the presence or absence of cells. They can be serous, hemorrhagic, fibrinous, membranous, or purulent. Often the exudate is composed of a combination of these types. *Serous exudates* are watery fluids low in protein content that result from plasma entering the inflammatory site. *Hemorrhagic exudates* occur when there is severe tissue injury that damages blood vessels or when there is significant leakage of red cells from the capillaries. *Fibrinous exudates* contain large amounts of fibrinogen and form a thick and sticky meshwork, much like the fibers of a blood clot. *Membranous* or *pseudomembranous exudates* develop on mucous membrane surfaces and are composed of necrotic cells enmeshed in a fibropurulent exudate.

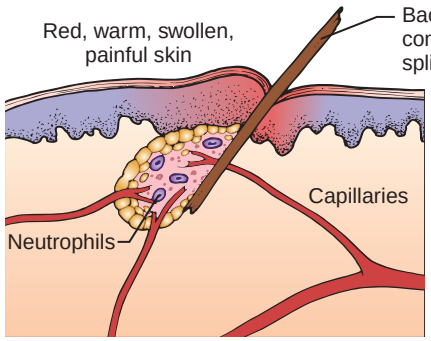
A *purulent* or *suppurative exudate* contains pus, which is composed of degraded white blood cells, proteins, and tissue debris. Certain microorganisms, such as *Staphylococcus*, are more likely to induce localized suppurative inflammation than others. An abscess is a localized area of inflammation containing a purulent exudate that may be surrounded by a neutrophil layer (Fig. 14.6). Fibroblasts may eventually enter the area and wall off the abscess. Because antimicrobial agents cannot penetrate the abscess wall, surgical incision and drainage may be required to effect a cure.

An *ulceration* refers to a site of inflammation where an epithelial surface (*e.g.,* skin or gastrointestinal epithelium) has become necrotic and eroded, often with associated subepithelial inflammation. Ulceration may occur as the result of traumatic injury to the epithelial surface (*e.g.,* peptic ulcer) or because of vascular compromise (*e.g.,* foot ulcers associated with diabetes).

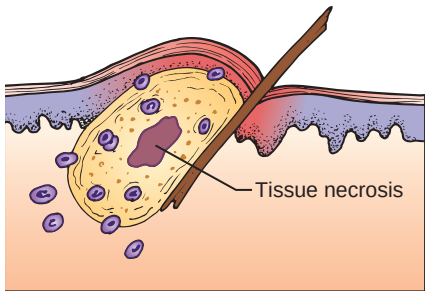
KEY POINTS

THE INFLAMMATORY RESPONSE

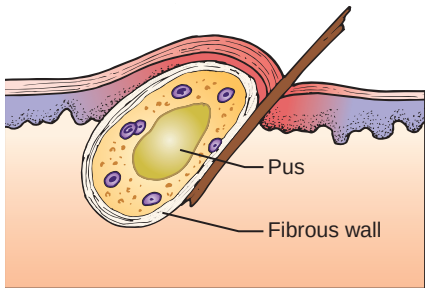
- The manifestations of an acute inflammatory response can be attributed to the immediate vascular changes that occur (vasodilation and increased capillary permeability), the influx of inflammatory cells such as neutrophils, and, in some cases, the widespread effects of inflammatory mediators, which produce fever and other systemic signs and symptoms.
- The manifestations of chronic inflammation are due to infiltration with macrophages, lymphocytes, and fibroblasts, leading to persistent inflammation, fibroblast proliferation, and scar formation.



A Inflammation
Capillary dilation, fluid exudation, neutrophil migration



B Suppuration
Development of suppurative or purulent exudate containing degraded neutrophils and tissue debris



C Abscess formation
Walling off of the area of purulent (pus) exudate to form an abscess

FIGURE 14.6 • Abscess formation. (**A**) Bacterial invasion and development of inflammation. (**B**) Continued bacterial growth, neutrophil migration, liquefaction tissue necrosis, and development of a purulent exudate. (**C**) Walling off of the inflamed area and its purulent exudate to form an abscess.

Chronic Inflammation

In contrast to acute inflammation, which is usually self-limited and of short duration, chronic inflammation is self-perpetuating and may last for weeks, months, or even years. It may develop as the result of a recurrent or progressive acute inflammatory process or from low-grade, smoldering responses that fail to evoke an acute response.

Characteristic of chronic inflammation is an infiltration by mononuclear cells (macrophages) and lymphocytes instead of the influx of neutrophils commonly seen in acute inflammation. Chronic inflammation also involves the proliferation of fibroblasts instead of exudates. As a result, the risk of scarring and deformity usually is greater than in acute inflammation. Agents that evoke chronic inflammation typically are low-grade, persistent infections or irritants that are unable to penetrate deeply or spread rapidly. Among the causes of chronic inflammation are foreign bodies such as talc, silica, asbestos, and surgical suture materials. Many viruses provoke chronic inflammatory responses, as do certain bacteria, fungi, and larger parasites of moderate to low virulence. Examples are the tubercle bacillus and the treponeme of syphilis. The presence of injured tissue such as that surrounding a healing fracture also may incite chronic inflammation. Immunologic mechanisms are thought to play an important role in chronic inflammation. The two patterns of chronic inflammation are a nonspecific chronic inflammation and granulomatous inflammation.

Nonspecific Chronic Inflammation

Nonspecific chronic inflammation involves a diffuse accumulation of macrophages and lymphocytes at the site of injury. Ongoing chemotaxis causes macrophages to infiltrate the inflamed site, where they accumulate owing to prolonged survival and immobilization. These mechanisms lead to fibroblast proliferation, with subsequent scar formation that in many cases replaces the normal connective tissue or the functional parenchymal tissues of the involved structures. For example, scar tissue resulting from chronic inflammation of the bowel causes narrowing of the bowel lumen.

Granulomatous Inflammation

A granulomatous lesion is a distinctive form of chronic inflammation. A granuloma typically is a small, 1- to 2-mm lesion in which there is a massing of macrophages surrounded by lymphocytes. These modified macrophages resemble epithelial cells and sometimes are called *epithelioid cells*.[1] Like other macrophages, the epithelioid cells are derived originally from blood monocytes.[10] Granulomatous inflammation is associated with foreign bodies such as splinters, sutures, silica, and asbestos and with microorganisms that cause tuberculosis, syphilis, sarcoidosis, deep fungal infections, and brucellosis. These types of agents have one thing in common: they are poorly digested and usually are not easily controlled by other inflammatory mechanisms. The epithelioid cells in granulomatous inflammation may clump in a mass or coalesce, forming a multinucleated giant cell that attempts to surround the foreign agent (Fig. 14.7). A dense membrane of connective tissue eventually encapsulates the lesion and isolates it. These cells are often referred to as *foreign body giant cells*.[10]

Systemic Manifestations of Inflammation

Under optimal conditions, the inflammatory response remains confined to a localized area. In some cases, however, local injury can result in prominent systemic manifestations as

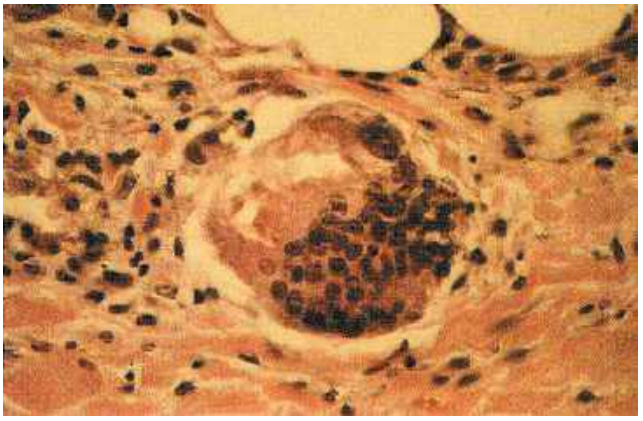


FIGURE 14.7 • Foreign body giant cell. The numerous nuclei are randomly arranged in the cytoplasm. (From Rubin E.& Farber J. L. (Eds.) (2012). *Rubin's Pathology: Clinicopathologic foundations of medicine* (6th ed., p. 81). Philadelphia, PA: Lippincott Williams & Wilkins.)

inflammatory mediators are released into the circulation. The most prominent systemic manifestations of inflammation include the acute-phase response, alterations in white blood cell count, and fever. Localized acute and chronic inflammation may extend to the lymphatic system and lead to a reaction in the lymph nodes that drain the affected area.

Acute-Phase Response

Along with the cellular responses that occur during the inflammatory response, a constellation of systemic effects called the *acute-phase response* occurs. The acute-phase response, which usually begins within hours or days of the onset of inflammation or infection, includes changes in the concentrations of plasma proteins (*i.e.,* acute-phase proteins), skeletal muscle catabolism, negative nitrogen balance, elevated erythrocyte sedimentation rate (ESR), and increased numbers of leukocytes. These responses are generated by the release of cytokines, particularly IL-1, IL-6, and TNF-α. These cytokines affect the thermoregulatory center in the hypothalamus to produce fever, the most obvious sign of the acute-phase response. IL-1 and other cytokines induce an increase in the number and immaturity of circulating neutrophils by stimulating their production in the bone marrow. Other manifestations of the acute-phase response include anorexia, somnolence, and malaise, probably because of the actions of IL-1 and TNF-α on the central nervous system. The metabolic changes, including skeletal muscle catabolism, provide amino acids that can be used in the immune response and for tissue repair. In general, the acute-phase response serves to coordinate the various changes in body function to enable an optimal host response.

In severe bacterial infections (sepsis), the large quantities of microorganisms in the blood result in an uncontrolled inflammatory response with the production and release of enormous quantities of inflammatory cytokines (most notably IL-1 and TNF-α) and development of what is referred to as the *systemic inflammatory response syndrome*.[25] These cytokines cause generalized vasodilation, increased vascular permeability, intravascular fluid loss, myocardial depression, and circulatory shock.

Acute-Phase Proteins. During the acute-phase response, the liver dramatically increases the synthesis of acute-phase proteins such as fibrinogen, C-reactive protein (CRP), and serum amyloid A protein (SAA).[1] The synthesis of these proteins is upregulated by cytokines, especially TNF-α, IL-1 (for SAA), and IL-6 (for fibrinogen and CRP).

CRP was named because it precipitated with the C fraction (C-polypeptide) of pneumococci. The function of CRP is thought to be protective, in that it binds to the surface of invading microorganisms and targets them for destruction by complement and phagocytosis.[26] Although everyone maintains a low level of CRP, this level rises when there is an acute inflammatory response.[26] Recent interest has focused on the use of high-sensitivity CRP (hsCRP) as a marker for increased risk of myocardial infarction in persons with coronary heart disease.[26] It is believed that inflammation involving atherosclerotic plaques in coronary arteries may predispose to thrombosis and myocardial infarction.[26]

During the acute-phase response, SAA protein replaces apolipoprotein A, a component of high-density lipoprotein (HDL) particles; this presumably increases the transfer of HDL from liver cells to macrophages, which can then use these particles for energy. The rise in fibrinogen causes red cells to form stacks (rouleaux) that sediment more rapidly than do individual erythrocytes. This is the basis for the accelerated ESR that occurs in disease conditions characterized by a systemic inflammatory response.

White Blood Cell Response

Leukocytosis, or increased white blood cells, is a frequent sign of an inflammatory response, especially one caused by bacterial infection. The white blood cell count commonly increases from a normal value of 4000 to 10,000 cells/μL to 15,000 to 20,000 cells/μL in acute inflammatory conditions. After being released from the bone marrow, circulating neutrophils have a life span of only about 10 hours and therefore must be constantly replaced if their numbers are to be adequate. With excessive demand for phagocytes, immature forms of neutrophils (bands) are released from the bone marrow.

Bacterial infections produce a relatively selective increase in neutrophils (neutrophilia), whereas parasitic and allergic responses induce eosinophilia. Viral infections tend to produce a decrease in neutrophils (neutropenia) and an increase in lymphocytes (lymphocytosis).[3] A decrease in white blood cells (leukopenia) may occur with overwhelming infections or an impaired ability to produce white blood cells.

Fever

Fever, or *pyrexia,* describes an elevation in body temperature that is caused by an upward displacement of the thermostatic set point of the hypothalamic thermoregulatory center. Temperature is one of the most frequent physiologic responses to be monitored during illness.

Mechanisms

Many proteins, breakdown products of proteins, and certain other substances released from bacterial cell membranes can cause a change in the set point to rise. Fever is resolved when the condition that caused the increase in the set point is removed. Fevers that are regulated by the hypothalamus usually do not rise above 41°C (105.8°F), suggesting a built-in thermostatic safety mechanism. Temperatures above that level are usually the result of superimposed activity, such as convulsions, hyperthermic states, or direct impairment of the temperature control center.

Pyrogens are exogenous or endogenous substances that produce fever. *Exogenous pyrogens* are derived from outside the body and include such substances as bacterial products, bacterial toxins, or whole microorganisms. Exogenous pyrogens induce host cells to produce fever-producing mediators called *endogenous pyrogens*. When bacteria or breakdown products of bacteria are present in blood or tissues, phagocytic cells of the immune system engulf them. These phagocytic cells digest the bacterial products and then release pyrogenic cytokines, principally interleukin-1 (IL-1), interleukin-6 (IL-6), and tumor necrosis factor-α (TNF-α), into the bloodstream for transport to the hypothalamus, where they exert their action.[3] These cytokines induce prostaglandin E_2 (PGE_2), which is a metabolite of arachidonic acid (an intramembrane fatty acid). It is hypothesized that when interleukin (IL-1B) interacts with the endothelial cells of the blood–brain barrier in the capillaries of the organum vasculosum laminae terminalis (OVLT), which is in the third ventricle above the optic chiasm, PGE_2 is released into the hypothalamus.[1]

At this point, PGE_2 binds to receptors in the hypothalamus to induce increases in the thermostatic set point through the second messenger cyclic adenosine monophosphate (cAMP). In response to the increase in its thermostatic set point, the hypothalamus initiates shivering and vasoconstriction that raise the body's core temperature to the new set point, and fever is established.

Although the central role of PGE_2 in raising the set point of the hypothalamic thermoregulatory center and producing fever is not questioned, research suggests that the febrile response to invading gram-negative bacteria and their products (mainly endotoxic lipopolysaccharides) is mediated by PGE_2.[1]

In addition to their fever-producing actions, the endogenous pyrogens mediate a number of other responses. For example, IL-1 and TNF-α are inflammatory mediators that produce other signs of inflammation such as leukocytosis, anorexia, and malaise. Many noninfectious disorders, such as myocardial infarction, pulmonary emboli, and neoplasms, produce fever. In these conditions, the injured or abnormal cells incite the production of endogenous pyrogens. For example, trauma and surgery can be associated with up to 3 days of fever. Some malignant cells, such as those of leukemia and Hodgkin disease, secrete chemical mediators that function as endogenous pyrogens.

A fever that has its origin in the central nervous system is sometimes referred to as a *neurogenic fever*. It usually is caused by damage to the hypothalamus due to central nervous system trauma, intracerebral bleeding, or an increase in intracranial pressure. Neurogenic fever is characterized by a high temperature that is resistant to antipyretic therapy and is not associated with sweating.

Purpose

The purpose of fever is not completely understood. However, from a purely practical standpoint, fever is a valuable index to health status. For many, fever signals the presence of an infection and may legitimize the need for medical treatment. There is little research to support the belief that fever is harmful unless the temperature rises above 40°C (104°F). However, animal studies have demonstrated a clear survival advantage in infected members with fever compared with animals that were unable to produce a fever. It has also been shown that small elevations in temperature such as those that occur with fever enhance immune function by T lymphocyte proliferation.[3] Many of the microbial agents that cause infection grow best at normal body temperatures, and their growth is inhibited by temperatures in the fever range.

Yet, fever is negative in many situations such as in older adults who have cardiac or pulmonary disease because it causes more of a demand for oxygen. For every elevated 1°C of temperature, the BMR increases by 7% and causes increased work of the heart. Fever can also produce confusion, tachycardia, and tachypnea. Cell damage can occur when temperatures are elevated greater than 42.2°C (108°F), and this can ultimately cause life-threatening acidosis, hypoxia, and hyperkalemia.[9]

Patterns

The patterns of temperature change in people with fever vary. Additionally, the average diurnal variation in temperature yields a peak rise in the late afternoon or early evening.[2] These patterns can be described as intermittent, remittent, sustained, or relapsing (Fig. 10.3). An *intermittent fever* is one in which temperature returns to normal at least once every 24 hours. In a *remittent fever,* the temperature does not return to normal and varies a few degrees in either direction. In a *sustained* or *continuous fever,* the temperature remains above normal with minimal variations (usually <0.55°C or 1°F). A *recurrent* or *relapsing fever* is one in which there is one or more episodes of fever, each as long as several days, with one or more days of normal temperature between episodes.

Critical to the analysis of a fever pattern is the relation of heart rate to the level of temperature elevation. Most people respond to an increase in temperature with an appropriate increase in heart rate. The observation that a rise in temperature is not accompanied by the anticipated change in heart rate can provide useful information about the cause of the fever. For example, a heart rate that is slower than would be anticipated can occur with Legionnaire disease and drug fever, and a heart rate that is more rapid than anticipated can be symptomatic of hyperthyroidism and pulmonary emboli.

Disorders of the Immune Response

15

Nancy A. Moriber

HYPERSENSITIVITY DISORDERS

Type I, Immediate Hypersensitivity Disorders
Anaphylactic (Systemic) Reactions
Atopic (Local) Reactions
Type II, Antibody-Mediated Disorders
Complement-Activated Cell Destruction
Antibody-Dependent Cell Cytotoxicity
Complement- and Antibody-Mediated Inflammation
Antibody-Mediated Cellular Dysfunction
Type III, Immune Complex–Mediated Disorders
Systemic Immune Complex Disorders
Localized Immune Complex Reactions
Type IV, Cell-Mediated Hypersensitivity Disorders
Allergic Contact Dermatitis
Hypersensitivity Pneumonitis

The human immune system is a complex, multidimensional system designed to protect the host against invasion by foreign substances, microorganisms, and toxins. In addition, it helps to protect against the proliferation of neoplastic cells and plays a key role in the process of inflammation and wound healing. Unfortunately, under certain circumstances the immune system can become inefficient or hyperactive, causing the development of debilitating and/or life-threatening diseases. These disease processes can take the form of immunodeficiency disorders, allergic or hypersensitivity reactions, transplant rejection, and autoimmune disorders. Regardless of the manifestation, the underlying cause can be traced back to an abnormality to one of the cellular or chemical components of the innate and adaptive immune responses.

HYPERSENSITIVITY DISORDERS

After completing this section of the chapter, you should be able to meet the following objectives:

- Describe the adaptive immune responses that protect against microbial agents and hypersensitivity responses.
- Discuss the immune response involved in the development of type I, type II, type III, and type IV hypersensitivity reactions.
- Describe the pathogenesis of common hypersensitivity reactions including allergic rhinitis, food allergy, serum sickness, Arthus reaction, contact dermatitis, and hypersensitivity pneumonitis.

Activation of the immune system normally results in the mobilization and coordination of T-cell and B-cell activity in order to protect the body from invading microorganisms and toxic substances. Unfortunately, this same system is capable of causing serious damage when it does not function as intended. Hypersensitivity is defined as an abnormal and excessive response of the activated immune system that causes injury and damage to host tissues. Disorders caused by immune responses are collectively referred to as *hypersensitivity reactions*. Hypersensitivity reactions are classified as one of four types: type I, IgE-mediated disorders; type II, antibody-mediated disorders; type III, complement-mediated immune disorders; and type IV, T cell–mediated disorders (Table 15.1). They differ with respect to the specific components of the immune response initiated, the onset of symptoms, and the eventual mechanism of injury.

Type I, Immediate Hypersensitivity Disorders

Type I hypersensitivity reactions are IgE-mediated reactions that develop rapidly upon exposure to an antigen. Type I hypersensitivity reactions represent the classic allergic response, and in this context, antigens are referred to as *allergens*. Environmental, medical, and pharmaceutical allergens are all capable of initiating a type I hypersensitivity reaction. Common allergens encountered include pollen proteins, house dust mites, animal dander, foods, household chemicals, and pharmaceutical agents like the antibiotic penicillin. Exposure to the allergen can be through inhalation, ingestion, injection, or skin contact. Depending on the portal of entry, type I reactions may be localized to a discrete area of the body (*e.g.,* contact dermatitis) or systemic causing significant disease (*e.g.,* asthma) and life-threatening anaphylaxis.

Two types of cells play a key role in the development of a type I hypersensitivity reactions: type 2 helper T (T_2H) cells and mast cells or basophils. Two distinct subtypes of helper T cells (T_1H or T_2H) develop from activated $CD4^+$ helper T cells based upon the cytokines expressed by the antigen-presenting cells (APCs) at the site of activation. Macrophages and dendritic cells direct the maturation of $CD4^+$ helper T cells toward the T_1H subtype, whereas mast cells and T cells induce differentiation toward the T_2H subtype. The T_1H cells stimulate the differentiation of B cells into IgM- and IgG-producing plasma cells. The T_2H cells direct B lymphocytes to switch class and produce the IgE antibodies necessary for an allergic or hypersensitivity response. In addition, T_2H cytokines are responsible for the mobilization and activation of mast cells, basophils, and eosinophils, inducing inflammatory responses that are distinct from T_1H reactions.[78]

Mast cells, basophils, and eosinophils are essential to the development of type I hypersensitivity reactions. They are members of the *granulocyte* class of leukocytes because they contain granules rich in chemical mediators such as histamine and heparin. These mediators may be preformed or are enzymatically activated in response to T_2H signaling. Once they are released they are capable of inducing a wide range of cellular responses. Mast cells and basophils are histologically similar and derived from $CD34^+$ progenitor cells.[79] However, basophils are confined to the bloodstream, and mast cells are distributed throughout connective tissue, especially in areas beneath the skin and mucous membranes of the respiratory, gastrointestinal, and genitourinary tracts and adjacent to blood and lymph vessels.[79] This places the mast cells in close proximity to surfaces with frequent exposure to allergen. Mast cells in different parts of the body and even in a single site can have significant differences in mediator content and sensitivity to agents that produce mast cell degranulation.

Type I hypersensitivity reactions are dependent upon IgE-mediated activation of mast cells and basophils (Fig. 15.2). During the initial exposure to an antigen, allergen-specific IgE is produced as part on the normal humoral response and IgE to the high-affinity IgE receptors known as FcεRI, expressed on the surface of mast cells and basophils.[79–81] In contrast, lymphocytes, eosinophils, and platelets bind IgE via low-affinity FcεRII receptors.[82] On subsequent exposure to an allergen, the multimeric cross-linkages between IgE antibodies are formed creating a bridge between two IgE molecules.[82] When IgE receptors aggregate, they induce a signal transduction that stimulates mast cell degranulation and release of vasoactive chemical mediators, the

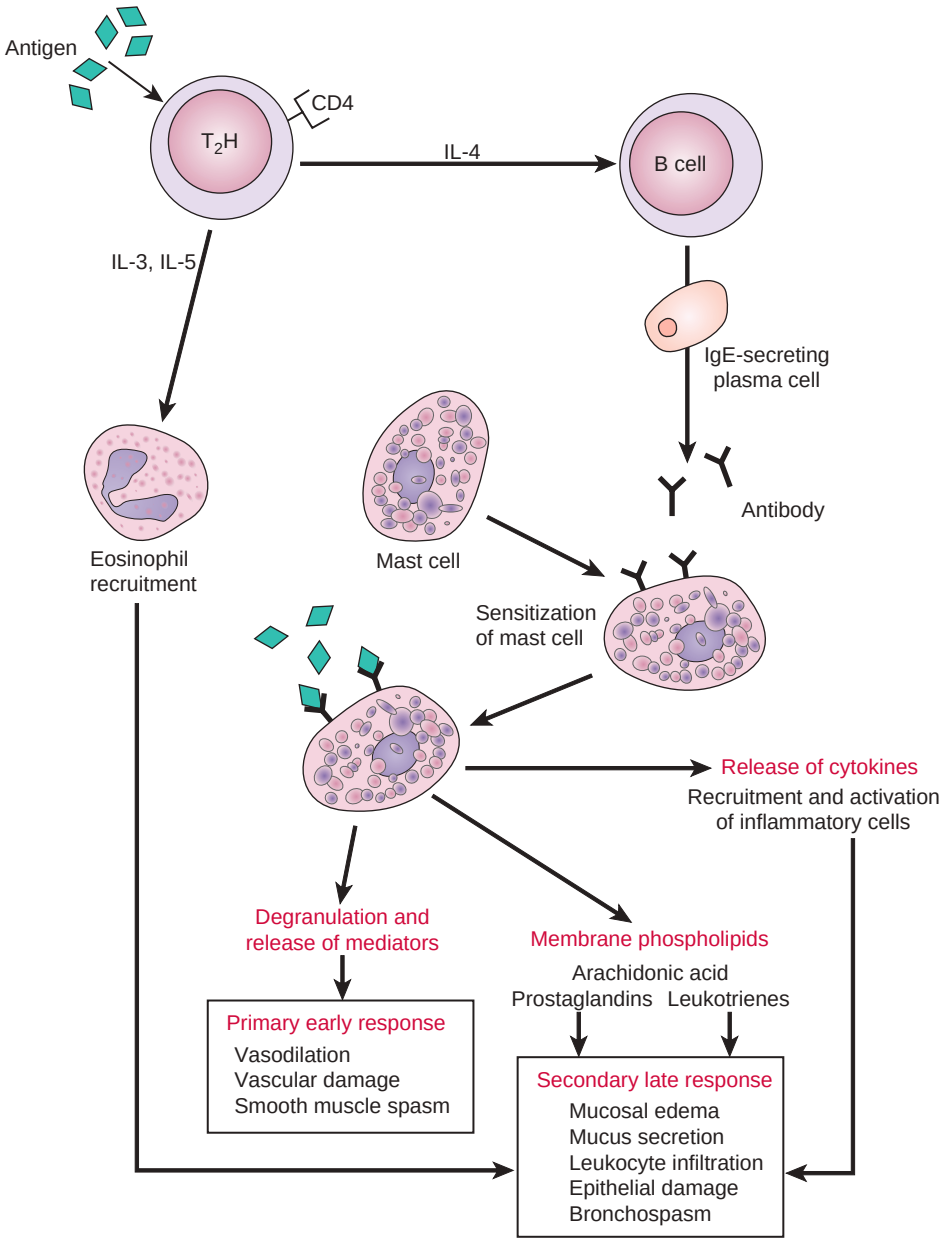


FIGURE 15.2 • Type I, IgE-mediated hypersensitivity reaction. The stimulation of B-cell differentiation by an antigen-stimulated type 2 helper (T_2H) T cell leads to plasma cell production of IgE and mast cell sensitization. Subsequent binding of the antigen produces degranulation of the sensitized mast cell with release of preformed mediators that leads to a primary or early-phase response. T_2H T-cell recruitment of eosinophils, along with the release of cytokines and membrane phospholipids from the mast cell, leads to a secondary or late-phase response.

synthesis and secretion of platelet-activating factor (PAF) and leukotrienes, and the secretion of many growth factors, cytokines, and chemokines.[79,80]

Most type I hypersensitivity reactions such as bronchial asthma develop in two distinct and well-defined phases: (1) a primary or initial-phase response characterized by vasodilation, vascular leakage, and smooth muscle contraction and (2) a secondary or late-phase response characterized by more intense infiltration of tissues with eosinophils and other acute and chronic inflammatory cells as well as tissue destruction in the form of epithelial cell damage.

The primary or initial-phase response usually begins with 5 to 30 minutes of exposure to an allergen and subsides within 60 minutes. It is mediated by acute mast cell degranulation and the release of preformed and/or enzymatically activated mediators. These mediators include histamine, serotonin, acetylcholine, adenosine, chemotactic mediators, growth factors, and neutral proteases such as chymase and trypsin that lead to generation of kinins.[80,82] Histamine is the most recognized mediator of type I hypersensitivity reactions. It is a potent vasoactive amine that increases nitric oxide production, relaxes vascular smooth muscle, increases the permeability of capillaries and venules, and causes smooth muscle contraction and bronchial constriction. Acetylcholine mimics many of the actions of histamine and produces bronchial smooth muscle contraction and dilation of small blood vessels via activation of the parasympathetic nervous system. The kinins are a group of potent inflammatory peptides that once activated through enzymatic modification, produce vasodilation and smooth muscle contraction as well.

The secondary or late phase of the type I hypersensitivity response occurs 2 to 8 hours after resolution of the initial phase and can last for several days. In some cases, the late phase may be significantly prolonged or only partially resolved as in the case of uncontrolled bronchial asthma. It results from the action of lipid mediators and cytokines released from immune cells as part of the normal inflammatory process. The lipid mediators, which are derived from phospholipids found in mast cell membranes, are broken down to form arachidonic acid during the process of mast cell degranulation. Arachidonic acid is then utilized in the synthesis of leukotrienes and prostaglandins, which produce end-organ effects similar to histamine and acetylcholine, except that they have a longer onset and prolonged duration of action. Mast cells also produce cytokines and chemotactic factors that promote migration of eosinophils and leukocytes to the site of allergen exposure, contributing to late-phase response.

It is important to point out that not all IgE-mediated reactions result in hypersensitivity or the development of disease. The IgE-mediated antibody response is a normal part of the immune response to parasitic infection. During the late phase of the response, IgE antibodies are directed against parasite larvae stimulating recruitment of large bodies of inflammatory cells including eosinophils and causing cell-mediated cytotoxicity. This type of type I hypersensitivity reaction is particularly important in developing countries where much of the population is infected with intestinal parasites.

Anaphylactic (Systemic) Reactions

Anaphylaxis is a catastrophic, systemic life-threatening IgE-mediated hypersensitivity reaction associated with the widespread release of histamine into the systemic circulation that produces massive vasodilation, hypotension, arterial hypoxia, and airway edema.[82] It results from the presence of even minute quantities of allergen that are introduced into the body via the airway, skin, blood, or gastrointestinal mucosa. The level of severity, therefore, depends on the preexisting degree of sensitization and not with the quantity of exposure.

Clinical manifestations occur along a continuum in severity and can be graded on a scale of I to IV.[82] Grade I reactions are usually confined to the cutaneous and mucosal tissues manifesting as erythema and urticaria, with or without angioedema. Grade II reactions progress to include moderate multisystem signs such as hypotension, tachycardia, dyspnea, and gastrointestinal disturbances (*e.g.*, nausea, vomiting, diarrhea, abdominal cramping from mucosal edema). Grade III reactions become life threatening because of the development of bronchospasm, cardiac dysrhythmias, and cardiac collapse. Once a hypersensitivity reaction reaches grade IV, cardiac arrest has occurred and management is purely resuscitative in nature.

Preventing exposure to potential triggers that cause anaphylaxis is essential because any reaction can be life threatening. All people with potential for anaphylaxis should be advised to wear or carry a medical alert bracelet, necklace, or other identification to inform emergency personnel of the possibility of anaphylaxis. In addition, people with a history of anaphylaxis should be provided with preloaded epinephrine syringes and instructed in their use.

The initial management of anaphylaxis is dependent upon the stage at which a person presents, but should always focuses on withdrawal of the offending allergen, maintenance of a patent airway, establishment of appropriate intravenous access, volume resuscitation, and administration of epinephrine.[82,83] It is important to explain to all people with a potential for anaphylaxis that if they have a reaction and self-treat with epinephrine, it is essential for them to seek immediate professional help regardless of their initial response to self-treatment because reactions can reoccur.

Atopic (Local) Reactions

Local hypersensitivity reactions usually occur when the offending allergen is confined to a particular site of exposure. The term atopy is frequently used to describe these reactions and refers to a genetic predisposition to the development of immediate, type I IgE-mediated hypersensitivity reactions upon exposure to common environmental antigens such as pollens, food, or animal dander. Atopic reactions most commonly manifest as urticarial (hives), allergic rhinitis, atopic dermatitis, and bronchial asthma. People prone to atopy frequently develop reactions to more than one environmental allergen with symptoms present at different times throughout the year.

The incidence of immediate hypersensitivity reactions tends to be greater in people with a family history of atopy, yet the genetic basis for these disorders is not completely understood. Because of underlying genetic differences in people with type I hypersensitivity, the exact genome has been difficult to delineate. However, several chromosomal regions have been shown to contain gene sequences linked to the development of asthma and atopy, including the cytokine cluster on chromosome 5q, IFNG (IFNg) and STAT6 on 12q, and IL4R on 16p.[84,85] People with atopic allergic conditions tend to have high total serum and allergen-specific levels of IgE as well as increased numbers of eosinophils, basophils, and mast cells. Although the IgE-triggered response is likely a key factor in the pathophysiology of atopic allergic disorders, it is not the only factor and may not be responsible for the development of all forms of atopic dermatitis and asthma.

Allergic Rhinitis. Allergic rhinitis is a common hypersensitivity disorder of the upper respiratory tract that affects between 20% and 40% of the western population.[86] Symptoms include rhinorrhea (runny nose), nasal obstruction, sneezing, nasal itching, and watery eyes (conjunctivitis). The diagnosis of allergic rhinitis is made based upon the person's clinical presentation and a positive skin prick test or the presence of serum-specific IgE antibodies to aeroallergens. People with allergic rhinitis frequently present with others forms of atopy such as allergic asthma and urticaria. Severe attacks may be accompanied by systemic malaise, fatigue, headache, and muscle soreness from sneezing. Fever is absent. The allergens

associated with the development of allergic rhinitis are airborne and are therefore deposited directly onto the nasal mucosa. Typical allergens include pollens from ragweed, grasses, trees, and weeds; fungal spores; house dust mites; animal dander; and feathers.

Clinical manifestations are dependent upon the timing and severity of exposure. In people who are chronically exposed to allergens, symptoms can be present throughout the year. This form of atopy is known as *perennial rhinitis*. In contrast, people who present with symptoms only when exposed to high allergen counts, such as in the fall or spring, are said to have *seasonal allergic rhinitis*. Symptoms that become worse at night suggest a household allergen, and symptoms that disappear on weekends suggest occupational exposure.

The allergic response in allergic rhinitis is located specifically in the nasal mucosa. When aeroallergens are inhaled, they are deposited mainly on the nasal mucosa where they are presented to T cells by APCs. In the presence of cellular cytokines, B-cell class switching occurs, resulting in an increase in IgE production.[86] Once the allergen–IgE complex is formed, infiltration of the nasal mucosa by T_2H cells, mast cells, basophils, eosinophils, and Langerhans cells takes place, inducing a full cell-mediated immune response.

Treatment of allergic rhinitis focuses on the institution of avoidance measures and control of symptoms. Whenever possible, the offending allergen should be removed from the environment, or exposure should be kept to a minimum. Most symptoms can be controlled with over-the-counter antihistamines and topical nasal decongestants. Tolerance and rebound congestion may occur with chronic administration of topical nasal decongestants, so their use should be limited to less than 1 week. More severe symptoms may require prescription medication including topical nasal corticosteroids (*e.g.,* mometasone or *Nasonex*) and antihistamines (*e.g.,* azelastine HCl). Mast cell stabilizers, such as intranasal cromolyn sodium, that prevent localized mast cell degranulation and release of intracellular mediators may be useful, especially when administered prophylactically. In people whose symptoms cannot be successfully controlled with these measures, a program of desensitization known as immunotherapy ("allergy shots") may be undertaken. Desensitization involves the frequent administration of progressively larger quantities of the offending antigen(s). The antigens stimulate production of high levels of IgG antibodies, which are capable of combining with the antigen and preventing activation of cell-bound IgE antibodies.

Food Allergies. Food allergy is very common in western countries around the world, often manifesting with life-threatening consequences. In fact, food-induced anaphylaxis is the leading cause of emergency room admissions, especially among children.[87] Currently, the prevalence rate of food allergy is between 3% and 6%, and, according to the Centers for Disease Control (CDC), this represents an increase of 18% over the past decade.[88,89] The exact etiology of the increase in cases is unknown. Any food is capable of inducing a hypersensitivity reaction in susceptible people, but the most commonly implicated foods include peanuts, tree nuts, and shellfish. In addition, milk is frequently implicated in children.[87,89] People with asthma, adolescents, and those with a personal or family history of food allergy are at increased risk of severe reactions.

The clinical manifestations of food allergy are dependent upon many factors including the amount of food ingested, the presence of an empty stomach, concurrent illness and medication, exercise, and the phase of the menstrual cycle.[87,90] Reactions may differ within a given person during different exposures, but the primary symptoms are seen in the skin, gastrointestinal tract, and respiratory system in approximately 80% of cases. The ability of a specific food to trigger a type I hypersensitivity reaction may be changed during the cooking process because heating can alter (denature) the protein structure of an allergen, so that it is no longer able to trigger the humoral response. Both acute reactions (hives and anaphylaxis) and chronic reactions (asthma, atopic dermatitis, and gastrointestinal disorders) to food allergens can occur.

Anaphylactic reactions to food allergens are common, and the presentation may differ between adults and children. Adults typically present with severe symptoms including cardiovascular collapse, whereas severe abdominal pain, hives, allergic rhinitis, conjunctivitis, and facial flushing are more common in children.[87] Within the pediatric population, wheezing and stridor are more common in preschoolers and older children, while hives and vomiting are usually seen in infants.[91] The majority of the reactions manifest within 1 hour of exposure, but delayed reactions are possible secondary to delayed absorption of the allergen. A rare form of anaphylaxis associated with food is known as *food-dependent exercise-induced anaphylaxis* (FDEIA). In FDEIA, both exercise and the food allergen are tolerated independently, and symptoms do not occur in the absence of exercise.[87,92] The pathophysiology is not completely understood but seems to suggest that a pliable state of immunologic tolerance exists in susceptible people.[92] Alterations in plasma osmolality and pH, tissue enzyme activity, blood blow distribution, and gastrointestinal permeability may occur during exercise, which result in facilitated allergen recognition and binding.[92]

Food allergies can occur at any age, but tend to manifest during childhood. The allergic response is activated when a specific food allergen comes in contact with IgE antibody present in the intestinal mucosa and subsequently stimulates local and systemic release of histamine and other cytokines necessary in the allergic response. Carbohydrates, lipids, proteins, or food additives, such as preservatives, colorings, or flavorings, can all serve as potential allergens in the allergic response. Cross-sensitivity to allergens between foods in closely related food groups is common. Therefore, a person can contain common cross-reacting allergens. For example, some people are allergic to all legumes (*i.e.,* beans, peas, and peanuts).

Diagnosis of food allergies is multifaceted and relies upon a careful food history and provocative diet testing. Provocative testing involves the systematic elimination of suspected allergen(s) from the diet for a time to see if the

symptoms disappear and then reintroducing the allergens(s) to the diet to determine if the symptoms reappear. Only one food should be tested at a time if definitive diagnosis is sought. Allergen-specific serum IgE levels can also be tested if the risk of provocative food testing is too great.

Treatment of food allergy focuses specifically on the avoidance of the offending allergen. However, this can be difficult, especially in people that are exquisitely sensitive to a particular food protein because foods (processed or fresh) may be contaminated with the protein during handling of the food. Foods that are prepared in the same processing plants that handle tree nuts may be potential allergen sources and, therefore, illicit an allergic response in susceptible people. As a result, warnings are placed on all goods processed in facilities that handle highly allergenic foods. People with severe allergies or a history of anaphylaxis should be educated to carry an EpiPen and to seek emergency care immediately after exposure.

Type II, Antibody-Mediated Disorders

Type II (antibody-mediated) hypersensitivity or *cytotoxic hypersensitivity* reactions are mediated by IgG or IgM antibodies directed against target antigens on specific host cell surfaces or tissues. The antigens may be either intrinsic, inherently part of the host cell, or extrinsic, incorporated into the cell surface upon exposure to a foreign substance or infectious agent. Thus, the tissues that express the target antigens determine the clinical manifestations of type II hypersensitivity reactions. These antigens are known as *tissue-specific antigens*.[93] There are four general mechanisms by which type II hypersensitivity reactions can be propagated, but regardless of the pathway, it is always initiated by the binding of IgG or IgM antibody to tissue-specific antigens. These mechanisms include complement-activated cell destruction, antibody-mediated cell cytotoxicity, complement- and antibody-mediated inflammation, and antibody-dependent modulation of normal cell surface receptors[93] (Fig. 15.3).

Complement-Activated Cell Destruction

The destruction of target cells in type II hypersensitivity reactions can occur as a result of activation of the complement system via the classic pathway. First, formation of the membrane attack complex (MAC) by activation of C5-C9 allows the passage of ions, small molecules, and water into the cell, causing direct lysis of the cell. In addition, IgG and the complement fragment C3b act as opsonins by binding to receptors located

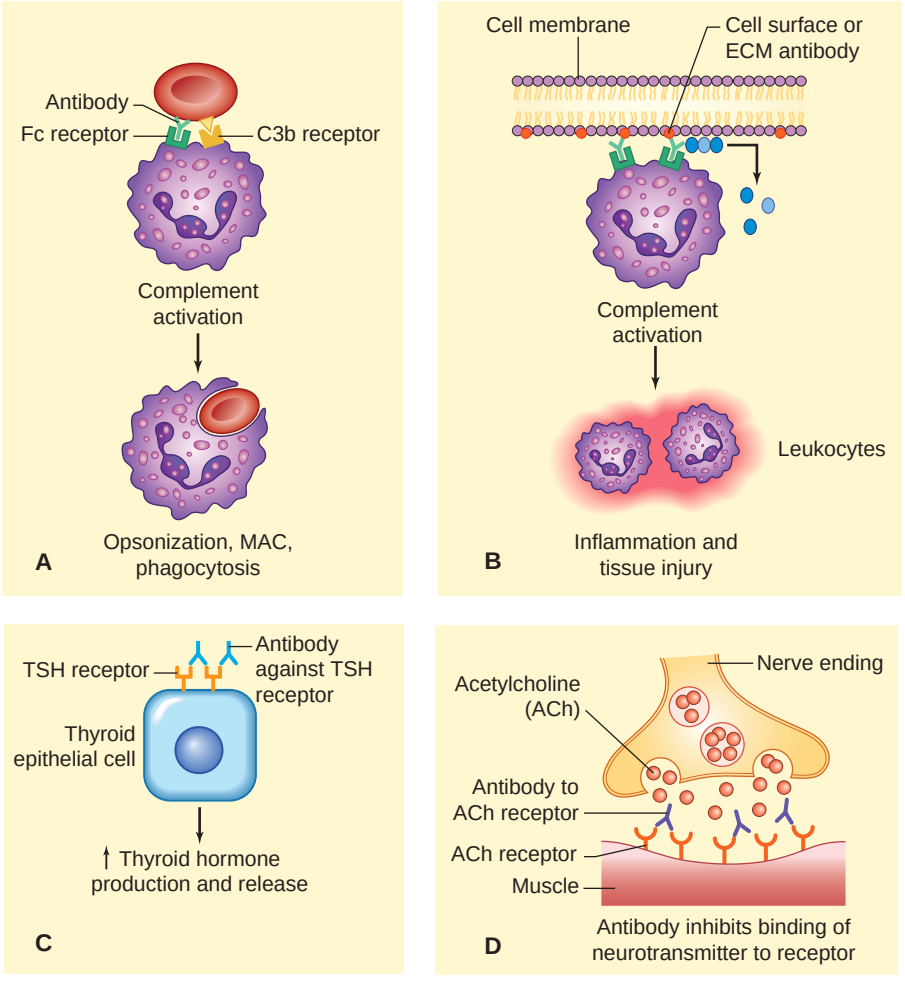


FIGURE 15.3 • Type II, hypersensitivity reactions result from binding of antibodies to normal or altered surface antigens. (**A**) Opsonization and complement- or antibody receptor–mediated phagocytosis or cell lysis through membrane attack complex (MAC). (**B**) Complement- and antibody receptor–mediated inflammation resulting from recruitment and activation of inflammation-producing leukocytes (neutrophils and monocytes). (**C**) Antibody-mediated cellular dysfunction, in which antibody against the thyroid-stimulating hormone (TSH) receptor increases thyroid hormone production, and (**D**) antibody to acetylcholine receptor inhibits receptor binding of the neurotransmitter in myasthenia gravis.

on the cell surfaces of macrophages. This process activates the macrophages, which then destroys the target cells by phagocytosis. Thus, activation of the complement system produces a twofold response that culminates in cell destruction.

In people with AIHA, autoantibodies target epitopes located on red blood cells.[94,95] Erythrocytes coated with these autoantibodies are destroyed by phagocytes in the liver or spleen. Some, but not all, autoantibody types also induce phagocytosis and cell lysis via the complement system. The same process occurs in utero in the development of *erythroblastosis fetalis* or Rh incompatibility. Women who are Rh negative lack RhD antigen on their erythrocytes but produce anti-D antibodies. In the Rh-positive fetus, maternal anti-D antibodies will coat fetal red blood cells containing RhD, allowing them to be removed from the fetal circulation by macrophage- and monocyte-mediated phagocytosis.[95]

Antibody-Dependent Cell Cytotoxicity

Antibody-dependent cell cytotoxicity (ADCC) incorporates components of both the innate and adaptive immune responses in the destruction of target cells, but is not dependent upon activation or utilization of complement proteins. Rather the mechanism relies upon the activity of nonspecific NK cells, but other cells such as macrophages and eosinophils have been implicated.[96] The Fc-fragment of the IgG antibody binds to Fc receptor (FcγR) on the surface of the effector cell, and the variable fragment binds to the epitope on the target cell surface, causing release of chemotactic substances and destruction of the target cell.[96] ADCC is a common antiviral mechanism. It has been implicated in the development of several autoimmune disorders including *pemphigus vulgaris*.

Complement- and Antibody-Mediated Inflammation

When antigens that are normally expressed on vessel walls or that circulate in the plasma are deposited on the surface of endothelial cells or extracellular tissues, the manifestations are the result of localized inflammation as opposed to cell destruction. The presence of antibody in the tissues activates the complement cascade, resulting in the release of the activated complement proteins C3a and C5a, which in turn attracts neutrophils to the area and stimulates the deposition of complement protein C3b.[93] Neutrophils bind to the Fc antibody fragment or to C3b, but rather than destroying cells via phagocytosis, undergo degranulation and release of chemical mediators (enzyme and oxidases) involved in the inflammatory response. Antibody-mediated inflammation is responsible for the tissue injury seen in Goodpasture disease, which is characterized by the presence of autoantibodies against the α3NC1 domain of collagen IV, an essential protein in the basement membranes of the kidneys and lungs.[97] The antibody-mediated neutrophil activation causes the development of glomerulonephritis, acute renal failure, and hemorrhagic lung disease if immunosuppressive therapy is not initiated.

Antibody-Mediated Cellular Dysfunction

In some type II reactions, the binding of antibody to specific target cell receptors causes the cell to malfunction in some way, rather than initiating the process of cell destruction. The antibody–receptor complex that is formed modulates the function of the receptor by preventing or enhancing interactions with normal ligands, by replacing ligand and directly stimulating receptors, or by destroying the receptor entirely. The symptoms of type II hypersensitivity reactions caused by antibody-mediated cellular dysfunction are dependent upon the specific receptor(s) that are targeted. In Graves disease, autoantibodies, known as *thyrotropin-binding inhibitory Ig,* bind to and activate thyroid-stimulating hormone (TSH) receptors on thyroid cells, stimulating thyroxine production and the development of hyperthyroidism.[98,99] In contrast, in myasthenia gravis, autoantibodies are directed toward the nicotinic acetylcholine receptors located on the motor end plates within the neuromuscular junction, where they block the action of acetylcholine and stimulate the destruction of the receptors, leading to decreased neuromuscular function.

KEY POINTS

ALLERGIC AND HYPERSENSITIVITY DISORDERS

- Type I hypersensitivity reactions are dependent upon IgE-mediated activation of mast cells and basophils and the subsequent release of chemical mediators of the inflammatory response.
- Type II (antibody-mediated) hypersensitivity or *cytotoxic hypersensitivity* reactions are mediated by IgG or IgM antibodies directed against target antigens on specific host cell surfaces or tissues and result in complement-mediated phagocytosis and cellular injury.
- Type III (immune complex) hypersensitivity is caused by the formation of antigen–antibody immune complexes in the bloodstream, which are subsequently deposited in vascular epithelium or extravascular tissues and which activate the complement system and induce a massive inflammatory response.
- Type IV (cell-mediated) hypersensitivity involves tissue damage in which cell-mediated immune responses with sensitized T lymphocytes cause cell and tissue injury. Although all are T cell mediated, the pathophysiologic mechanisms and sensitized T-cell populations involved differ.

Type III, Immune Complex–Mediated Disorders

Immune complex allergic disorders are caused by the formation of antigen–antibody immune complexes in the bloodstream, which are later deposited in vascular epithelium or

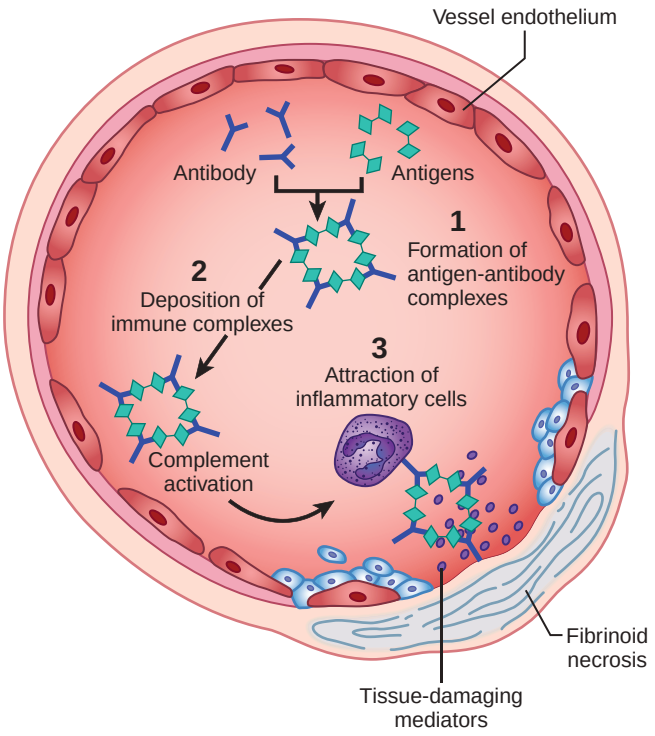


FIGURE 15.4 • Type III, immune complex reactions involving complement-activating IgG or IgM immunoglobulins with (*1*) formation of blood-borne immune complexes that are (*2*) deposited in tissues. Complement activation at the site of immune complex deposition (*3*) leads to attraction of leukocytes that are responsible for vessel and tissue injury.

extravascular tissues (Fig. 15.4). The deposition of these complexes in the tissues activates the complement system and induces a massive inflammatory response. Like type II hypersensitivity reactions, IgG and IgM antibodies activate immune complex–mediated disorders. However, in type III reactions, the antibody–antigen complexes are formed first in the plasma and then deposited in the tissues. The clinical manifestations may therefore have little to do with the particular antigenic target, but rather with the site of immune complex deposition. Immune complexes formed in the circulation can produce damage in any end-organ vessels including those feeding the renal glomerulus, skin, lung, and joint synovium. They can be generalized if the immune complexes are deposited in many organs or localized to a particular organ, such as the kidney, joints, or small blood vessels of the skin. Once deposited, the immune complexes elicit an inflammatory response by activating complement and generating chemotactic factors that recruit neutrophils and other cells of the inflammatory response. The activation of these inflammatory cells by immune complexes and complement, accompanied by the release of potent inflammatory mediators, is directly responsible for the injury. Type III reactions are responsible for the vasculitis seen in many autoimmune diseases including *systemic lupus erythematosus* (SLE) and acute glomerulonephritis.

Systemic Immune Complex Disorders

Serum sickness is a clinical syndrome that results from the formation of insoluble antigen–antibody immune complexes in the presence of antigen excess and subsequent generalized deposition in target tissues such as blood vessels, joints, and the heart and kidneys. The deposited immune complexes activate the complement cascade, increase vascular permeability, and stimulate the recruitment of phagocytic cells. The net result is generalized tissue damage and edema. Clinical manifestations include rash, fever, generalized lymphadenopathy, and arthralgias, which usually begin approximately 1 to 2 weeks after the initial antigen exposure and subside upon withdrawal of the offending agent.[100] In previously sensitized people, severe and life-threatening reactions have been reported. Serum sickness was first described in people receiving foreign serum, such as horse serum, for the treatment of diphtheria and scarlet fever. This antigen load was capable of stimulating the production of large quantities of immune complexes that were deposited in tissues causing activation of mast cells, monocytes, polymorphonuclear leukocyte, and platelets.[101] Today, large volume injection of foreign proteins is rarely indicated, but a variety of drugs including beta-lactam antibiotics and sulfonamides are capable of causing similar reactions.[100]

Treatment of serum sickness usually is directed toward removal of the sensitizing antigen and providing symptom relief. This may include aspirin for joint pain and antihistamines for pruritus. Epinephrine or systemic corticosteroids may be used for severe reactions.

Localized Immune Complex Reactions

The *Arthus reaction* is a localized immune complex reaction associated with discrete tissue necrosis, usually in the skin. It is caused by repeated local exposure to an antigen, where high levels of preformed circulating antibodies exist. Symptoms usually begin within 1 hour and peak within 6 to 12 hours of an exposure.[93] Lesions are typically red, raised, and inflamed. Ulcers may often form at the center of the lesions because of the release of inflammatory cytokines. The mechanism of the Arthus reaction is not completely understood but is believed to be the result of localized contact of injected antigen with circulating IgG antibody. This reaction is the prototypical model for the development of localized vasculitis associated with certain drug reactions in humans.

Type IV, Cell-Mediated Hypersensitivity Disorders

Type IV hypersensitivity reactions differ from type I to III hypersensitivity reactions in that they are cell-mediated and delayed, rather than antibody-mediated and immediate immune responses (Fig. 15.5). The cell-mediated immune response is normally the principal mechanism of defense against a variety of microorganisms, including intracellular pathogens such as *Mycobacterium tuberculosis* and viruses, as well as extracellular agents such as fungi, protozoa, and

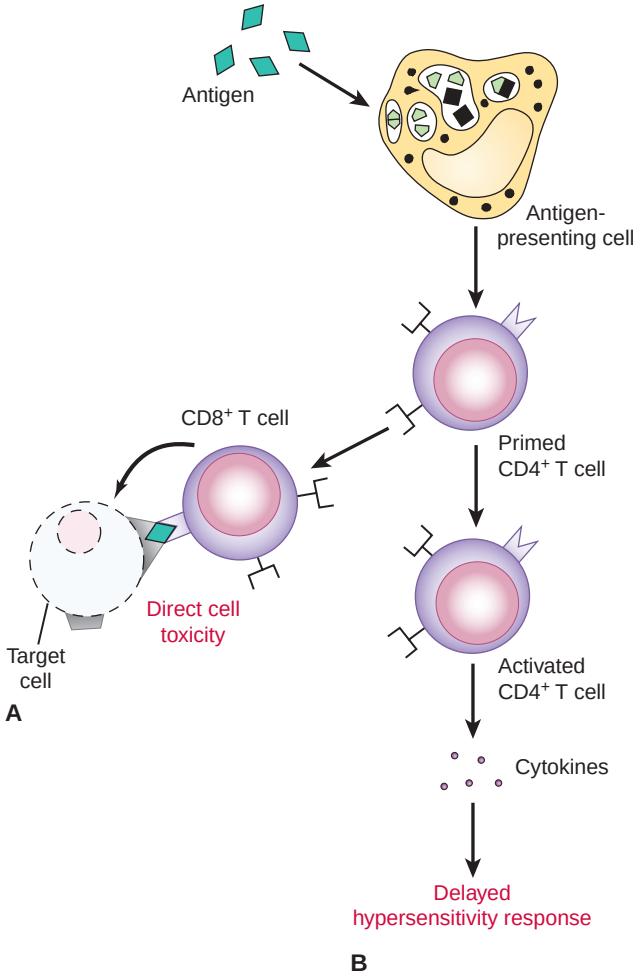


FIGURE 15.5 • Type IV, cell-mediated hypersensitivity reactions, which include (**A**) direct cell-mediated cytotoxicity in which $CD8^+$ T cells kill the antigen-bearing target cells and (**B**) delayed-type hypersensitivity reactions in which presensitized $CD4^+$ cells release cell-damaging cytokines.

parasites. However, it can cause cell death and tissue injury in sensitized people in response to topically administered chemical antigens (contact dermatitis), systemic antigen exposure, or as part of the autoimmune process.

Type IV hypersensitivity reactions are comprised of a spectrum of disorders that range from mild to severe in clinical presentation. Although all are T cell–mediated, the pathophysiologic mechanisms and sensitized T-cell populations involved differ. Because of the heterogeneity of delayed hypersensitivity reactions, current immunology experts subdivide type IV reactions into four distinct subtypes (IVa, IVb, IVc, and IVd) based upon the immune response, T-cell population, and pathologic characteristics involved.[102,103] In addition, depending upon the reaction, different T-cell subsets with different cytotoxic and regulatory functions can be activated at different stages of the disease process.

In type IVa hypersensitivity reactions (*e.g.*, eczema), the $CD4^+$-T_1H cells activate monocytes and macrophages through the secretion of large amounts of interferon (IFN-γ). Activated monocytes stimulate the production of complement-fixing antibodies and activate proinflammatory (*e.g.*, tumor necrosis factor [TNF]-α and interleukin [IL-12]) and $CD8^+$ responses.[102,103] Because type IVa responses require the synthesis of effector molecules, they can take up to 24 to 72 hours to develop, which is why they are called "delayed-type" hypersensitivity disorders.

Type IVb and IVd reactions are also considered to be delayed hypersensitivity reactions. Type IVb reactions (*e.g.*, *maculopapular exanthema and bullous exanthema*) are the result of T_2H cell activation and eosinophilic infiltration of the tissues.[103] T_2H cells secrete the cytokines IL-4 and IL-5, which are necessary for activation of mast cell and eosinophilic responses. In addition, these cytokines deactivate macrophages and promote the production of IgE and IgG antibodies by the B lymphocytes. Type IVd reactions are very rare and involve the recruitment and activation of neutrophils by T lymphocytes that specifically secrete IL-8.[103] The only disorder of this subtype is *acute generalized exanthematous pustulosis* (AGEP), which presents with neutrophil-filled sterile pustules of the skin, fever, and massive leukocytosis.[102]

Type IVc hypersensitivity reactions are cytotoxic responses mediated by $CD4^+$ and $CD8^+$ lymphocytes that secrete perforin and granzyme B.[102] Cytotoxic lymphocytes (CTLs) bind antigen fragments that are displayed on MHC molecules found on the surface of APCs. Peptides derived from cytosolic antigens (*e.g.*, viral) are presented by MHC class I molecules and activate $CD8^+$ T cells, which kill any cell displaying the foreign antigen. Peptides derived from proteins degraded as a result of phagocytic ingestion (*e.g.*, bacteria) are presented on MHC class II molecules, which activate $CD4^+$ T cells. Once activated in this manner, $CD4^+$ T cells can be considered cytotoxic because they are capable of activating other effector cells including cytotoxic $CD8^+$, macrophages, and B lymphocytes.

In viral infections, cell damage is frequently the result of CTL responses rather than cytotoxic effects of the invading organism. While some viruses directly injure infected cells and are said to be cytopathic, other noncytopathic viruses do not. Because CTLs cannot distinguish between cytopathic and noncytopathic viruses, they destroy virtually all cells that are infected regardless of whether or not the virus is dangerous to the cell. In certain forms of hepatitis, for example, the destruction of liver cells is due to the host CTL response and not the virus.

Allergic Contact Dermatitis

Allergic contact dermatitis is a type IV hypersensitivity reaction associated with the activation of T_1H and T helper (17) lymphocytes.[104] The inflammatory response takes place in two phases, sensitization and elicitation. It is usually confined to sites on the skin that have come in direct contact with a hapten (*e.g.,* cosmetics, hair dyes, metals, topical drugs, plant oils).[104] During the sensitization phase, haptens are

captured by dendritic cells, which then migrate to regional lymph nodes and stimulate T cell production. In addition, local keratinocytes sense haptens and initiate and amplify the local immune response. Reexposure to the specific hapten results in rapid recruitment and activation of memory-specific T cells. The most common form of this condition is the dermatitis that follows an intimate encounter with poison ivy or poison oak antigens, although many other substances can trigger a reaction.

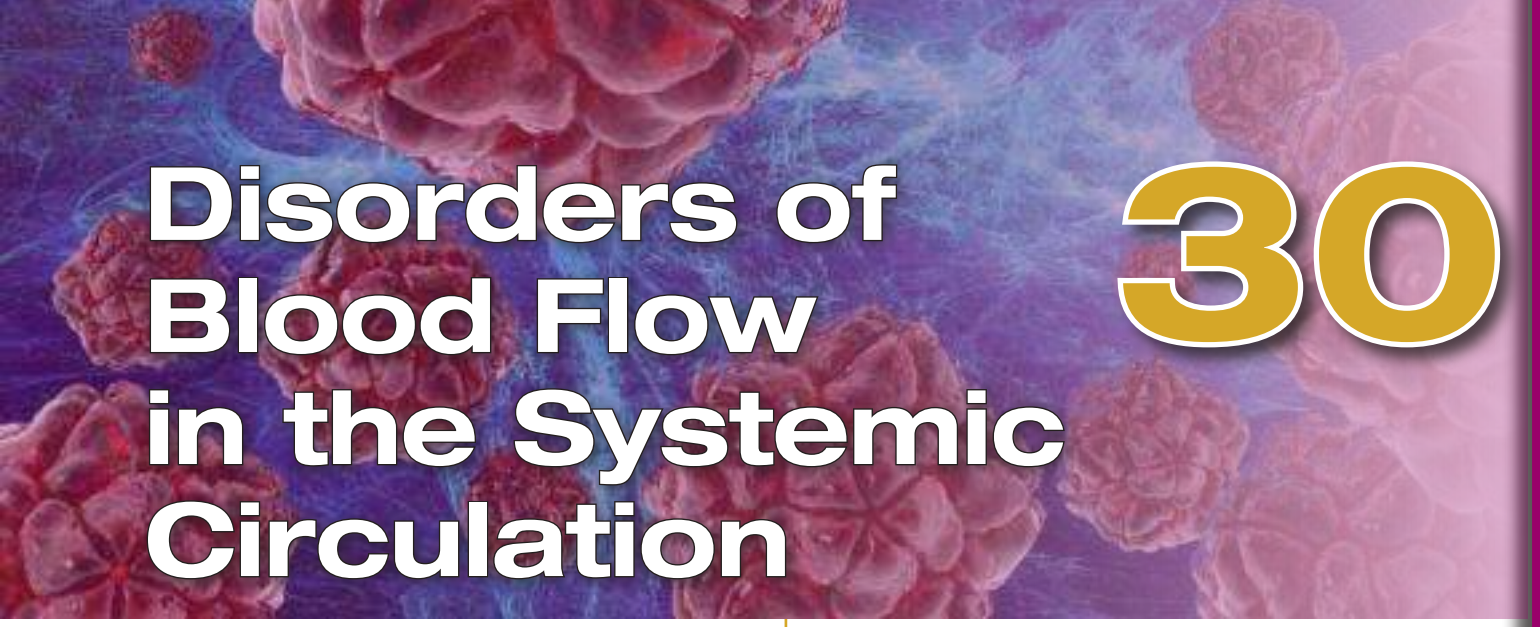
Clinical manifestations of contact dermatitis include an erythematous, papular, and vesicular rash that is associated with intense pruritus and weeping. The affected area often becomes swollen and warm, with exudate formation and crusting. It is not uncommon for a secondary infection to develop. The location of the lesions often provides a clue about the antigen causing the disorder. However, with same cases of contact dermatitis (*e.g.,* poison ivy), the allergen can be unknowingly spread from one part of the body to another. The severity of the reactions ranges from mild to intense, depending on the person and the allergen. Symptoms usually appear approximately 12 to 24 hours after exposure. Depending on the antigen and the duration of exposure, the reaction may last from days to weeks.

Diagnosis of contact dermatitis is made based upon the characteristics and distribution of the rash as well as the temporal relationship of exposure to the suspected allergen. Patch tests can be performed to confirm the diagnosis. Treatment involves removal of the offending agent followed by application of topical preparations (*e.g.,* ointments, corticosteroid creams) to relieve symptomatic skin lesions and prevent secondary bacterial infections. Severe reactions may require the administration of systemic corticosteroid therapy.

Hypersensitivity Pneumonitis

Hypersensitivity pneumonitis, also known as extrinsic allergic alveolitis, is a form of inflammatory lung disease that results from an exaggerated immune response after exposure to a multitude of inhaled organic particles or related occupational antigens.[105] It was first described by Pepys et al.[106] after exposure to moldy grains and hay and termed "farmer's lung." The offending agent was found to be *Actinomyces*, a bacterium commonly found in moldy foliage. The exact pathophysiologic mechanism of hypersensitivity pneumonitis remains unclear, but evidence supports a role for both type III and type IV immune responses.[105] People demonstrate both high levels of antigen-specific serum IgG levels and combined cellular infiltration and granuloma formation.[107] T_1H cells appear to play a critical role in the development of the disease through the production and release of TNF, IFN-γ, IL-12, and IL-18 in lung tissue.[108,109] Symptoms, including labored breathing, dry cough, chills and fever, headache, and malaise, usually begin several hours after exposure and subside within hours after the sensitizing antigens are removed. However, long-term sequelae have been reported.

Diagnosis of hypersensitivity pneumonitis is based upon a good history (occupational and otherwise) of exposure to possible antigens. CAT scans of the chest demonstrate areas of lobar vascularity and the presence of centrilobular nodules.[107] Removal of the offending agent and oral corticosteroids are the only treatments available.



Disorders of Blood Flow in the Systemic Circulation

30

Jaclyn Conelius

BLOOD VESSEL STRUCTURE AND FUNCTION
- Endothelial Cells
- Vascular Smooth Muscle Cells

DISORDERS OF THE ARTERIAL CIRCULATION
- Hyperlipidemia
  - *Classification of Lipoproteins*
  - *Etiology and Pathogenesis of Hyperlipidemia*
  - *Diagnosis of Hyperlipidemia*
  - *Treatment of Hyperlipidemia*
- Atherosclerosis
  - *Etiology and Risk Factors*
  - *Pathogenesis*
  - *Clinical Manifestations*
- Vasculitis
- Polyarteritis Nodosa
  - *Etiology*
  - *Clinical Manifestations*
  - *Diagnosis and Treatment*
- Giant Cell Temporal Arteritis
- Arterial Disease of the Extremities
- Acute Arterial Occlusion
  - *Etiology and Pathogenesis*
  - *Clinical Manifestations*
  - *Diagnosis and Treatment*
- Atherosclerotic Occlusive Disease
  - *Etiology*
  - *Clinical Manifestations*
  - *Diagnosis*
  - *Treatment*
- Thromboangiitis Obliterans
  - *Etiology and Pathogenesis*
  - *Clinical Manifestations*
  - *Diagnosis and Treatment*
- Raynaud Disease and Phenomenon
  - *Etiology and Pathogenesis*
  - *Clinical Manifestations*
  - *Diagnosis and Treatment*
- Aneurysms
- Aortic Aneurysms
  - *Etiology*
  - *Clinical Manifestations*
  - *Diagnosis and Treatment*
- Aortic Dissection
  - *Etiology and Pathogenesis*
  - *Clinical Manifestations*
  - *Diagnosis and Treatment*

DISORDERS OF THE VENOUS CIRCULATION
- Varicose Veins
  - *Etiology and Pathogenesis*
  - *Clinical Manifestations*
  - *Diagnosis and Treatment*
- Chronic Venous Insufficiency
- Venous Thrombosis
  - *Etiology and Pathogenesis*
  - *Clinical Manifestations*
  - *Diagnosis and Treatment*

Blood flow in the arterial and venous systems depends on a system of patent blood vessels and adequate perfusion pressure. Unlike disorders of the respiratory system or central circulation that cause hypoxia and impair oxygenation of tissues throughout the body, the effects of blood vessel disease usually are limited to local tissues supplied by a particular vessel or group of vessels.

With arterial disorders, there is decreased blood flow to the tissues along with impaired delivery of oxygen and nutrients. With venous disorders, there is interference with the outflow of blood and removal of waste products. Disturbances in blood flow can result from pathologic changes in the vessel wall (*i.e.,* atherosclerosis and vasculitis), acute vessel obstruction due to thrombus or embolus, vasospasm (*i.e.,* Raynaud phenomenon), or abnormal vessel dilation (*i.e.,* arterial aneurysms or varicose veins).

BLOOD VESSEL STRUCTURE AND FUNCTION

After completing this section of the chapter, you should be able to meet the following objectives:

- Describe the functions of the endothelial cells and define the term *endothelial dysfunction.*
- Describe the function of vascular smooth muscle and its role in vascular repair.

The heart is the pump of the cardiovascular system. It pumps blood via the blood vessels so that blood is transported throughout the body. The walls of all blood vessels, except the very smallest, are composed of three distinct layers—an outer layer of loosely woven collagen tissue, the *tunica externa,* which is composed of loose connective tissue; a middle layer, the *tunica media,* which consists primarily of circumferentially arranged layers of smooth muscle cells (SMCs); and an inner layer, the *tunica intima,* which consists of a single layer of endothelial cells that line the lumen of the vessel, and the underlying subendothelial connective tissue (Fig. 30.1). Table 30.1 describes the structure and function of the blood vessels. As the main cellular components of the blood vessel wall, the endothelial and smooth muscle cells play an important role in the pathogenesis of many disorders of the arterial circulation. Figure 30.2 illustrates the microanatomy of the vein, artery, and capillary beds.

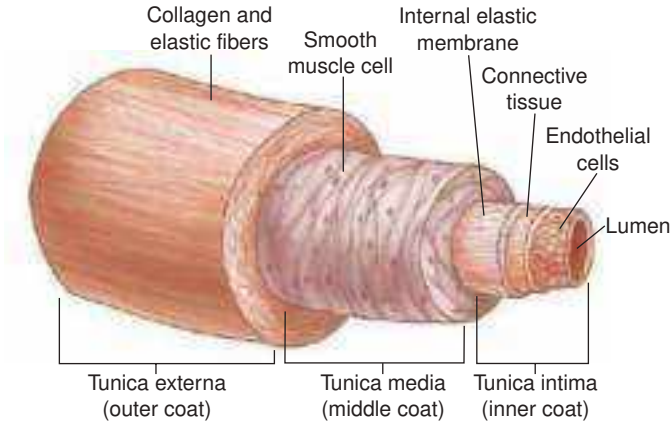


FIGURE 30.1 • Diagram of a typical artery showing the tunica externa, tunica media, and tunica intima.

Endothelial Cells

Endothelial cells form a continuous lining for the entire vascular system called the *endothelium.* The endothelium is made up of approximately 60,000 miles of squamous epithelium that lines the various-sized vessels.[1] Endothelium is a versatile, multifunctional tissue that plays an active role in controlling vascular function.[2] This semipermeable membrane controls the transfer of molecules across the vascular wall and has an essential role in homeostasis. The endothelium also plays a role in the control of platelet adhesion and blood clotting, modulation of blood flow and vascular resistance, metabolism of hormones, regulation of immune and inflammatory reactions, and elaboration of factors that influence the growth of other cell types, particularly vascular SMCs.

Structurally intact endothelial cells respond to various abnormal stimuli by adjusting their usual functions and by expressing newly acquired functions.[1,3] The term *endothelial dysfunction* describes several types of potentially reversible changes in endothelial function that occur in response to environmental stimuli. Inducers of endothelial dysfunction include cytokines and bacterial, viral, and

TABLE 30.1 THE STRUCTURE AND FUNCTION OF THE BLOOD VESSELS

VESSEL	STRUCTURE	FUNCTION
Artery	Three-layered wall with thick tunica media, which gives it the properties of contractility and elasticity	Transport of blood away from the heart, maintenance of blood pressure
Arteriole	Three-layered wall, with much thinner layers and smaller lumen than in arteries	Transport of blood away from the heart, help control blood pressure by regulation of peripheral resistance through vasoconstriction and vasodilation
Capillary	Microscopic size with single-layered wall of endothelium	Thin walls permit the exchange of materials between the blood and interstitial fluid
Venule	Three-layered wall with very thin layers, which gradually enlarge as they near the heart	Transport of blood from capillary beds toward the heart
Vein	Three-layered wall, with thinner tunica media and larger lumen than in arteries. Include internal valves to aid in the unidirectional flow of blood toward the heart	Transport of blood from venules toward the heart

From Wingerd B. (2014). *The human body: Concepts of anatomy and physiology*. Philadelphia, PA: Lippincott Williams & Wilkins.

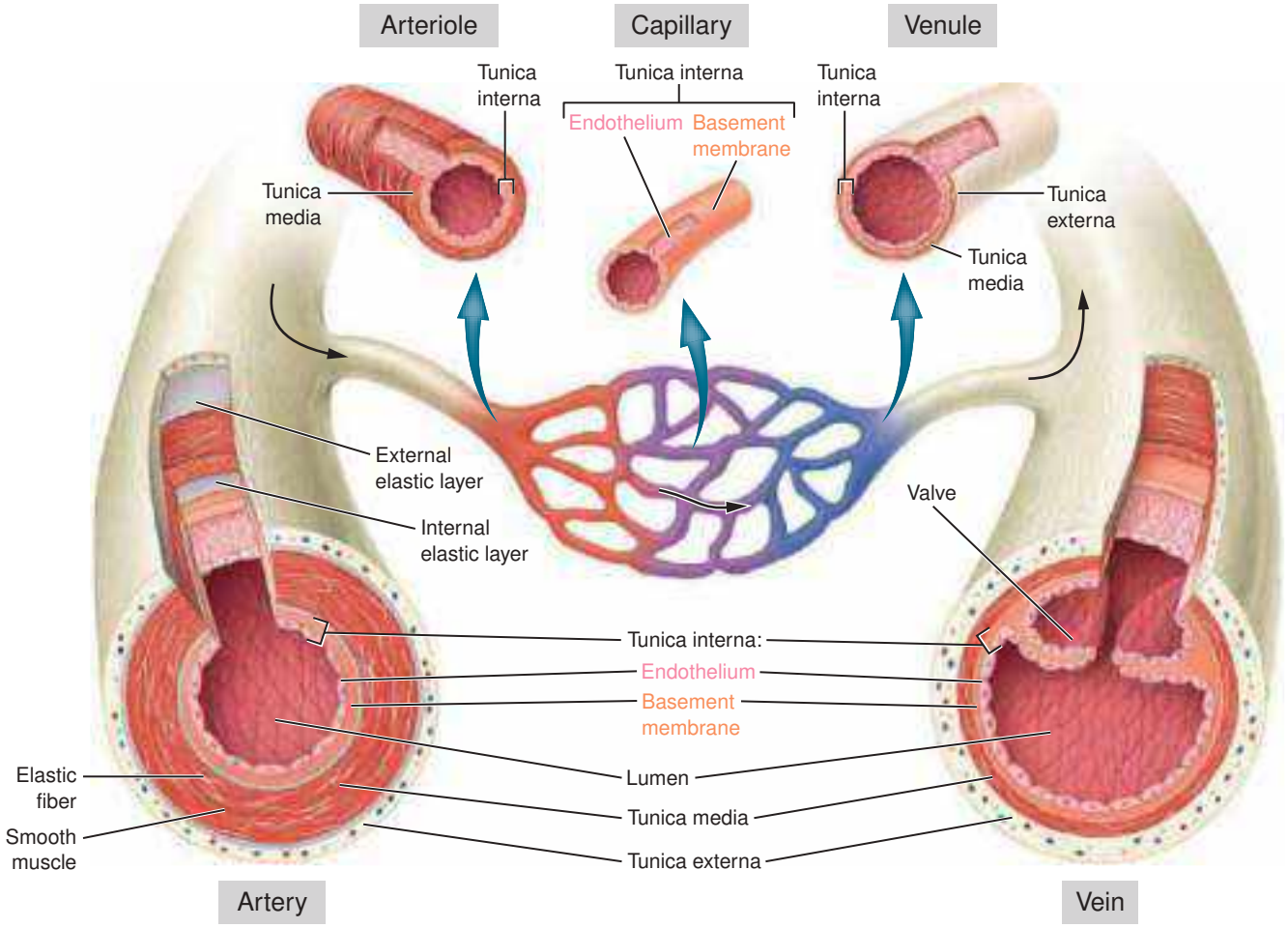


FIGURE 30.2 • Blood vessels—microanatomy of the artery, vein, and capillary beds. (From McConnell T. H., Hull K. L. (2011). *Human form human function: Essentials of anatomy & physiology* (p. 433, Figure 11–12). Philadelphia, PA: Wolters Kluwer/Lippincott Williams & Wilkins.)

parasitic products that cause inflammation; hemodynamic stresses and lipid products that are critical to the pathogenesis of atherosclerosis; and hypoxia. Dysfunctional endothelial cells, in turn, produce other cytokines, growth factors, procoagulant or anticoagulant substances, and a variety of other biologically active products. They also influence the reactivity of underlying SMCs through production of both relaxing factors (*e.g.,* nitric oxide) and contracting factors (*e.g.,* endothelins).

Vascular Smooth Muscle Cells

Vascular SMCs, which form the predominant cellular layer in the tunica media, produce vasoconstriction and/or dilation of blood vessels. A network of vasomotor nerves of the sympathetic component of the autonomic nervous system supplies the smooth muscle in the blood vessels. These nerves and circulating hormones are responsible for vasoconstriction of the vessel walls. Because they do not enter the tunica media of the blood vessel, the nerves do not synapse directly on the SMCs. Instead, they release the neurotransmitter, norepinephrine, which diffuses into the media and acts on the nearby SMCs. The resulting impulses are propagated along the SMCs through their gap junctions, causing contraction of the entire muscle cell layer and thus reducing the radius of the vessel lumen. This in turn increases the systemic circulation.[1]

Vascular SMCs also synthesize collagen, elastin, and other components of the extracellular matrix (ECM); elaborate growth factors and cytokines; and after vascular injury migrate into the intima and proliferate.[3] Thus, SMCs are important in both normal vascular repair and pathologic processes such as atherosclerosis. Growth promoters and inhibitors stimulate the migratory and proliferative activities of vascular SMCs. Promoters include platelet-derived growth factor, thrombin, fibroblast growth factor, and cytokines such as interferon gamma and interleukin-1. Growth inhibitors include nitric oxide. Other regulators include the renin–angiotensin system (angiotensin II) and the catecholamines.

IN SUMMARY

The walls of blood vessels are composed of three layers—an outer layer of loosely woven collagen tissue, a middle layer of vascular smooth muscle, and an inner layer of endothelial cells. The endothelium controls the transfer of molecules across the vascular wall and plays a role in the control of platelet adhesion and blood clotting, modulation of blood flow and vascular resistance, metabolism of hormones, regulation of immune and inflammatory reactions, and elaboration of factors that influence the growth of other cell types, particularly the SMCs. The term *endothelial dysfunction* describes several types of potentially reversible changes in endothelial function that occur in response to environmental stimuli. Vascular SMCs not only control dilation and constriction of blood vessels but elaborate growth factors and synthesize collagen, elastin, and other components of the ECM that are important in both normal vascular repair and pathologic processes such as atherosclerosis.

DISORDERS OF THE ARTERIAL CIRCULATION

After completing this section of the chapter, you should be able to meet the following objectives:

- Describe possible mechanisms involved in the development of atherosclerosis.
- Describe the pathology associated with the vasculitides and relate it to four disease conditions associated with vasculitis.
- Distinguish between the pathology and manifestations of aortic aneurysms and dissection of the aorta.

The arterial system distributes blood to all the tissues in the body. There are three types of arteries—large elastic arteries, including the aorta and its distal branches; medium-sized arteries, such as the coronary and renal arteries; and small arteries and arterioles that pass through the tissues. The large arteries function mainly in transport of blood. The medium-sized arteries are composed predominantly of circular and spirally arranged SMCs. Distribution of blood flow to the various organs and tissues of the body is controlled by contraction and relaxation of the smooth muscle of these vessels. The small arteries and arterioles regulate capillary blood flow. Each of these different types of arteries tends to be affected by different disease processes.

Disease of the arterial system affects body function by impairing blood flow. The effect of impaired blood flow on the body depends on the structures involved and the extent of altered flow. The term *ischemia* denotes a reduction in arterial flow to a level that is insufficient to meet the oxygen demands of the tissues. *Infarction* refers to an area of ischemic necrosis in an organ produced by occlusion of its arterial blood supply or its venous drainage. The discussion in this section focuses on blood lipids and hypercholesterolemia, atherosclerosis, vasculitis, arterial disease of the extremities, and arterial aneurysms.

KEY POINTS

DISORDERS OF THE ARTERIAL CIRCULATION

- Atherosclerosis is a progressive disease characterized by the formation of fibrofatty plaques in the intima of large and medium-sized vessels, including the aorta, coronary arteries, and cerebral vessels. The major risk factors for atherosclerosis are hypercholesterolemia and inflammation.
- Vasculitis is an inflammation of the blood vessel wall resulting in vascular tissue injury and necrosis. Arteries, capillaries, and veins may be affected. The inflammatory process may be initiated by direct injury, infectious agents, or immune processes.
- Aneurysms represent an abnormal localized dilatation of an artery due to a weakness in the vessel wall. As the aneurysm increases in size, the tension in the wall of the vessel increases, and it may rupture. The increased size of the vessel also may exert pressure on adjacent structures.

Hyperlipidemia

Hyperlipidemia is an excess of lipids in the blood. Lipids are classified as triglycerides or neutral fat, phospholipids, and cholesterol. They are a diverse group of compounds that have many key biological functions. Triglycerides, which are used in energy metabolism, are combinations of three fatty acids condensed with a single glycerol molecule. Phospholipids, which contain a phosphate group, are important structural constituents of lipoproteins, blood clotting components, the myelin sheath, and cell membranes. Although cholesterol is not composed of fatty acids, its steroid nucleus is synthesized from fatty acids, and thus, its chemical and physical activity is similar to that of other lipid substances.[2]

Elevated levels of blood cholesterol (*hypercholesterolemia*) are implicated in the development of atherosclerosis with its attendant risk of heart attack and stroke. This is a major public health issue that is underscored by statistics released by the American Heart Association (AHA). An estimated 102.2 million Americans have a serum cholesterol of greater than 200 mg/dL, and 37.7 million Americans have high-risk serum cholesterol levels (240 mg/ dL or greater) that could contribute to a heart attack, stroke, or other cardiovascular event associated with atherosclerosis.[3,4]

Classification of Lipoproteins

Because cholesterol and triglyceride are insoluble in plasma, they are encapsulated by a stabilizing coat of water-soluble phospholipids and proteins (called *apoproteins*). These particles, which are called *lipoproteins,* transport cholesterol and triglyceride to various tissues for energy utilization, lipid deposition, steroid hormone production, and bile acid formation. There are five types of lipoproteins, classified according to their densities as measured by ultracentrifugation: chylomicrons, very–low-density lipoprotein (VLDL), intermediate-density lipoprotein (IDL), low-density lipoprotein (LDL), and high-density lipoprotein (HDL). Chylomicrons are apparent in the blood about 1 hour after a meal and carry primarily triglycerides but also a small amount of phospholipids, cholesterol, and apoprotein B.[6] VLDL carries large amounts of triglycerides that have a lower density than cholesterol. LDL is the main carrier of cholesterol, whereas HDL actually is 50% protein (Fig. 30.3).

Each type of lipoprotein consists of a large molecular complex of lipids combined with apoproteins.[5,6] The major lipid constituents are cholesterol esters, triglycerides, nonesterified (or free) cholesterol, and phospholipids. The insoluble cholesterol esters and triglycerides are located in the hydrophobic core of the lipoprotein macromolecule, surrounded by the soluble phospholipids, nonesterified cholesterol, and apoproteins (Fig. 30.4). Nonesterified cholesterol and phospholipids provide a negative charge that allows the lipoprotein to be soluble in plasma.

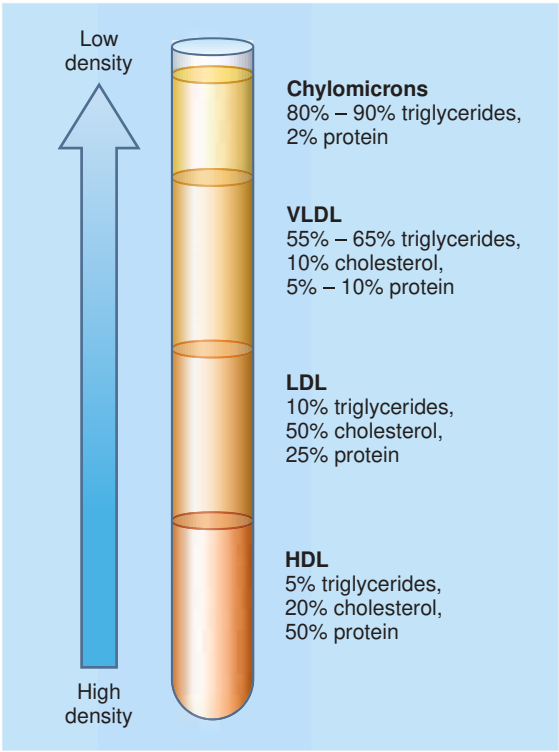


FIGURE 30.3 • Lipoproteins are named based on their protein content, which is measured in density. Because fats are less dense than proteins, as the proportion of triglycerides decreases, the density increases.

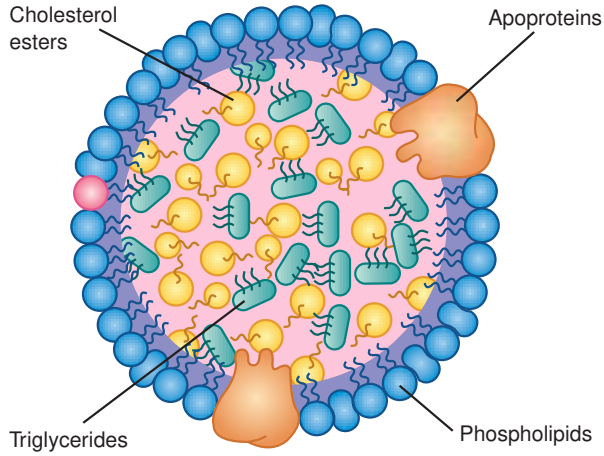


FIGURE 30.4 • General structure of a lipoprotein. The cholesterol esters and triglycerides are located in the hydrophobic core of the macromolecule, surrounded by phospholipids and apoproteins.

There are four major classes of apoproteins: A (*i.e.,* apoA-I, apoA-II, and apoA-IV), B (*i.e.,* apoB-48, apoB-100), C (*i.e.,* apoC-I, apoC-II, and apoC-III), and apoE.[5] The apoproteins control the interactions and ultimate metabolic fate of the lipoproteins. Some of the apoproteins activate the lipolytic enzymes that facilitate the removal of lipids from the lipoproteins. Others serve as a reactive site that cellular receptors can recognize and use in the endocytosis and metabolism of the lipoproteins. The major apoprotein in LDL is apoB-100, whereas in HDL it is apoA-I. Research findings suggest that genetic defects in the apoproteins may be involved in hyperlipidemia and accelerated atherosclerosis.[1,5,6,7]

There are two sites of lipoprotein synthesis—the small intestine and the liver. The chylomicrons, which are the largest of the lipoprotein molecules, are synthesized in the wall of the small intestine. They are involved in the transport of dietary (exogenous pathway) triglycerides and cholesterol that have been absorbed from the gastrointestinal tract. Chylomicrons transfer their triglycerides to the cells of adipose and skeletal muscle tissue. Cholesterol remains in the remnant chylomicron particles after the triglycerides are removed. Ultimately the residual cholesterol is then taken up by the liver, which synthesizes it for the development of VLDL and/or excretes it in bile.[2]

The liver synthesizes and releases VLDL and HDL. The VLDLs contain large amounts of triglycerides and lesser amounts of cholesterol esters.[3] They provide the primary pathway for transport of the endogenous triglycerides produced in the liver, as opposed to those obtained from the diet. They are also the body's main source of energy during prolonged fasting. Like chylomicrons, VLDLs carry their triglycerides to fat and muscle cells, where the triglycerides are removed. The resulting IDL fragments are reduced in triglyceride content and enriched in cholesterol. They are taken to the liver and recycled to form VLDL, or converted to LDL in the vascular compartment. IDLs are the main source of LDL (Fig. 30.5).

LDL, sometimes called the *bad cholesterol,* is the main carrier of cholesterol. LDL is removed from the circulation by

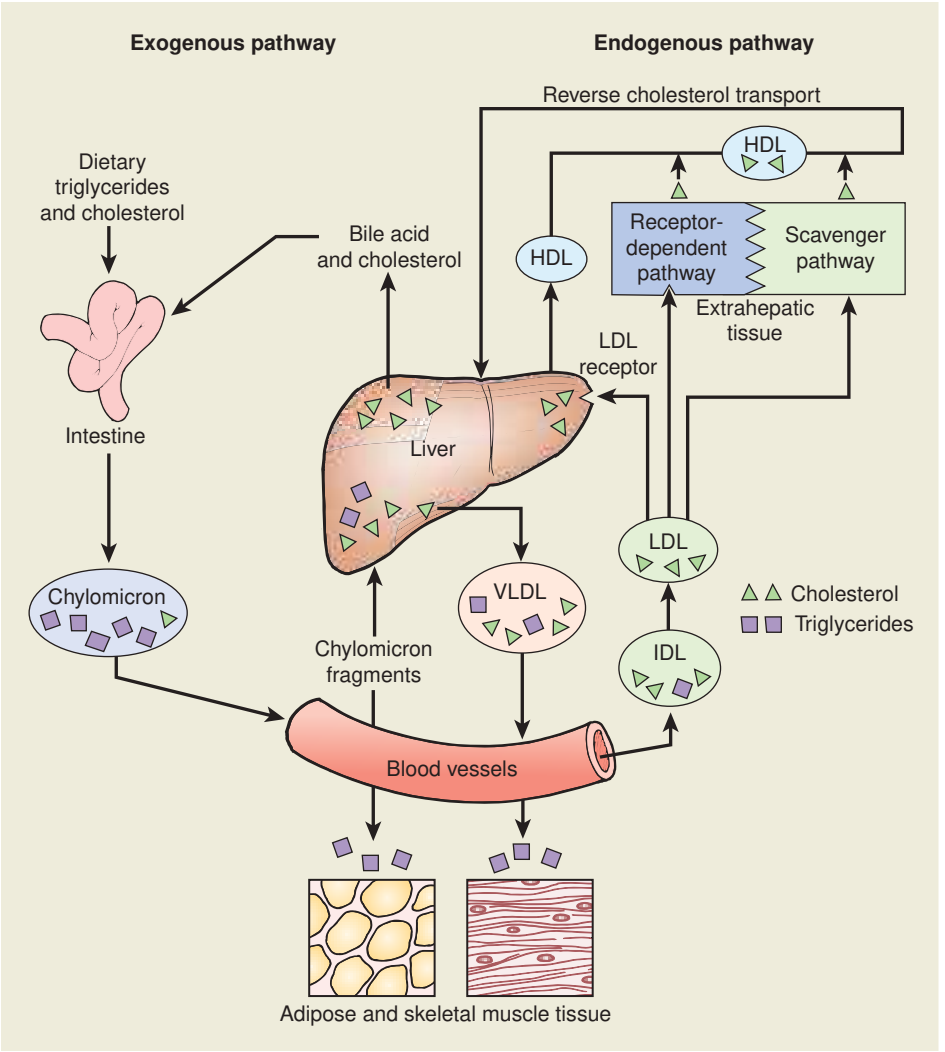


FIGURE 30.5 • Schematic representation of the exogenous and endogenous pathways for triglyceride and cholesterol transport.

either LDL receptors or by scavenger cells such as monocytes or macrophages. Approximately 70% of LDL is removed through the LDL receptor–dependent pathway, and the rest is removed by the scavenger pathway.[3] Although LDL receptors are widely distributed, approximately 75% are located on hepatocytes. Thus, the liver plays an extremely important role in LDL metabolism. LDL receptor–mediated removal involves binding of LDL to cell surface receptors, followed by *endocytosis,* a phagocytic process in which LDL is engulfed and moved into the cell in the form of a membrane-covered endocytic vesicle. Within the cell, the endocytic vesicles fuse with lysosomes, and the LDL molecule is enzymatically degraded, causing free cholesterol to be released into the cytoplasm.

Other, nonhepatic tissues (*i.e.,* adrenal glands, SMCs, endothelial cells, and lymphoid cells) also use the LDL receptor–dependent pathway to obtain cholesterol needed for membrane and hormone synthesis. These tissues can control their cholesterol intake by adding or removing LDL receptors. The scavenger pathway involves ingestion by phagocytic monocytes and macrophages. These scavenger cells have receptors that bind LDL that has been oxidized or chemically modified. The amount of LDL that is removed by the scavenger pathway is directly related to the plasma cholesterol level. When there is a decrease in LDL receptors or when LDL levels exceed receptor availability, the amount of LDL that is removed by scavenger cells is greatly increased. The uptake of LDL by macrophages in the arterial wall can result in the accumulation of insoluble cholesterol esters, the formation of foam cells, and the development of atherosclerosis.

HDL is synthesized in the liver and often is referred to as the *good cholesterol.* HDL participates in the reverse transport of cholesterol by carrying cholesterol from the peripheral tissues back to the liver. Epidemiologic studies have shown an inverse relation between HDL levels and the development of atherosclerosis.[8] It is thought that HDL, which is low in cholesterol and rich in surface phospholipids, facilitates the clearance of cholesterol from the periphery (including atheromatous plaques) and transports it to the liver, where it may be excreted rather than reused in the formation of VLDL (reverse cholesterol transport). The mechanism whereby HDL promotes the movement of cholesterol from peripheral cells to lipid-poor HDL involves a specialized

lipid transporter called the *ATP-binding cassette transporters* (ABCA1 and ABCG1).[8,9] These transporters play a pivotal role in the anti-inflammatory effects of HDL. Defects in this system (resulting from mutations in the ABCA1 transporter) are responsible for Tangier disease, which is characterized by accelerated atherosclerosis and little or no HDL. HDL is also believed to inhibit cellular uptake of LDL by reducing oxidation, thereby preventing uptake of oxidized LDL by the scavenger receptors on macrophages. It has been observed that regular exercise, moderate alcohol consumption, and certain lipid medications increase HDL levels, while smoking and the metabolic syndrome are associated with decreased levels of HDL.[3,8]

Etiology and Pathogenesis of Hyperlipidemia

Serum cholesterol levels may be elevated as a result of an increase in any of the lipoproteins—the chylomicrons, VLDL, IDL, LDL, or HDL. The commonly used classification system for hyperlipidemia is based on the type of lipoprotein involved. Several factors, including nutrition, genetics, medications, comorbid conditions, and metabolic diseases, can raise blood lipid levels. Most cases of elevated levels of cholesterol are probably multifactorial. Some people may have increased sensitivity to dietary cholesterol, others have a lack of LDL receptors, and still others have an altered synthesis of the apoproteins, including oversynthesis of apoB-100, the major apoprotein in LDL.

Hypercholesterolemia (hyperlipoproteinemia) can be classified as either primary or secondary hypercholesterolemia. Primary hypercholesterolemia describes elevated cholesterol levels that develop independent of other health problems or lifestyle behaviors, whereas secondary hypercholesterolemia is associated with other health problems and behaviors.

Many types of primary hypercholesterolemia have a genetic basis. There may be a defective synthesis of the apoproteins, a lack of receptors, defective receptors, or defects in the handling of cholesterol in the cell that are genetically determined.[5] For example, the LDL receptor is deficient or defective in the genetic disorder known as *familial hypercholesterolemia (type 2A).* This autosomal dominant type of hyperlipoproteinemia results from a mutation in the gene specifying the receptor for LDL. Because most of the circulating cholesterol is removed by receptor-dependent mechanisms, blood cholesterol levels are markedly elevated in people with this disorder. The disorder is probably one of the most common of all mendelian disorders. Plasma LDL levels in people with the heterozygote form of the disease range between 250 and 500 mg/dL. However, in people with the homozygote form of the disease, LDL cholesterol levels may rise to 1000 mg/dL. Although people with the heterozygote form of the disease commonly have an elevated cholesterol level from birth, they do not develop symptoms until adult life, when they often develop *xanthomas* (*i.e.,* cholesterol deposits) along the tendons, and atherosclerosis appears (Fig. 30.6). Myocardial infarction is seen in this population, but at a later age (40 to 45 years of age in men) compared to those with the homozygote form of the disease. Those with the homozygote form are much more severely affected; they have cutaneous xanthomas in childhood and may experience myocardial infarction early.[5,10]

Causes of secondary hyperlipoproteinemia include obesity with high-calorie intake and diabetes mellitus. High-calorie diets increase the production of VLDL, with triglyceride elevation and high conversion of VLDL to LDL. Excess ingestion of cholesterol may reduce the formation of LDL receptors and thereby decrease LDL removal. Diets that are high in triglycerides and saturated fats increase cholesterol synthesis and suppress LDL receptor activity.[11]

In diabetes mellitus and the metabolic syndrome, typical dyslipidemia is seen with elevation of triglycerides, low HDL, and minimal or modest elevation of LDL.[8,12,13] Other systemic disorders that can elevate lipids include hypothyroidism,

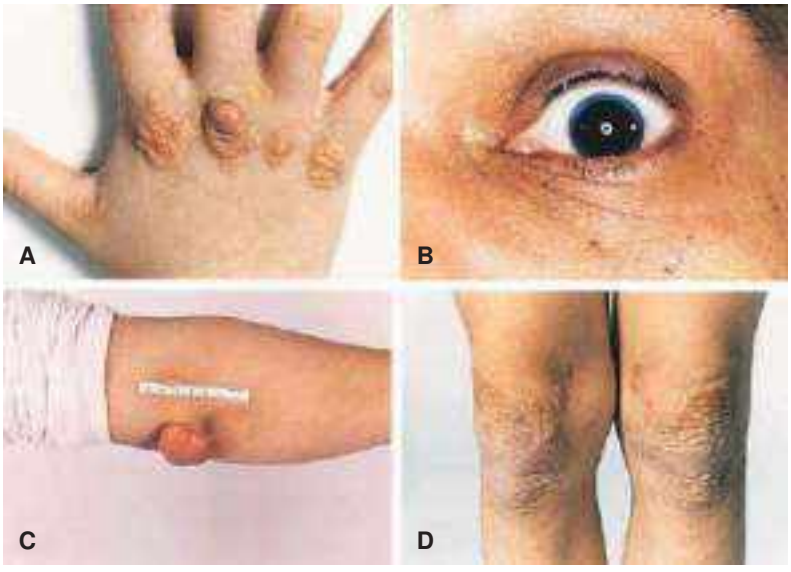


FIGURE 30.6 • Xanthomas in the skin and tendons (**A, C, D**). Arcus lipoides represents the deposition of lipids in the peripheral cornea (**B**). (From Rubin R., Strayer, D. (Eds.). (2012). *Rubin's pathology: Clinicopathologic foundations of medicine* (6th ed., p. 459). Philadelphia, PA: Lippincott Williams & Wilkins.)

nephrotic syndrome, and obstructive liver disease. Medications such as beta-blockers, estrogens, and protease inhibitors (used in the treatment of human immunodeficiency virus [HIV] infection) can also increase lipid levels.

Diagnosis of Hyperlipidemia

The Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults includes a classification system for hyperlipidemia that describes optimal to very high levels of LDL cholesterol, desirable to high levels of total cholesterol, and low and high levels of HDL cholesterol.[12] The NCEP recommends that all adults 20 years of age and older should have a fasting lipoprotein profile (total cholesterol, LDL cholesterol, HDL cholesterol, and triglycerides) measured once every 5 years. Normal ranges can be found at the NCEP website at http://old.nhlbi.nih.gov/chd/index.htm. If testing is done in the nonfasting state, only the total cholesterol and HDL are considered useful. A follow-up lipoprotein profile should be done on people with nonfasting total cholesterol levels of 200 mg/dL or more or HDL levels lower than 40 mg/dL. Lipoprotein measurements are particularly important in people at high risk for developing coronary heart disease (CHD) since there are few if any early clinical manifestations of hyperlipidemia.

Treatment of Hyperlipidemia

The NCEP continues to identify reduction in LDL cholesterol as the primary target for cholesterol-lowering therapy, particularly in people at risk for CHD. The major risk factors for CHD, exclusive of LDL cholesterol levels, that modify LDL cholesterol goals include cigarette smoking, hypertension, family history of premature CHD in a first-degree relative, age (men ≥45 years; women ≥55 years), and an HDL cholesterol level less than 40 mg/dL. Accordingly, the NCEP has updated the 2001 guidelines for management of LDL cholesterol based on risk factors.[12] The updated guidelines recommend that persons with zero or no major risk factors should have an LDL cholesterol goal of 160 mg/dL or less; those with two or more of the major risk factors should have an LDL cholesterol goal of less than 130 mg/dL; people with *high-risk* factors (*i.e.,* those with CHD, other forms of atherosclerotic disease, or diabetes) should have an LDL cholesterol goal of less than 100 mg/dL; and persons with *very–high-risk* factors (*i.e.,* acute coronary syndromes or CHD with other risk factors) should have an LDL cholesterol of less than 70 mg/dL.[13] The guidelines also recommend that people with a greater than 20% 10-year risk of experiencing myocardial infarction or coronary death, as determined by the risk assessment tool developed from Framingham Heart Study data, should have an LDL cholesterol goal of less than 100 mg/dL. (To calculate a risk score, see www.nhlbi.nih.gov/guidelines/cholesterol.)

The management of hypercholesterolemia focuses on dietary and therapeutic lifestyle changes; when these are unsuccessful, pharmacologic treatment may be necessary. Therapeutic lifestyle changes include an increased emphasis on physical activity, dietary measures to reduce LDL cholesterol levels, smoking cessation, and weight reduction for people who are overweight.

Several dietary elements affect cholesterol and its lipoprotein fractions: (1) excess calorie intake, (2) saturated and trans fats, and (3) cholesterol. Excess calories consistently lower HDL and less consistently elevate LDL. Saturated fats in the diet can strongly influence cholesterol levels. Each 1% of saturated fat relative to caloric intake increases the cholesterol level an average of 2.8 mg/dL.[12,13] Depending on individual differences, it raises the VLDL and the LDL. Trans fats, which are manufactured from vegetable oils and are used to enhance the taste and extend the shelf life of fast foods, are more atherogenic than saturated fats. Dietary cholesterol tends to increase LDL cholesterol. On average, each 100 mg of ingested cholesterol raises the serum cholesterol 8 to 10 mg/dL.[13]

The aim of dietary therapy is to reduce total and LDL cholesterol levels and increase HDL cholesterol by reduction in total calories and to reduce the percentage of total calories from saturated fat and cholesterol. The AHA has issued new dietary guidelines that focus on an overall plan of healthy food choices and increased physical activity to decrease the risk for development of cardiovascular disease (CVD).[14] The specific guidelines are intended to assist the general public in the maintenance of a body mass index lower than 25 (weight in kilograms divided by body surface area in square meters), to achieve and maintain a low total cholesterol and LDL and a high HDL, and to maintain a blood pressure within normal limits. In general, the dietary guidelines emphasize an increased intake of fruits, vegetables, and fish and a decreased intake of fat, cholesterol, sugars, alcohol, and salt. For people who already have an elevated LDL, the AHA recommends that saturated fat be restricted to less than 7% of the total daily intake, trans fat to less than 1% of the total daily intake, and cholesterol to less than 300 mg/day.[14] However, even with strict adherence to the diet, drug therapy is usually necessary. Clinical data suggest that drug therapy may be efficacious even for those with normal LDL cholesterol since some of the cardioprotective effects of the statin drugs are not just related to LDL lowering, but to their anti-inflammatory effects.[6]

Lipid-lowering drugs work in several ways, including decreasing cholesterol production, decreasing cholesterol absorption from the intestine, and removing cholesterol from the bloodstream. Drugs that act directly to decrease cholesterol levels also have the beneficial effect of further lowering cholesterol levels by stimulating the production of additional LDL receptors. Unless lipid levels are severely elevated, it is recommended that a minimum of 3 months of intensive diet therapy be undertaken before drug therapy is considered.[12,13] However, certain high-risk groups (*e.g.,* people with diabetes who are at increased cardiovascular risk) are now started on statin therapy at the same time as therapeutic lifestyle changes are initiated.[15,16]

There currently are five major types of medications available for treating hypercholesterolemia: HMG CoA reductase inhibitors (statins), bile acid–binding resins, cholesterol absorption inhibitor agents, niacin, and the fibrates.[10]

Inhibitors of HMG CoA reductase (*e.g.,* atorvastatin, rosuvastatin, simvastatin), a key enzyme in the cholesterol biosynthetic pathway, can reduce or block the hepatic synthesis of cholesterol and are the cornerstone of LDL-reducing therapy. Statins also reduce triglyceride levels and increase HDL levels. Statin therapy has been shown to reduce the risk for acute coronary syndromes and stoke in secondary prevention.[12]

The bile acid–binding resins (*e.g.,* cholestyramine, colestipol, colesevelam) bind and sequester cholesterol-containing bile acids in the intestine. This leads to increased production of LDL receptors by the liver, with resulting increased removal of cholesterol from the blood for synthesis of new bile acids. These agents are typically used as adjuncts to statin therapy for patients requiring further reductions in LDL and a 3% to 5% increase in HDL cholesterol.

Nicotinic acid, a niacin congener, blocks the synthesis and release of VLDL by the liver, thereby lowering not only VLDL levels but also IDL and LDL levels. Nicotinic acid also increases HDL concentrations up to 15% to 35%.[12] The fibrates (*e.g.,* fenofibrate and gemfibrozil) also decrease the synthesis of VLDL by the liver, but also enhance the clearance of triglycerides from the circulation resulting in a triglyceride decrease of 20% to 50%.

Atherosclerosis

Atherosclerosis is a type of arteriosclerosis or hardening of the arteries. The term *atherosclerosis,* which comes from the Greek words *atheros* ("gruel" or "paste") and *sclerosis* ("hardness"), denotes the formation of fibrofatty lesions in the intimal lining of the large- and medium-sized arteries such as the aorta and its branches, the coronary arteries, and the large vessels that supply the brain (Fig. 30.7).

Although there has been a gradual decline in deaths from atherosclerosis over the past several decades, one complication of atherosclerosis, CVD, remains the leading cause of death among men and women in the United States.[3] The reported decline in death rate probably reflects new and improved methods of medical treatment and improved health care practices resulting from an increased public awareness of the factors that predispose to the development of this disorder. In 2011, the major complications of atherosclerosis, including ischemic heart disease, stroke, and peripheral vascular disease, accounted for approximately 33.6% of the deaths in the United States.[17]

Atherosclerosis begins as an insidious process, and clinical manifestations of the disease typically do not become evident for 20 to 40 years or longer. Fibrous plaques commonly begin to appear in the arteries of Americans in their third decade.

Etiology and Risk Factors

The major risk factor for atherosclerosis is hypercholesterolemia, which can be modified. Other risk factors, such as increasing age, family history of premature CHD, and male sex, cannot be changed. The tendency toward the development of atherosclerosis appears to run in families. People who come from families with a strong history of heart disease or stroke due to atherosclerosis are at greater risk for developing atherosclerosis than those with a negative family history. Several genetically determined alterations in lipoprotein and cholesterol metabolism have been identified, and it seems likely that others will be identified in the future.[5] The incidence of atherosclerosis increases with age. Other factors being equal, men are at greater risk for development of CVD than are premenopausal women, probably because of the protective effects of natural estrogens. After menopause, the incidence of atherosclerosis-related diseases in women increases, and the frequency of myocardial infarction in the two sexes tends to equalize.[3]

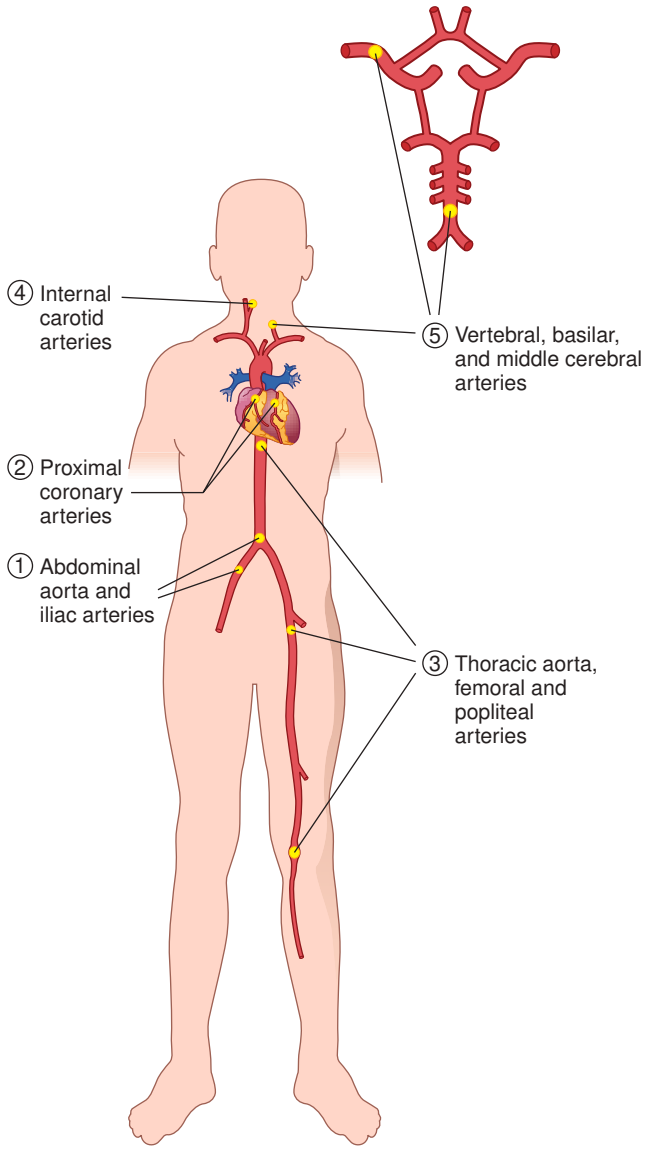


FIGURE 30.7 • Sites of severe atherosclerosis in order of frequency. (From Rubin R., Strayer D. (Eds.). (2012). *Rubin's pathology: Clinicopathologic foundations of medicine* (6th ed., p. 452). Philadelphia, PA: Lippincott Williams & Wilkins.)

The major risk factors for atherosclerosis that can be affected by a change in health care behaviors include high blood cholesterol levels (specifically high LDL cholesterol levels), cigarette smoking, obesity and visceral fat, hypertension, and diabetes mellitus (traditional cardiovascular risk factors). Cigarette smoking is closely linked with CVD and sudden death. Endothelial damage may be worsened by cigarette smoke. Prolonged smoking of years of one pack or more per day doubles the damage to the endothelium. However, stopping smoking reduces the risk of endothelial damage significantly.[3]

Hypertension or high blood pressure increases the risk of atherosclerotic coronary artery disease by twofold. Type 2 diabetes mellitus increases the risk for disease greater then twofold. When a person has hypertension *and* type 2 diabetes, his or her risk for atherosclerotic coronary artery disease increases by eightfold.[1]

However, not all atherothrombotic vascular disease can be explained by the established genetic and environmental risk factors. Other, so-called nontraditional, cardiovascular risk factors can be associated with an increased risk for development of atherosclerosis, including C-reactive protein (CRP), serum homocysteine, serum lipoprotein(a), and infectious agents.[5,11]

Considerable interest in the role of inflammation in the etiology of atherosclerosis has emerged over the last few years.[10,17,18] In particular, CRP is now considered a major risk factor marker.[19,20] CRP is a serum marker for systemic inflammation (see Chapter 14). Several prospective studies have indicated that elevated CRP levels are associated with vascular disease. The pathophysiologic role of CRP in atherosclerosis has not yet been defined. High-sensitivity CRP (hs-CRP) may be a better predictor of cardiovascular risk than lipid measurement alone.[20] Furthermore, greater than 75% of cardiovascular events occur in women with a normal LDL (<160mg/dL).[3] In the Heart Protection Study, statin therapy decreased cardiovascular complications even in people with a normal LDL. This was thought to be due to the anti-inflammatory effects of these agents. Inflammation (as assessed by a decrease in hs-CRP) can be reduced by using certain lifestyle changes (exercise and reducing stress) and by drugs (including statins, fibrates, and thiazolidinediones). Serum hs-CRP levels of less than 1, 1 to 3, and over 3 mg/L correspond, respectively, to low-, moderate-, and high-risk groups for future cardiovascular events.[20] In most clinical settings, a single hs-CRP assessment is likely to be adequate as long as levels less than 10 mg/L are observed. Because CRP is an acute inflammatory phase reactant, major infections, trauma, or acute hospitalization can elevate CRP levels (usually 100-fold or more). Thus, CRP levels to determine cardiovascular risk should be performed when the person is clinically stable. If the level remains markedly elevated, an alternative source of systemic inflammation should be considered.[20]

Homocysteine is derived from the metabolism of dietary methionine, an amino acid that is abundant in animal protein. The normal metabolism of homocysteine requires adequate levels of folate, vitamin B_6, vitamin B_{12}, and riboflavin. Homocysteine inhibits elements of the anticoagulant cascade and is associated with endothelial damage, which is thought to be an important first step in the development of atherosclerosis.[3,5] However, supplementation with folic acid, vitamin B_6, and vitamin B_{12} to decrease plasma homocysteine levels is not generally recommended for either primary or secondary prevention of CVD based on recent clinical evidence.

Lipoprotein(a) is similar to LDL in composition and is an independent risk factor for the development of premature CHD lipoprotein(a) and can cause atherosclerosis by binding to macrophages through a high-affinity receptor that promotes foam cell formation and the deposition of cholesterol in atherosclerotic plaques. Lipoprotein(a) levels should be determined in people who have premature coronary artery disease or a positive family history since they are not altered by traditional cholesterol-lowering drugs.[18] Lipoprotein levels have been shown to be reduced with the use of nicotinic therapy.[5] The desirable level is less than 14 mg/dL.

There also has been increased interest in the possible connection between infectious agents (*e.g., Chlamydia pneumoniae,* herpesvirus, cytomegalovirus) and the development of vascular disease. The presence of these organisms in atheromatous lesions has been demonstrated by immunocytochemistry, but no cause-and-effect relationship has been established. The organisms may play a role in atherosclerotic development by initiating and enhancing the inflammatory response.[2]

Pathogenesis

The lesions associated with atherosclerosis are of three types—the fatty streak, the fibrous atheromatous plaque, and the complicated lesion. The latter two are responsible for the clinically significant manifestations of the disease.

Fatty streaks are thin, flat, yellow intimal discolorations that progressively enlarge by becoming thicker and slightly elevated as they grow in length. Histologically, they consist of macrophages and SMCs that have become distended with lipid to form foam cells. Fatty streaks are present in children, often in the first year of life.[1,3] This occurs regardless of geographic setting, sex, or race. They increase in number until about 20 years of age, and then they remain static or regress. Damage to the endothelium is an early marker that can later become atherosclerotic. Once the endothelium is damaged circulating monocytes and lipids begin to adhere to the area. This *fibrous atheromatous plaque* is characterized by the gray to pearly white appearance due to the macrophages that ingest and oxidize accumulated lipoproteins and form a visible fatty steak. Over time the fatty steaks grow larger and proliferate in to the smooth muscle. As the lesions increase in size, they encroach on the lumen of the artery. The macrophages release substances that cause inflammation and eventually may occlude the vessel or predispose to thrombus formation, causing a reduction of blood flow (Fig. 30.8).[2] Because blood flow is related to the fourth power of the vessel radius, the reduction in blood flow becomes increasingly greater as the disease progresses.

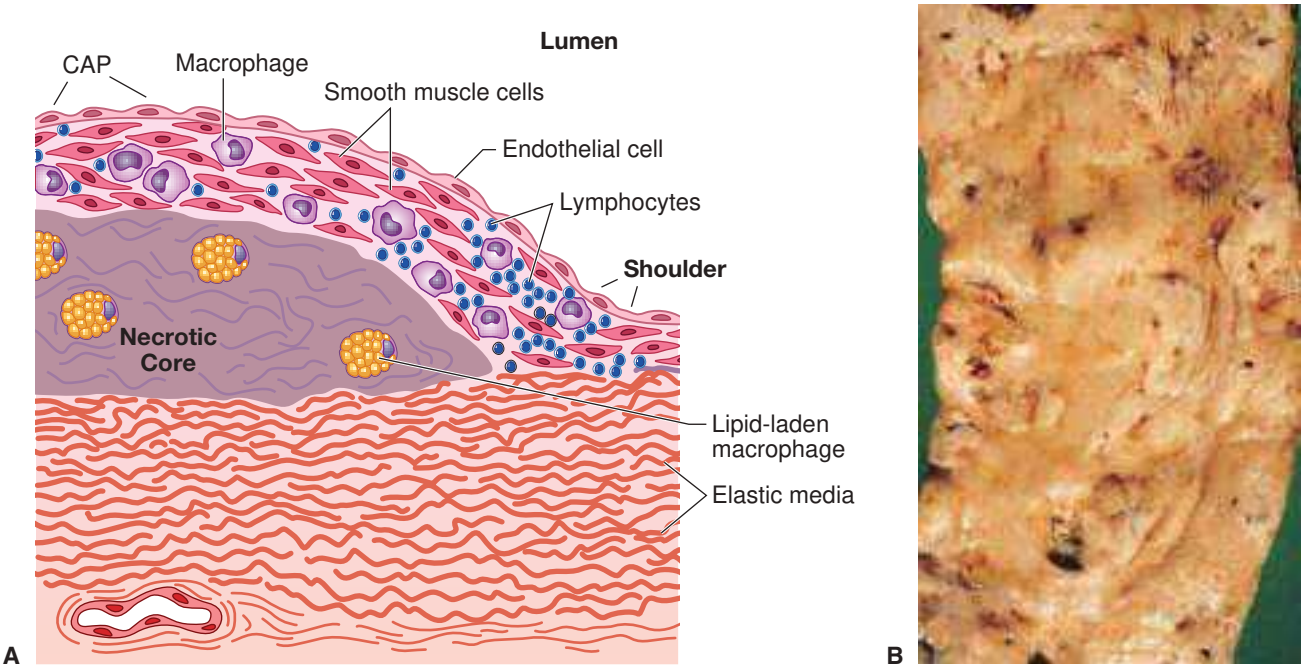


FIGURE 30.8 • Fibrofatty plaque of atherosclerosis. (**A**) In this fully developed fibrous plaque, the core contains lipid-filled macrophages and necrotic smooth muscle cell (SMC) debris. The "fibrous" cap is composed largely of SMCs, which produce collagen, small amounts of elastin, and glycosaminoglycans. Also shown are infiltrating macrophages and lymphocytes. Note that the endothelium over the surface of the fibrous cap frequently appears intact. (**B**) The aorta shows discrete raised, tan plaques. Focal plaque ulcerations are also evident. (From Rubin R., Strayer D. (Eds.). (2012). *Rubin's pathology: Clinicopathologic foundations of medicine* (6th ed., p. 447–448). Philadelphia, PA: Lippincott Williams & Wilkins.)

The more advanced complicated lesions contain hemorrhage, ulceration, and scar tissue deposits. Thrombosis is the most important complication of atherosclerosis. Slowing and turbulence of blood flow in the region of the plaque and ulceration of the plaque cause it. The thrombus may cause occlusion of small vessels in the heart and brain. In addition, aneurysms may develop in arteries weakened by extensive plaque formation.

Although the risk factors associated with atherosclerosis have been identified through epidemiologic studies, many unanswered questions remain regarding the mechanisms by which these risk factors contribute to the development of atherosclerosis. The vascular endothelial layer, which consists of a single layer of cells with cell-to-cell attachments, normally serves as a selective barrier that protects the subendothelial layers by interacting with blood cells and other blood components. One hypothesis of plaque formation suggests that injury to the endothelial vessel layer is the initiating factor in the development of atherosclerosis.[3,5] A number of factors are regarded as possible injurious agents, including products associated with smoking, immune mechanisms, and mechanical stress such as that associated with hypertension. The fact that atherosclerotic lesions tend to form where vessels branch or where there is turbulent flow suggests that hemodynamic factors also play a role.

Hyperlipidemia, particularly LDL with its high cholesterol content, is also believed to play an active role in the pathogenesis of the atherosclerotic lesion. Interactions between the endothelial layer of the vessel wall and white blood cells, particularly the monocytes (blood macrophages), normally occur throughout life; these interactions increase when blood cholesterol levels are elevated. One of the earliest responses to elevated cholesterol levels is the attachment of monocytes to the endothelium.[5] The monocytes have been observed to emigrate through the cell-to-cell attachments of the endothelial layer into the subendothelial spaces, where they are transformed into macrophages.

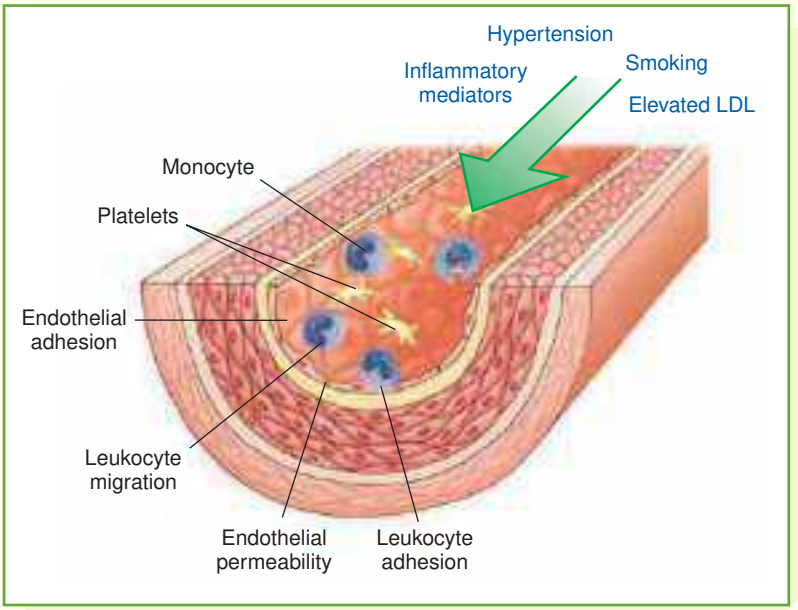
Activated macrophages release free radicals that oxidize LDL. Oxidized LDL is toxic to the endothelium, causing endothelial loss and exposure of the subendothelial tissue to blood components. This leads to platelet adhesion and aggregation and fibrin deposition. Platelets and activated macrophages release various factors that are thought to promote growth factors that modulate the proliferation of SMCs and deposition of ECM in the lesions.[3,5] Activated macrophages also ingest oxidized LDL (by uptake through the scavenger receptor) to become foam cells, which are present in all stages of atherosclerotic plaque formation. Lipids released from necrotic foam cells accumulate to form the lipid core of unstable plaques. Unstable plaques typically are characterized histologically by a large central lipid core, inflammatory infiltrate, and a thin fibrous cap.[21] These "vulnerable plaques" are at risk of rupture (plaque rupture), often at the shoulder of the plaque (see Fig. 30.8A) where the fibrous cap is thinnest (because of the presence of local inflammatory cells and mediators that degrade the cap) and the mechanical stresses highest.[19]

Understanding The Development of Atherosclerosis

Atherosclerosis is characterized by the development of atheromatous lesions within the intimal lining of the large and medium-sized arteries that protrude into and can eventually obstruct blood flow. The development of atherosclerotic lesions is a progressive process involving (1) endothelial cell injury, (2) migration of inflammatory cells, (3) SMC proliferation and lipid deposition, and (4) gradual development of the atheromatous plaque with a lipid core.

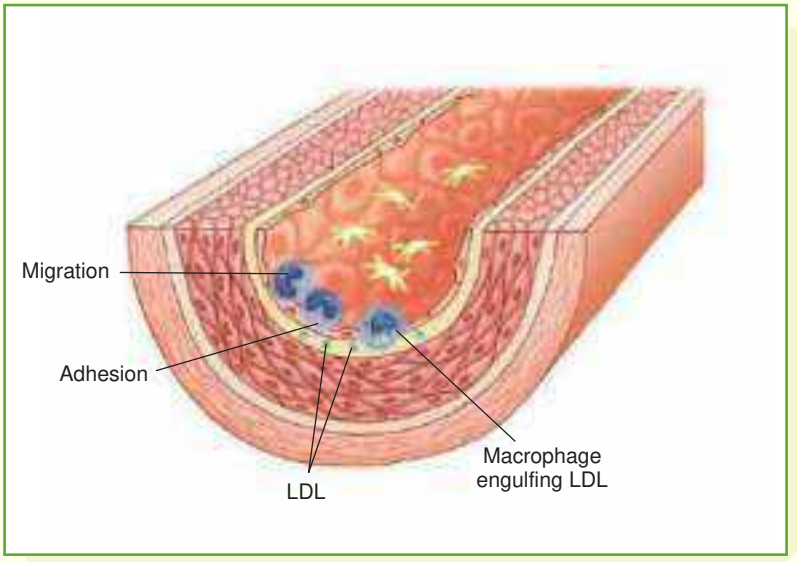
Endothelial Cell Injury

The vascular endothelium consists of a single layer of cells with cell-to-cell attachments, which normally protects the subendothelial layers from interacting with blood cells and other blood components. Agents such as smoking, elevated LDL levels, immune mechanisms, and mechanical stress associated with hypertension share the potential for causing endothelial injury with adhesion of monocytes and platelets.



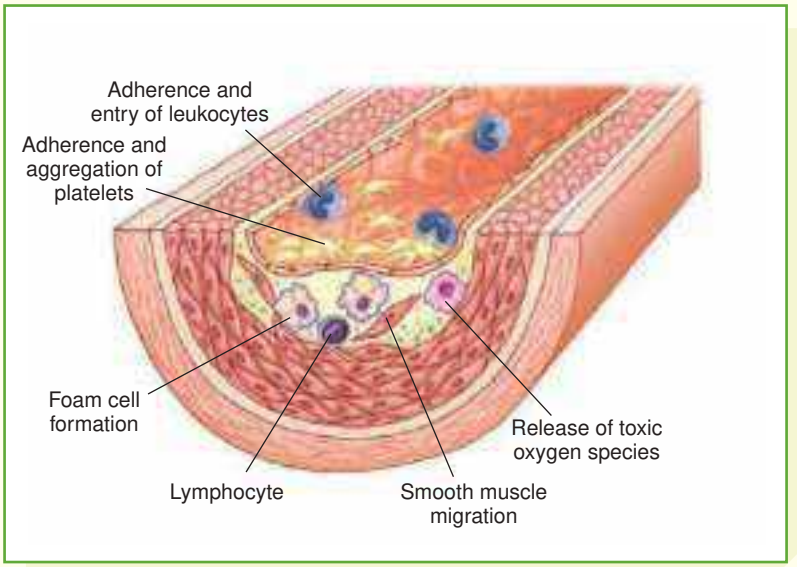
Migration of Inflammatory Cells

Early in the development of atherosclerotic lesions, endothelial cells begin to express selective adhesion molecules that bind monocytes and other inflammatory cells that initiate the atherosclerotic lesions. After monocytes adhere to the endothelium, they migrate between the endothelial cells to localize in the intima, transform into macrophages, and engulf lipoproteins, largely LDL.



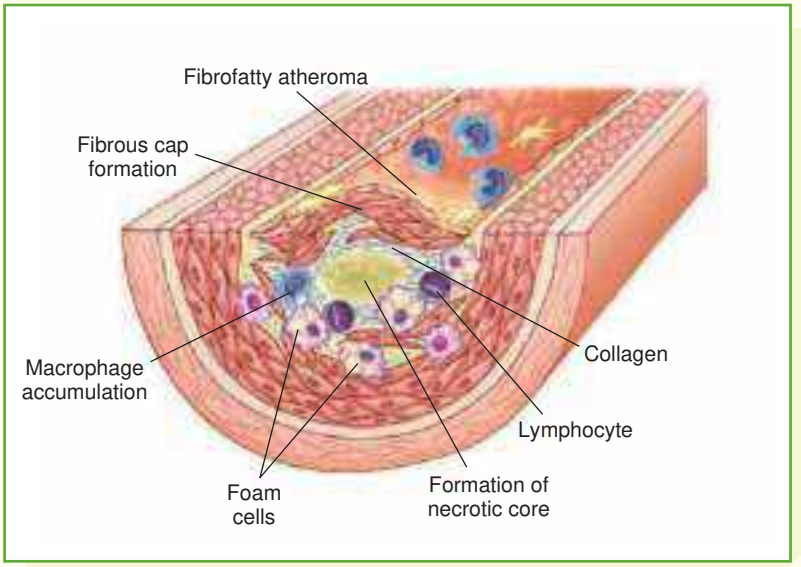
Lipid Accumulation and Smooth Muscle Cell Proliferation

Although the recruitment of monocytes, their differentiation into macrophages and subsequent ingestion of lipids, and their ultimate transformation into foam cells is protective in that it removes excess lipids from the circulation, progressive accumulation eventually leads to lesion progression. Activated macrophages release toxic oxygen species that oxidize LDL; they then ingest the oxidized LDL to become foam cells. They also produce growth factors that contribute to the migration and proliferation of SMCs and the elaboration of ECM.



Plaque Structure

Atherosclerotic plaques consist of an aggregation of SMCs, macrophages, and other leukocytes; ECM, including collagen and elastic fibers; and intracellular and extracellular lipids. Typically, the superficial fibrous cap is composed of SMCs and dense ECM. Immediately beneath and to the side of the fibrous cap is a cellular area (the shoulder) consisting of macrophages, SMCs, and lymphocytes. Below the fibrous cap is a central core of lipid-laden foam cells and fatty debris. Rupture, ulceration, or erosion of an unstable or vulnerable fibrous cap may lead to hemorrhage into the plaque or thrombotic occlusion of the vessel lumen.



Clinical Manifestations

Atherosclerosis begins as an insidious process, and clinical manifestations of the disease typically do not become evident for 20 to 40 years or longer. Fibrous plaques commonly begin to appear in the arteries of Americans in their third decade. The clinical manifestations of atherosclerosis depend on the vessels involved and the extent of vessel obstruction.

Atherosclerotic plaques (lesions) produce their effects through

- Narrowing of the vessel and production of ischemia
- Sudden vessel obstruction due to plaque hemorrhage or rupture
- Thrombosis and formation of emboli resulting from damage to the vessel endothelium
- Aneurysm formation due to weakening of the vessel wall[3]

In larger vessels, such as the aorta, the important complications are those of thrombus formation and weakening of the vessel wall. In medium-sized arteries, such as the coronary and cerebral arteries, ischemia and infarction due to vessel occlusion are more common. Although atherosclerosis can affect any organ or tissue, the arteries supplying the heart, brain, kidneys, lower extremities, and small intestine are most frequently involved.

Vasculitis

The vasculitides are a group of vascular disorders that cause inflammatory injury and necrosis of the blood vessel wall (*i.e.,* vasculitis). The vasculitides, which are a common pathway for tissue and organ involvement in many different disease conditions, involve the endothelial cells and SMCs of the vessel wall.[3] Vessels of any type (arteries, veins, and capillaries) in virtually any organ can be affected. Because they may affect veins and capillaries, the terms *vasculitis, angiitis,* and *arteritis* often are used interchangeably. Clinical manifestations often include fever, myalgia, arthralgia, and malaise. Vasculitis may result from direct injury to the vessel, infectious agents, or immune processes, or they may be secondary to other disease states such as systemic lupus erythematosus. Physical agents such as cold (*i.e.,* frostbite), irradiation (*i.e.,* sunburn), mechanical injury, immune mechanisms, and toxins may secondarily cause vessel damage, often leading to necrosis of the vessels. Small vessel vasculitides are sometimes associated with *antineutrophil cytoplasmic antibodies* (ANCA). ANCA are antibodies directed against certain proteins in the cytoplasm of neutrophils. These autoantibodies may cause endothelial damage.[3] Serum ANCA titers, which can correlate with disease activity, may serve as a useful quantitative diagnostic marker for these disorders.

The vasculitides are commonly classified based on etiology, pathologic findings, and prognosis. One classification system divides the conditions into three groups: (1) small vessel, (2) medium-sized vessel, and (3) large vessel vasculitides[3,5] (Table 30.2). Small vessel refers to small arteries (ANCA-associated disease only), arterioles, venules, and capillaries; medium vessels refer to medium- and small-sized arteries and arterioles; and large vessel refers to the aorta and its major tributaries. The small vessel vasculitides are involved in a number of different diseases, most of which are mediated by type III immune complex hypersensitivity reaction. They commonly involve the skin and are often a complication of an underlying disease (*i.e.,* vasculitis associated with neoplasms or connective tissue disease) and exposure to environmental agents (*i.e.,* serum sickness and urticarial vasculitis). ANCA-positive small vessel vasculitis includes microscopic polyangiitis, Wegener granulomatosis, and the Churg-Strauss syndrome. These ANCA-positive vasculitides are treated by similar regimens.[22]

Medium-sized vessel vasculitides produce necrotizing damage to medium-sized muscular arteries of major organ systems. This group includes polyarteritis nodosa, Kawasaki disease, and thromboangiitis obliterans. Large vessel vasculitides involve large elastic arteries. They include giant cell (temporal) arteritis, polymyalgia rheumatica, and Takayasu arteritis. The following discussion focuses on two of the vasculitides: polyarteritis nodosa and giant cell (temporal) arteritis.

Polyarteritis Nodosa

Polyarteritis nodosa, so named because of the numerous nodules found along the course of muscular arteries, is a primary multisystem inflammatory disease of smaller and medium-sized

TABLE 30.2 CLASSIFICATION OF THE VASCULITIDES

GROUP	EXAMPLES	CHARACTERISTICS
Small vessel vasculitis	Microscopic polyangiitis	Necrotizing vasculitis with few or no immune deposits affecting medium and small blood vessels, including capillaries, venules, and arterioles; necrotizing glomerulonephritis and involvement of the pulmonary capillaries are common
	Wegener granulomatosis	Granulomatous inflammation involving the respiratory tract and necrotizing vasculitis affecting capillaries, venules, arterioles, and arteries; necrotizing glomerulonephritis is common
Medium-sized vessel vasculitis	Polyarteritis nodosa	Necrotizing inflammation of medium-sized or small arteries without vasculitis in arteries, capillaries, or venules; usually associated with underlying disease or environmental agents
	Kawasaki disease	Involves large, medium-sized, and small arteries (frequently the coronaries) and is associated with mucocutaneous lymph node syndrome; usually occurs in small children
	Thromboangiitis obliterans	Segmental, thrombosing, acute and chronic inflammation of the medium-sized and small arteries, principally the tibial and radial arteries but sometimes extending to the veins and nerves of the extremities; occurs almost exclusively in men who are heavy smokers
Large vessel vasculitis	Giant cell (temporal) arteritis	Granulomatous inflammation of the aorta and its major branches with predilection for extracranial vessels of the carotid artery; infiltration of vessel wall with giant cells and mononuclear cells; usually occurs in people older than 50 years of age and is often associated with polymyalgia rheumatica
	Takayasu arteritis	Granulomatous inflammation of the aorta and its branches; usually occurs in people younger than 50 years of age

blood vessels, especially those of the kidney, liver, intestine, peripheral nerve, skin, and muscle. The disease is seen more commonly in men than women.

Etiology

The cause of polyarteritis nodosa remains unknown. It can occur in drug abusers and may be associated with the use of certain drugs such as allopurinol and the sulfonamides. There is an association between polyarteritis nodosa and hepatitis B, with 10% to 30% of people with the disease having antibodies to hepatitis B. Other associations include serous otitis media, hairy cell leukemia, and hyposensitization therapy for allergies. People with connective tissue diseases such as systemic lupus erythematosus, rheumatoid arthritis, and primary Sjögren syndrome may have manifestations similar to those of primary polyarteritis nodosa.[23]

Clinical Manifestations

The onset of polyarteritis nodosa course can be acute, subacute, or chronic with long periods of symptom-free intervals.[3] Clinical signs and symptoms may vary due to the widely varied vascular involvement. It usually begins complaints of anorexia, weight loss, fever, and fatigue often accompanied by signs of organ involvement. The kidney is the most frequently affected organ, and hypertension is a common manifestation of the disorder. Gastrointestinal involvement may manifest as abdominal pain, nausea, vomiting, or diarrhea. Myalgia, arthralgia, and arthritis are common, as are peripheral neuropathies such as paresthesias, pain, and weakness. Central nervous system complications include thrombotic and hemorrhagic stroke. Cardiac manifestations result from involvement of the coronary arteries. Skin lesions also may occur and are highly variable. They include reddish blue, mottled areas of discoloration of the skin of the extremities called *livedo reticularis,* purpura (*i.e.,* black and blue discoloration from bleeding into the skin), urticaria (*i.e.,* hives), and ulcers.

Diagnosis and Treatment

Laboratory findings, although variable, include an elevated erythrocyte sedimentation rate, leukocytosis, anemia, and signs of organ involvement such as hematuria and abnormal liver function test results. The diagnosis is confirmed through biopsy specimens demonstrating necrotizing vasculitis of the small and large arteries. Treatment involves use of high-dose corticosteroid therapy and often-cytotoxic immunosuppressant agents (*e.g.,* azathioprine, cyclophosphamide). Typically, 3 months of a cytotoxic immunosuppressant is given, followed by tapered glucocorticoids over the next 4 months.[22] Before the availability of corticosteroids and immunosuppressive agents, the disease commonly was fatal. For people with polyarteritis nodosa associated with hepatitis B, aggressive simultaneous treatment of the hepatitis with antiviral agents is indicated.

Giant Cell Temporal Arteritis

Temporal arteritis (*i.e.,* giant cell arteritis), the most common of the vasculitides, is a focal inflammatory condition of medium-sized and large arteries. It predominantly affects

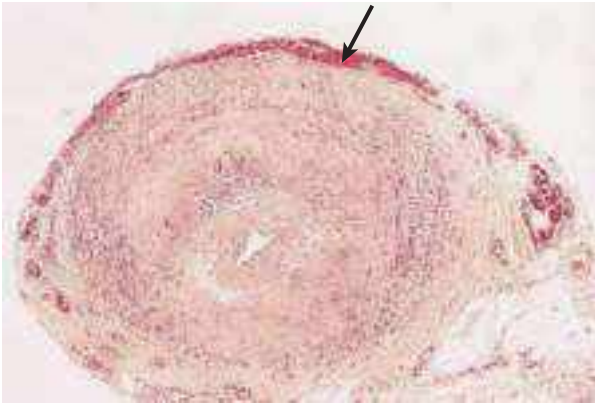


FIGURE 30.9 • Temporal arteritis. A cross-sectional photograph of a temporal artery shows inflammation throughout the wall, giant cells (*arrow*), and a lumen severely narrowed by intimal thickening. (From Rubin R., Strayer D. (Eds.). (2012). *Rubin's pathology: Clinicopathologic foundations of medicine* (6th ed., p. 466). Philadelphia, PA: Lippincott Williams & Wilkins.)

branches of arteries originating from the aortic arch, including the superficial temporal, vertebral, ophthalmic, and posterior ciliary arteries. The disorder progresses to involve the entire artery wall with focal necrosis and granulomatous inflammation involving multinucleated giant cells (Fig. 30.9). It is more common in older adults, with a 2:1 female-to-male ratio. The cause is unknown. However, an autoimmune origin, such as an initial T cell–mediated immune response, has been suggested.[3]

The disorder often is insidious in onset and may be heralded by the sudden onset of headache, tenderness over the artery, swelling and redness of the overlying skin, blurred vision or diplopia, and facial pain. Almost one half of affected persons have systemic involvement in the form of polymyalgia rheumatica. Up to 10% of people with giant cell arteritis go on to develop aortic aneurysm (especially thoracic).

Diagnosis is based on the clinical manifestations, a characteristically elevated erythrocyte sedimentation rate and CRP, and temporal artery biopsy. Treatment includes use of high-dose corticosteroids without delay because of the significant risk of visual symptoms. Before people with the disorder were treated with corticosteroids, blindness developed in almost 80% of cases due to involvement of the posterior ciliary artery. A typical starting dose of prednisolone is 40 to 60 mg/day for 4 weeks. [22,23]

Arterial Disease of the Extremities

Disorders of the circulation in the extremities often are referred to as *peripheral vascular disorders.* In many respects, the disorders that affect arteries in the extremities are the same as those affecting the coronary and cerebral arteries in that they produce ischemia, pain, impaired function, and in some cases infarction and tissue necrosis. Not only are the effects similar, but the pathologic conditions that impair circulation in the extremities are identical. This section focuses on acute arterial occlusion of the extremities, atherosclerotic occlusive disease, thromboangiitis obliterans, and Raynaud disease and phenomenon.

Acute Arterial Occlusion

Acute arterial occlusion is a sudden event that interrupts arterial flow to the affected tissues or organ. Most acute arterial occlusions are the result of an embolus or a thrombus. Although much less common than emboli and thrombus, trauma or arterial spasm caused by arterial cannulation can be another cause of acute arterial occlusion.

Etiology and Pathogenesis

An embolus is a freely moving particle such as a blood clot that breaks loose and travels in the larger vessels of the circulation until lodging in a smaller vessel and occluding blood flow. Most emboli arise in the heart and are caused by conditions that cause blood clots to develop on the wall of a heart chamber or valve surface. Emboli usually are a complication of heart disease: ischemic heart disease with or without infarction, atrial fibrillation, or rheumatic heart disease. Prosthetic heart valves can be another source of emboli. Other types of emboli are fat emboli that originate from bone marrow of fractured bones, air emboli from the lung, and amniotic fluid emboli that develop during childbirth. Acute arterial embolism is associated with a 5% to 25% risk of affected limb loss and a 25% to 30% increase in hospital mortality. Heart disease is responsible for over half these deaths.

A thrombus is a blood clot that forms on the wall of a vessel and continues to grow until reaching a size that obstructs blood flow. Thrombi often arise as the result of erosion or rupture of the fibrous cap of an arteriosclerotic plaque.

Clinical Manifestations

The signs and symptoms of acute arterial occlusion depend on the artery involved and the adequacy of the collateral circulation. Emboli tend to lodge in bifurcations of the major arteries, including the aorta and iliac, femoral, and popliteal arteries.[5] The presentation of acute arterial embolism is often described as that of the seven "Ps":

- Pistol shot (acute onset)
- Pallor
- Polar (cold)
- Pulselessness
- Pain
- Paresthesia
- Paralysis

Occlusion in an extremity causes sudden onset of acute pain with numbness, tingling, weakness, pallor, and coldness. There often is a sharp line of demarcation between the oxygenated tissue above the line of obstruction and the ischemic tissue below the line of obstruction. Pulses are absent below the level of the occlusion. These changes are followed rapidly by cyanosis, mottling, and loss of sensory, reflex, and motor function. Tissue death occurs unless blood flow is restored.

Diagnosis and Treatment

Diagnosis of acute arterial occlusion is based on signs of impaired blood flow. It uses visual assessment, palpation of pulses, and methods to assess blood flow. Treatment of acute arterial occlusion is aimed at restoring blood flow. An embolectomy, surgical removal of the embolus, is the optimal therapy when a large artery is occluded.

Thrombolytic therapy (*i.e.,* streptokinase or tissue plasminogen activator) may be used in an attempt to dissolve the clot. Anticoagulant therapy (*i.e.,* heparin) usually is given to prevent extension of the embolus and to prevent progression of the original thrombus. Application of cold should be avoided, and the extremity should be protected from injury resulting from hard surfaces and overlying bedclothes.

Atherosclerotic Occlusive Disease

Atherosclerosis is an important cause of peripheral artery disease (PAD) and is seen most commonly in the vessels of the lower extremities. The condition is sometimes referred to as *arteriosclerosis obliterans.* The superficial femoral and popliteal arteries are the most commonly affected vessels. When lesions develop in the lower leg and foot, the tibial, common peroneal, or pedal vessels are the arteries most commonly affected. The disease is seen most commonly in men and women as they advance in age. Approximately 20% of people in their 70s have PAD.[24,25]

Etiology

The risk factors for this disorder are similar to those for atherosclerosis. Cigarette smoking contributes to the progress of the atherosclerosis of the lower extremities and to the development of symptoms of ischemia. People with diabetes mellitus develop more extensive and rapidly progressive vascular disease than do people who do not have diabetes.

Clinical Manifestations

As with atherosclerosis in other locations, the signs and symptoms of vessel occlusion are gradual. Usually, there is at least a 50% narrowing of the vessel before symptoms of ischemia arise. The primary symptom of chronic obstructive arterial disease is *intermittent claudication* or pain with walking.[25] Typically, people with the disorder complain of calf pain because the gastrocnemius muscle has the highest oxygen consumption of any muscle group in the leg during walking. Some people may complain of a vague aching feeling or numbness, rather than pain. Other activities such as swimming, bicycling, and climbing stairs use other muscle groups and may not incite the same degree of discomfort as walking.

Other signs of ischemia include atrophic changes and thinning of the skin and subcutaneous tissues of the lower leg and diminution in the size of the leg muscles. The foot often is cool, and the popliteal and pedal pulses are weak or absent. Limb color blanches with elevation of the leg because of the effects of gravity on perfusion pressure and becomes deep red when the leg is in the dependent position because of an autoregulatory increase in blood flow and a gravitational increase in perfusion pressure.

When blood flow is reduced to the extent that it no longer meets the minimal needs of resting muscle and nerves, ischemic pain at rest, ulceration, and gangrene develop. As tissue

necrosis develops, there typically is severe pain in the region of skin breakdown, which is worse at night with limb elevation and is improved with standing.[25]

Diagnosis

Diagnostic methods include inspection of the limbs for signs of chronic low-grade ischemia such as subcutaneous atrophy, brittle toenails, hair loss, pallor, coolness, or dependent rubor. Palpation of the femoral, popliteal, posterior tibial, and dorsalis pedis pulses allows for an estimation of the level and degree of obstruction. The ratio of ankle to arm (*i.e.,* tibial and brachial arteries) systolic blood pressure is used to detect significant obstruction, with a ratio of less than 0.9 indicating occlusion. Normally, systolic pressure in the ankle exceeds that in the brachial artery because systolic pressure and pulse pressure tend to increase as the pressure wave moves away from the heart. Blood pressures may be taken at various levels on the leg to determine the level of obstruction. A Doppler ultrasound stethoscope may be used for detecting pulses and measuring blood pressure. Ultrasound imaging, magnetic resonance imaging (MRI) arteriography, spiral computed tomographic (CT) arteriography, and invasive contrast angiography may also be used as diagnostic methods.[24,25]

Treatment

The two goals of treatment in people with PAD are (1) to decrease their considerable cardiovascular risk and (2) to reduce symptoms. People with PAD should be evaluated for coexisting coronary and cerebrovascular atherosclerosis. The risk of death, mainly from coronary and cerebrovascular events, is higher than if they did not have PAD.[25] It is also important to address other cardiovascular risk factors, including smoking, hypertension, high lipid levels, and diabetes. Smoking cessation should be encouraged, and the coexisting health conditions should be treated appropriately.

Antiplatelet agents (aspirin or clopidogrel) reduce the vascular death rate in people with PAD by about 25%.[25] Other medications that are useful include statins, cilostazol (a phosphodiesterase inhibitor), and pentoxifylline (an adenosine diphosphate [ADP] receptor antagonist that decreases blood viscosity and improves erythrocyte flexibility). The tissues of extremities affected by atherosclerosis are easily injured and slow to heal. Treatment includes measures directed at protection of the affected tissues and preservation of functional capacity. Walking (slowly) to the point of claudication usually is encouraged because it increases collateral circulation.

Percutaneous or surgical intervention is typically reserved for the person with disabling claudication or limb-threatening ischemia. Surgery (*i.e.,* femoropopliteal bypass grafting using a section of saphenous vein) may be indicated in severe cases. In people with diabetes, the peroneal arteries between the knees and ankles commonly are involved, making revascularization difficult. Thromboendarterectomy with removal of the occluding core of atherosclerotic tissue may be done if the section of diseased vessel is short. Percutaneous transluminal angioplasty and stent placement, in which a balloon catheter is inserted into the area of stenosis and the balloon inflated to increase vessel diameter, is another form of treatment.[24,25]

Thromboangiitis Obliterans

Thromboangiitis obliterans, or Buerger disease, is an inflammatory (*i.e.,* vasculitis) arterial disorder that causes thrombus formation. The disorder affects the medium-sized arteries, usually the plantar and digital vessels in the foot and lower leg. Arteries in the arm and hand also may be affected. It is characterized by segmental, thrombosing, acute and chronic inflammation. Although primarily an arterial disorder, the inflammatory process often extends to involve adjacent veins and nerves. Usually the disease is seen in people less than 35 years of age who are heavy cigarette smokers.

Etiology and Pathogenesis

The pathogenesis of Buerger disease remains speculative. However, cigarette smoking and in some instances tobacco chewing seem to be involved. It has been suggested that the nicotine has a direct effect on the endothelial cell toxicity and may trigger an immune response. [3] Genetic influences are suggested since it is more prevalent in certain ethnic groups.

Clinical Manifestations

Pain is the predominant symptom of the disorder. It usually is related to distal arterial ischemia. During the early stages of the disease, there is intermittent claudication in the arch of the foot and the digits. In severe cases, pain is present even when the person is at rest. The impaired circulation increases sensitivity to cold. The peripheral pulses are diminished or absent, and there are changes in the color of the extremity. In moderately advanced cases, the extremity becomes cyanotic when the person assumes a dependent position, and the digits may turn reddish blue even when in a nondependent position. With lack of blood flow, the skin assumes a thin, shiny look and hair growth and skin nutrition suffer. Chronic ischemia causes thick, malformed nails. If the disease continues to progress, tissues eventually ulcerate and gangrenous changes arise that may necessitate amputation.

Diagnosis and Treatment

Diagnostic methods are similar to those for atherosclerotic disease of the lower extremities. As part of the treatment program for thromboangiitis obliterans, it is mandatory that the person stop smoking cigarettes or using tobacco. Even passive smoking and nicotine replacement therapy should be eliminated. Other treatment measures are of secondary importance and focus on methods for producing vasodilation and preventing tissue injury. Sympathectomy may be done to alleviate the vasospastic manifestations of the disease.

Raynaud Disease and Phenomenon

Raynaud disease or phenomenon is a functional disorder caused by intense vasospasm of the arteries and arterioles in the fingers and, less often, the toes. This is a common disorder affecting 3% to 5% of the population and is more common in

women than men. The disorder is divided into two types—the primary type, called *Raynaud disease,* occurs without demonstrable cause, and the secondary type, called *Raynaud phenomenon,* is associated with other disease states or known causes of vasospasm.[3,5,26]

Etiology and Pathogenesis

Vasospasm implies an excessive vasoconstrictor response to stimuli that normally produce only moderate vasoconstriction. In contrast to other regional circulations that are supplied by vasodilator and vasoconstrictor fibers, the cutaneous vessels of the fingers and toes are innervated only by sympathetic vasoconstrictor fibers. In these vessels, vasodilation occurs by withdrawal of sympathetic stimulation. Cooling of specific body parts such as the head, neck, and trunk produces a sympathetic-mediated reduction in digital blood flow, as does emotional stress.

Raynaud disease is precipitated by exposure to cold or by strong emotions and usually is limited to the fingers. It also follows a more benign course than Raynaud phenomenon, seldom causing tissue necrosis. The cause of vasospasm in primary Raynaud disease is unknown. Raynaud phenomenon is associated with previous vessel injury, such as frostbite, occupational trauma associated with the use of heavy vibrating tools, collagen diseases, neurologic disorders, and chronic arterial occlusive disorders. Another occupation-related cause is the exposure to alternating hot and cold temperatures such as that experienced by butchers and food preparers.[26] Raynaud phenomenon often is the first symptom of collagen diseases, such as scleroderma and systemic lupus erythematosus.[5]

Clinical Manifestations

In Raynaud disease and Raynaud phenomenon, ischemia due to vasospasm causes changes in skin color that progress from pallor to cyanosis, a sensation of cold, and changes in sensory perception, such as numbness and tingling. The color changes usually are first noticed in the tips of the fingers, later moving into one or more of the distal phalanges (Fig. 30.10). After the ischemic episode, there is a period of hyperemia with intense redness, throbbing, and paresthesias. The period of hyperemia is followed by a return to normal color. Although all of the fingers usually are affected symmetrically, in some cases only one or two digits are involved, or only a portion of the digit is affected.

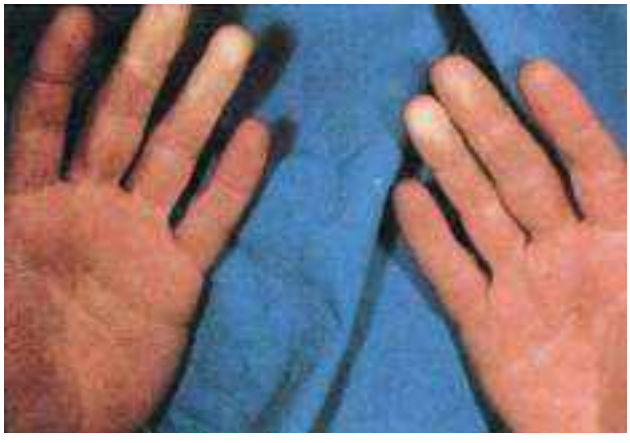


FIGURE 30.10 • Raynaud phenomenon. The tips of the fingers show marked pallor. (From Rubin R., Strayer D. (Eds.). (2012). *Rubin's pathology: Clinicopathologic foundations of medicine* (4th ed., p. 463). Philadelphia, PA: Lippincott Williams & Wilkins.)

In severe, progressive cases usually associated with Raynaud phenomenon, trophic changes may develop. The nails may become brittle, and the skin over the tips of the affected fingers may thicken. Nutritional impairment of these structures may give rise to arthritis. Ulceration and superficial gangrene of the fingers, although infrequent, may occur.

Diagnosis and Treatment

The initial diagnosis is based on history of vasospastic attacks supported by other evidence of the disorder. Immersion of the hand in cold water may be used to initiate an attack as an aid to diagnosis. Laser Doppler flow velocimetry may be used to quantify digital blood flow during changes in temperature. Raynaud disease is differentiated from Raynaud phenomenon by excluding secondary disorders known to cause vasospasm.[26]

Treatment measures are directed toward eliminating factors that cause vasospasm and protecting the digits from trauma during an ischemic episode. Abstinence from smoking and protection from cold are priorities. The entire body must be protected from cold, not just the extremities. Avoidance of emotional stress is another important factor in controlling the disorder because anxiety and stress may precipitate a vascular spasm in predisposed people. Vasoconstrictor medications, such as the decongestants contained in allergy and cold preparations, should be avoided. Treatment with vasodilator drugs may be indicated, particularly if episodes are frequent, because frequency encourages the potential for development of thrombosis and gangrene. The calcium channel–blocking drugs (*e.g.,* nifedipine, diltiazem) decrease the severity and frequency of attacks. Prazosin, an α-adrenergic receptor–blocking drug, also may be used. Surgical interruption of sympathetic nerve pathways (sympathectomy) may be used for people with severe symptoms.[26]

Aneurysms

An *aneurysm* is an abnormal localized dilation of a blood vessel. Aneurysms can occur in arteries and veins, but they are most common in the aorta. There are two types of aneurysms—true aneurysms and false aneurysms. A *true aneurysm* is one in which the aneurysm is bounded by a complete vessel wall. The blood in a true aneurysm remains within the vascular compartment. *False aneurysm* or *pseudoaneurysm* represents a localized *dissection* or tear in the inner wall of the artery with formation of an extravascular hematoma that causes vessel enlargement. Unlike true aneurysms, false aneurysms are bounded only by the outer layers of the vessel wall or supporting tissues (Fig. 30.11).

Aneurysms can assume several forms and may be classified according to their cause, location, and anatomic features.

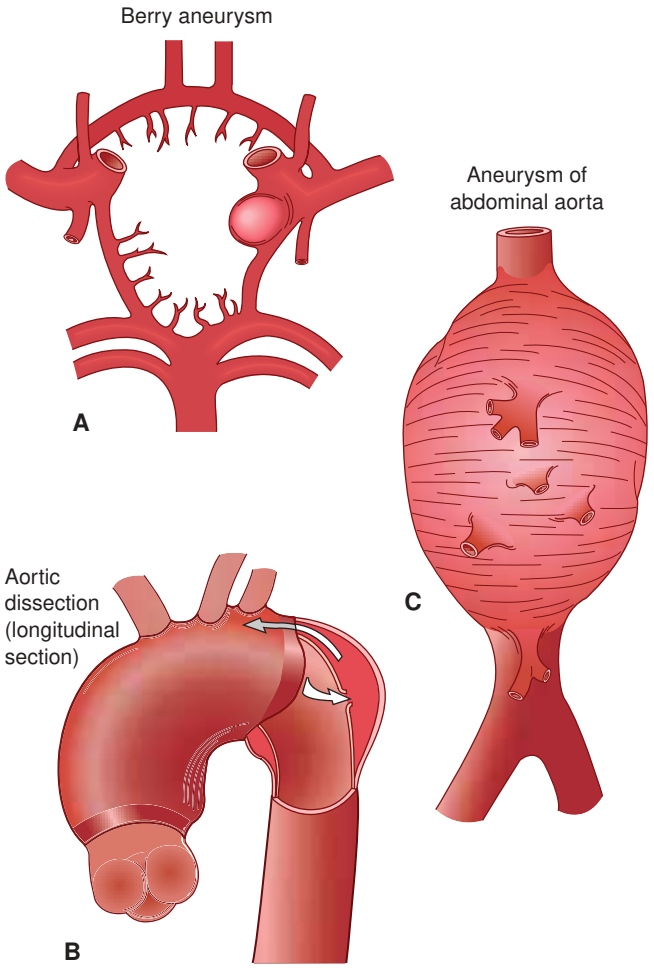


FIGURE 30.11 • Three forms of aneurysms: (**A**) berry aneurysm in the circle of Willis, (**B**) aortic dissection, and (**C**) fusiform-type aneurysm of the abdominal aorta.

A *berry aneurysm* is a true aneurysm that consists of a small, spherical dilation of the vessel at a bifurcation.[3,5] This type of aneurysm usually is found in the circle of Willis in the cerebral circulation. A *fusiform aneurysm* is a true aneurysm that involves the entire circumference of the vessel and is characterized by a gradual and progressive dilation of the vessel. These aneurysms, which vary in diameter (up to 20 cm) and length, may involve the entire ascending and transverse portions of the thoracic aorta or may extend over large segments of the abdominal aorta. A *saccular aneurysm* is a true aneurysm that extends over part of the circumference of the vessel and appears saclike.[5] A *dissecting aneurysm* is a false aneurysm resulting from a tear in the intimal layer of the vessel that allows blood to enter the vessel wall, dissecting its layers to create a blood-filled cavity (see Fig. 30.11). If dissection occurs in the aorta, it is a life-threatening condition (see below).

The weakness that leads to aneurysm formation may be caused by several factors, including congenital defects, trauma, infections, and atherosclerosis. Once initiated, the aneurysm grows larger as the tension in the vessel increases. This is because the tension in the wall of a vessel is equal to the pressure multiplied by the radius (*i.e.,* tension = pressure × radius; see Chapter 21). In this case, the pressure in the segment of the vessel affected by the aneurysm does not change but remains the same as that of adjacent portions of the vessel. As an aneurysm increases in diameter, the tension in the wall of the vessel increases in direct proportion to its increased size. If untreated, the aneurysm may rupture because of the increased tension. Even an unruptured aneurysm can cause damage by exerting pressure on adjacent structures and interrupting blood flow.

Aortic Aneurysms

Aortic aneurysms may involve any part of the aorta—the ascending aorta, aortic arch, descending aorta, thoracoabdominal aorta, or abdominal aorta. Multiple aneurysms may be present.

Etiology

The two most common causes of aortic aneurysms are atherosclerosis and degeneration of the vessel media. Half of the people with aortic aneurysms have hypertension.[5] Aortic aneurysms usually develop more frequently in men after the age of 50 years who smoke cigarettes.

Clinical Manifestations

The signs and symptoms of aortic aneurysms depend on the size and location. An aneurysm also may be asymptomatic, with the first evidence of its presence being associated with vessel rupture. Aneurysms of the thoracic aorta are less common than abdominal aortic aneurysms. They account for less than 10% of aortic aneurysms and may present with substernal, back, and neck pain. There also may be dyspnea, stridor, or brassy cough caused by pressure on the trachea. Hoarseness may result from pressure on the recurrent laryngeal nerve, and there may be difficulty swallowing because of pressure on the esophagus.[27] The aneurysm also may compress the superior vena cava, causing distention of neck veins and edema of the face and neck.

Abdominal aortic aneurysms are located most commonly below the level of the renal artery (>90%) and involve the bifurcation of the aorta and proximal end of the common iliac arteries.[3,5] The infrarenal aorta is normally 2 cm in diameter; an aneurysm is defined as an aortic diameter greater than 3 cm. They can involve any part of the vessel circumference (saccular) or extend to involve the entire circumference (fusiform). Most abdominal aneurysms are asymptomatic. Because an aneurysm is of arterial origin, a pulsating mass may provide the first evidence of the disorder. Typically, aneurysms larger than 4 cm are palpable. The mass may be discovered during a routine physical examination, or the affected person may complain of its presence. Calcification, which frequently exists on the wall of the aneurysm, may be detected during abdominal radiologic examination. Pain may be present and varies from mild mid-abdominal or lumbar

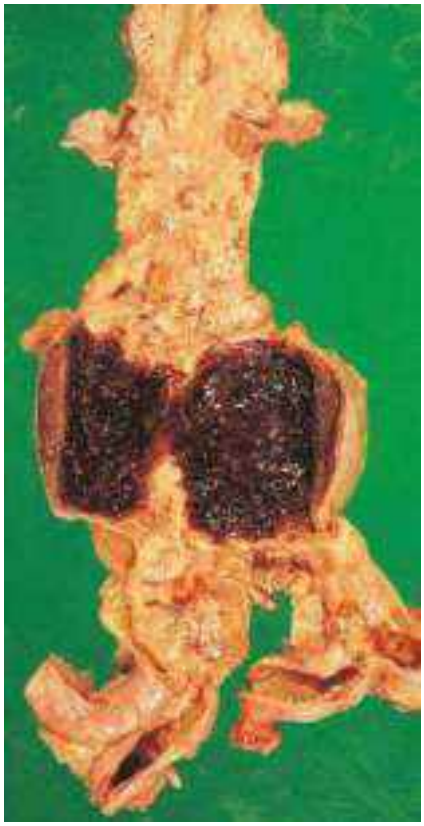


FIGURE 30.12 • Atherosclerotic aneurysm of the abdominal aorta. The aneurysm has been opened longitudinally to reveal a large thrombus in the lumen. The aorta and common iliac arteries display complicated lesions of atherosclerosis. (From Rubin R., Strayer D. (Eds.). (2012). *Rubin's pathology: Clinicopathologic foundations of medicine* (6th ed., p. 471). Philadelphia, PA: Lippincott Williams & Wilkins.)

discomfort to severe abdominal and back pain. As the aneurysm expands, it may compress the lumbar nerve roots, causing lower back pain that radiates to the posterior aspects of the legs. The aneurysm may extend to and impinge on the renal, iliac, or mesenteric arteries or to the vertebral arteries that supply the spinal cord. An abdominal aneurysm also may cause erosion of vertebrae. Stasis of blood favors thrombus formation along the wall of the vessel (Fig. 30.12), and peripheral emboli may develop, causing symptomatic arterial insufficiency.

With thoracic and abdominal aneurysms, the most dreaded complication is rupture. The likelihood of rupture correlates with increasing aneurysm size. The risk of rupture rises from less than 2% for small abdominal aneurysms (<4 cm in diameter) to 5% to 10% per year for aneurysms larger than 5 cm in diameter.[3]

Diagnosis and Treatment

Diagnostic methods include use of ultrasonography, echocardiography, CT scans, and MRI. Surgical repair, in which the involved section of the aorta is replaced with a synthetic graft of woven Dacron, frequently is the treatment of choice.[27]

Aortic Dissection

Aortic dissection (dissecting aneurysm) is an acute, life-threatening condition. It involves hemorrhage into the vessel wall with longitudinal tearing of the vessel wall to form a blood-filled channel (Fig. 30.13). Unlike atherosclerotic aneurysms, aortic dissection often occurs without evidence of previous vessel dilation. More than 95% of the cases

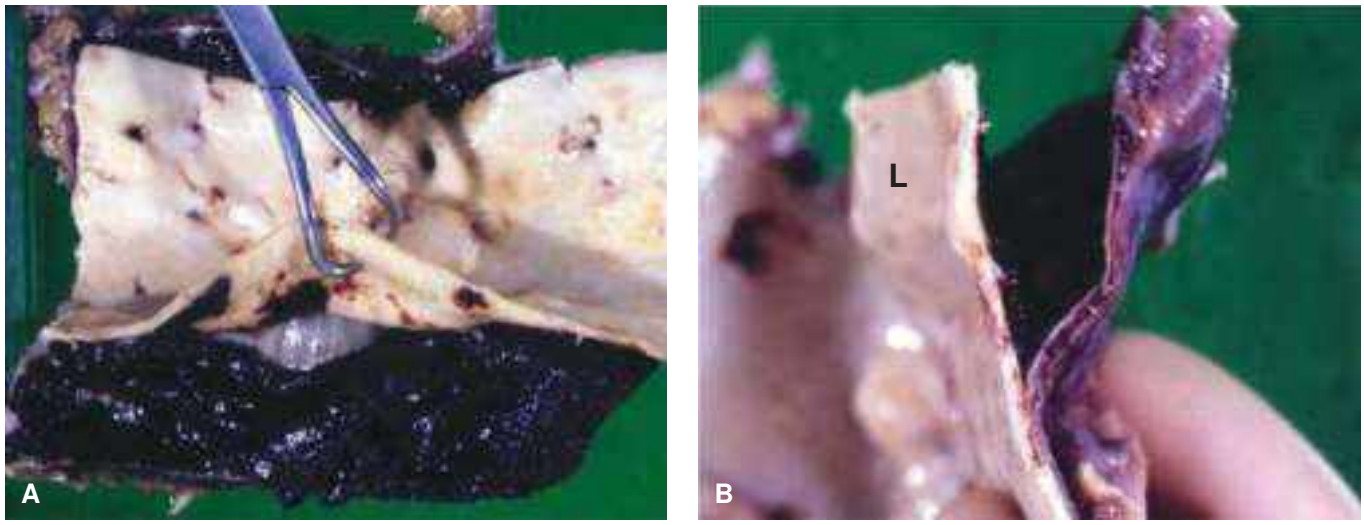


FIGURE 30.13 • Dissecting aneurysm of the aorta. (**A**) A transverse tear is present in the aortic arch. The orifices of the great vessels are on the left. (**B**) The thoracic aorta has been open longitudinally and reveals clotted blood dissecting the media of the vessel. The luminal surface shows extensive complicated lesions of atherosclerosis. (From Rubin R., Strayer D. (Eds.). (2012). *Rubin's pathology: Clinicopathologic foundations of medicine* (6th ed., p. 472). Philadelphia, PA: Lippincott Williams & Wilkins.)

of dissecting aneurysm show transverse tear in the intima and internal media. The dissection can originate anywhere along the length of the aorta. The majority of the dissections involve the ascending aorta.[5] The second most common site is the thoracic aorta just distal to the origin of the subclavian artery.

Etiology and Pathogenesis

Aortic dissection is caused by conditions that weaken or cause degenerative changes in the elastic and smooth muscle of the layers of the aorta. It is most common in the 40- to 60-year-old age group and more prevalent in men than in women.[3] Two risk factors predispose to aortic dissection—hypertension and degeneration of the medial layer of the vessel wall. There is a history of hypertension in most cases.[3] Aortic dissection also is associated with connective tissue diseases, such as Marfan syndrome. It also may occur during pregnancy because of changes in the aorta that occur during this time. Other factors that predispose to dissection are congenital defects of the aortic valve (*i.e.,* bicuspid or unicuspid valve structures) and aortic coarctation. Aortic dissection is a potential complication of cardiac surgery or catheterization. Surgically related dissection may occur at the points where the aorta has been incised or cross-clamped. It also has been reported at the site where the saphenous vein was sutured to the aorta during coronary artery bypass surgery.

Aortic dissections are commonly classified into two types, type A and type B, as determined by the level of dissection. The more common (and potentially more serious in terms of complications) proximal lesions, involving the ascending aorta only or both the ascending and the descending aorta, are designated type A. Those not involving the ascending aorta and usually beginning distal to the subclavian artery are designated type B.[3] Aortic dissections are also classified according to time of onset as acute or chronic.[28] Chronic dissections are defined as the persistence of the dissection flap or channel for greater than 2 weeks after initial event. Dissections usually extend distally from the intimal tear. When the ascending aorta is involved, expansion of the wall of the aorta may impair closure of the aortic valve. There also is the risk of aortic rupture with blood moving into the pericardium and compressing the heart. Although the length of dissection varies, it is possible for the abdominal aorta to be involved with progression into the renal, iliac, or femoral arteries. Partial or complete occlusion of the arteries that arise from the aortic arch or the intercostal or lumbar arteries may lead to stroke, ischemic peripheral neuropathy, or impaired blood flow to the spinal cord.

Clinical Manifestations

A major symptom of a dissecting aneurysm is the abrupt presence of excruciating pain, described as tearing or ripping. The location of the pain may point to the site of dissection.[3] Pain associated with dissection of the ascending aorta frequently is located in the anterior chest, and pain associated with dissection of the descending aorta often is located in the back. In the early stages, blood pressure typically is moderately or markedly elevated. Later, the blood pressure and the pulse rate become unobtainable in one or both arms as the dissection disrupts arterial flow to the arms. Syncope, hemiplegia, or paralysis of the lower extremities may occur because of occlusion of blood vessels that supply the brain or spinal cord. Heart failure may develop when the aortic valve is involved.

Diagnosis and Treatment

Diagnosis of aortic dissection is based on history and physical examination. Aortic angiography, transesophageal echocardiography, CT scans, and MRI studies aid in the diagnosis.[28]

The treatment of dissecting aortic aneurysm may be medical or surgical depending on the type and whether it is acute or chronic. Because aortic dissection is a life-threatening emergency, people with a probable diagnosis are stabilized medically even before the diagnosis is confirmed. Two important factors that participate in propagating the dissection are high blood pressure and the steepness of the pulse wave. Without intervention, these forces continue to cause extension of the dissection. Therefore, medical treatment focuses on control of hypertension and the use of drugs that lessen the force of systolic blood ejection from the heart. Two commonly used drugs, given in combination, are intravenous sodium nitroprusside and a β-adrenergic–blocking drug. Surgical treatment consists of resection of the involved segment of the aorta and replacement with a prosthetic graft. The mortality rate due to untreated dissecting aneurysm is high.[28]

IN SUMMARY

The arterial system distributes blood to all the tissues of the body, and lesions of the arterial system exert their effects through ischemia or impaired blood flow. There are two types of arterial disorders: diseases such as atherosclerosis, vasculitis, and peripheral arterial diseases that obstruct blood flow and disorders such as aneurysms that weaken the vessel wall.

Cholesterol relies on lipoproteins (LDLs and HDLs) for transport in the blood. The LDLs, which are atherogenic, carry cholesterol to the peripheral tissues. The HDLs, which are protective, remove cholesterol from the tissues and carry it back to the liver for disposal (reverse cholesterol transport). LDL receptors play a major role in removing cholesterol from the blood; persons with reduced numbers of receptors are at particularly high risk for development of atherosclerosis.

Atherosclerosis, a leading cause of death in the United States, affects large and medium-sized arteries, such as the coronary and cerebral arteries. It has an insidious onset, and its lesions usually are far advanced before symptoms appear. Although the mechanisms of atherosclerosis are uncertain, risk factors associated with its development have been identified. These include factors such as heredity, sex,

and age, which cannot be controlled, and factors such as smoking, high blood pressure, high serum cholesterol levels, diabetes, obesity, and inflammation, which can be controlled or modified.

The vasculitides are a group of vascular disorders characterized by vasculitis or inflammation and necrosis of the blood vessels in various tissues and organs of the body. They can be caused by injury to the vessel, infectious agents, or immune processes or can occur secondary to other disease states such as systemic lupus erythematosus.

Occlusive disorders interrupt arterial flow of blood and interfere with the delivery of oxygen and nutrients to the tissues. Occlusion of flow can result from a thrombus, emboli, vessel compression, vasospasm, or structural changes in the vessel. Peripheral arterial diseases affect blood vessels outside the heart and thorax. They include Raynaud disease or phenomenon, caused by vessel spasm, and thromboangiitis obliterans (Buerger disease), characterized by an inflammatory process that involves medium-sized arteries.

Aneurysms are localized areas of vessel dilation caused by weakness of the arterial wall. A berry aneurysm, most often found in the circle of Willis in the brain circulation, consists of a small, spherical vessel dilation. Fusiform and saccular aneurysms, most often found in the thoracic and abdominal aorta, are characterized by gradual and progressive enlargement of the aorta. They can involve part of the vessel circumference (saccular) or extend to involve the entire circumference of the vessel (fusiform). A dissecting aneurysm is an acute, life-threatening condition. It involves hemorrhage into the vessel wall with longitudinal tearing (dissection) of the vessel wall to form a blood-filled channel. The most serious consequence of aneurysms is rupture.

DISORDERS OF THE VENOUS CIRCULATION

After completing this section of the chapter, you should be able to meet the following objectives:

- Describe venous return of blood from the lower extremities, including the function of the muscle pumps and the effects of gravity, and relate to the development of varicose veins.
- Characterize the pathology of venous insufficiency and relate to the development of stasis dermatitis and venous ulcers.
- List the four most common causes of lower leg ulcer.

Veins are low-pressure, thin-walled vessels that rely on the ancillary action of skeletal muscle pumps and changes in abdominal and intrathoracic pressure to return blood to the heart. The venous system in the legs consists of two components—the superficial veins (*i.e.,* saphenous vein and its tributaries) and the deep venous channels. Perforating, or communicating, veins connect these two systems. Blood from the skin and subcutaneous tissues in the leg collects in the superficial veins and is then transported across the communicating veins into the deeper venous channels for return to the heart.

Unlike the arterial system, the venous system is equipped with valves that prevent retrograde flow of blood. These valves play an important role in the function of the venous system. Although these valves are irregularly located along the length of the veins, they almost always are found at junctions where the communicating veins merge with the larger deep veins and where two veins meet. The number of venous valves differs somewhat from one person to another, as does their structural competence. These factors may help explain the familial predisposition to development of varicose veins.

The action of the leg muscles assists in moving venous blood from the lower extremities back to the heart. When a person walks, the action of the leg muscles serves to increase flow in the deep venous channels and return venous blood to the heart (Fig. 30.14). The function of the so-called *muscle pump,* located in the gastrocnemius and soleus muscles of the lower extremities, can be compared with the pumping action of the heart.[5] During muscle contraction, which is similar to systole, valves in the communicating channels close to prevent backward flow of blood into the superficial system, as blood in the deep veins is moved forward by the action of the contracting muscles. During muscle relaxation, which is similar to diastole, the communicating valves open, allowing blood from the superficial veins to move into the deep veins.

Although its structure enables the venous system to serve as a storage area for blood, it also renders the system susceptible to problems related to stasis and venous insufficiency.

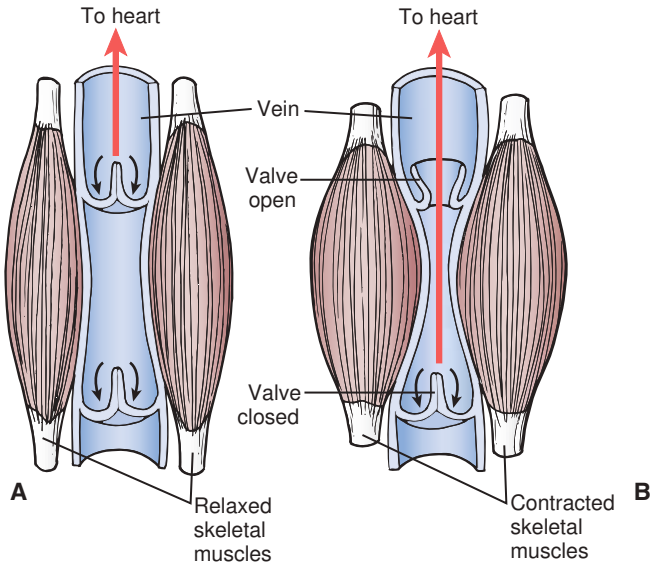


FIGURE 30.14 • The skeletal muscle pumps and their function in promoting blood flow in the deep and superficial calf vessels of the leg.

This section focuses on three common problems of the venous system—varicose veins, venous insufficiency, and venous thrombosis.

KEY POINTS

DISORDERS OF THE VENOUS CIRCULATION

- Veins are thin-walled, distensible vessels that collect blood from the tissues and return it to the heart. The venous system is a low-pressure system that relies on the pumping action of the skeletal muscles to move blood forward and the presence of venous valves to prevent retrograde flow.
- Disorders of the venous system produce congestion of the affected tissues and predispose to clot formation because of stagnation of flow and activation of the clotting system.

Varicose Veins

Varicose, or dilated, tortuous veins of the lower extremities are common and often lead to secondary problems of venous insufficiency. Varicose veins are described as being primary or secondary. Primary varicose veins originate in the superficial saphenous veins, and secondary varicose veins result from impaired flow in the deep venous channels. Approximately 80% to 90% of venous blood from the lower extremities is transported through the deep channels. The development of secondary varicose veins becomes inevitable when flow in these deep channels is impaired or blocked. The most common cause of secondary varicose veins is deep vein thrombosis (DVT). Other causes include congenital or acquired arteriovenous (AV) fistulas, congenital venous malformations, and pressure on the abdominal veins caused by pregnancy or a tumor.

The incidence of varicose veins rises with age. The prevalence of varicose veins is 50% in people older than 50 years. The condition is more common in female between 30 and 50 years of age, especially if there is a strong familial predisposition.[5] There is also a higher incidence in obese people due to the increase in intra-abdominal pressure and among people who stand for the majority of their day due to an occupation (*e.g.*, nurses).

Etiology and Pathogenesis

Prolonged standing and increased intra-abdominal pressure are important contributing factors in the development of primary varicose veins. Prolonged standing increases venous pressure and causes dilation and stretching of the vessel wall. One of the most important factors in the elevation of venous pressure is the hydrostatic effect associated with the standing position. When a person is in the erect position, the full weight of the venous columns of blood is transmitted to the leg veins. The effects of gravity are compounded in people who stand for long periods without using their leg muscles to assist in pumping blood back to the heart.

Because there are no valves in the inferior vena cava or common iliac veins, blood in the abdominal veins must be supported by the valves located in the external iliac or femoral veins. When intra-abdominal pressure increases, as it does during pregnancy, or when the valves in these two veins are absent or defective, the stress on the saphenofemoral junction is increased. The high incidence of varicose veins in women who have been pregnant also suggests a hormonal effect on venous smooth muscle contributing to venous dilation and valvular incompetence. Lifting also increases intra-abdominal pressure and decreases flow of blood through the abdominal veins. Occupations that require repeated heavy lifting also predispose to development of varicose veins.

Prolonged exposure to increased pressure causes the venous valves to become incompetent so they no longer close properly. When this happens, the reflux of blood causes further venous enlargement, pulling the valve leaflet apart and causing more valvular incompetence in sections of adjacent distal veins. Another consideration in the development of varicose veins is the fact that the superficial veins have only subcutaneous fat and superficial fascia for support, but the deep venous channels are supported by muscle, bone, and connective tissue. Obesity reduces the support provided by the superficial fascia and tissues, increasing the risk for development of varicose veins.

Clinical Manifestations

The signs and symptoms associated with primary varicose veins vary. Most women with superficial varicose veins complain of their unsightly appearance. In many cases, aching in the lower extremities and edema, especially after long periods of standing, may occur. The edema usually subsides at night when the legs are elevated. When the communicating veins are incompetent, symptoms are more common.

Diagnosis and Treatment

The diagnosis of varicose veins often can be made after a thorough history and physical examination, especially inspection of the extremities involved. Several procedures are used to assess the extent of venous involvement associated with varicose veins but have been shown to have limited value. One of the more helpful tests is the Perthes test. In this test, a tourniquet is applied to the affected knee while the person is instructed to complete 10 heel raises, and the leg is evaluated. If the varicosities empty, the site of reflux is above the tourniquet. If the veins remain distended, the site of reflux is below the tourniquet. The Doppler ultrasonic flow probe also may be used to assess flow in the large vessels. Angiographic studies using a radiopaque contrast medium also are used to assess venous function.[29]

After the venous channels have been repeatedly stretched and the valves rendered incompetent, little can be done to restore normal venous tone and function. Ideally, measures should be taken to prevent the development and progression of varicose veins. This would include weight loss and measures

center on avoiding activities such as continued standing that produce prolonged elevation of venous pressure.

Treatment measures for varicose veins focus on improving venous flow and preventing tissue injury. When correctly fitted, elastic support stockings or leggings compress the superficial veins and prevent distention. Prescription stockings measured to fit properly afford the most precise control. These stockings should be applied before the standing position is assumed, when the leg veins are empty.[29]

Sclerotherapy, which often is used in the treatment of small residual varicosities, involves the injection of a sclerosing agent into the collapsed superficial veins to produce fibrosis of the vessel lumen. Surgical treatment consists of removing the varicosities and the incompetent perforating veins, but it is limited to people with patent deep venous channels.

Chronic Venous Insufficiency

The term *venous insufficiency* refers to the physiologic consequences of DVT, valvular incompetence, or a combination of both conditions. The most common cause is DVT, which causes deformity of the valve leaflets, rendering them incapable of closure. In the presence of valvular incompetence, effective unidirectional flow of blood and emptying of the deep veins cannot occur. The muscle pumps also are ineffective, often driving blood in retrograde directions. Secondary failure of the communicating and superficial veins subjects the subcutaneous tissues to high pressures.

With venous insufficiency, there are signs and symptoms associated with impaired blood flow. In contrast to the ischemia caused by arterial insufficiency, venous insufficiency leads to tissue congestion, edema, and eventual impairment of tissue nutrition.[30] The edema is exacerbated by long periods of standing. Necrosis of subcutaneous fat deposits occurs, followed by skin atrophy. Brown pigmentation of the skin caused by hemosiderin deposits resulting from the breakdown of red blood cells is common. Secondary lymphatic insufficiency occurs, with progressive sclerosis of the lymph channels in the face of increased demand for clearance of interstitial fluid.

In advanced venous insufficiency, impaired tissue nutrition causes stasis dermatitis and the development of stasis or venous ulcers (Fig. 30.15). Stasis dermatitis is characterized by the presence of thin, shiny, bluish brown, irregularly pigmented desquamative skin that lacks the support of the underlying subcutaneous tissues. Minor injury leads to relatively painless ulcerations that are difficult to heal. The lower part of the leg is particularly prone to development of stasis dermatitis and venous ulcers. Most lesions are located medially over the ankle and lower leg, with the highest frequency just above the medial malleolus. Venous insufficiency is the most common cause of lower leg ulcers, accounting for nearly 80% of all cases.[31] The other common causes of lower extremity ulcers are arterial insufficiency, neuropathy (often due to diabetes), and pressure ulcers. People with long-standing venous insufficiency may also experience stiffening of the ankle joint and loss of muscle mass and strength.

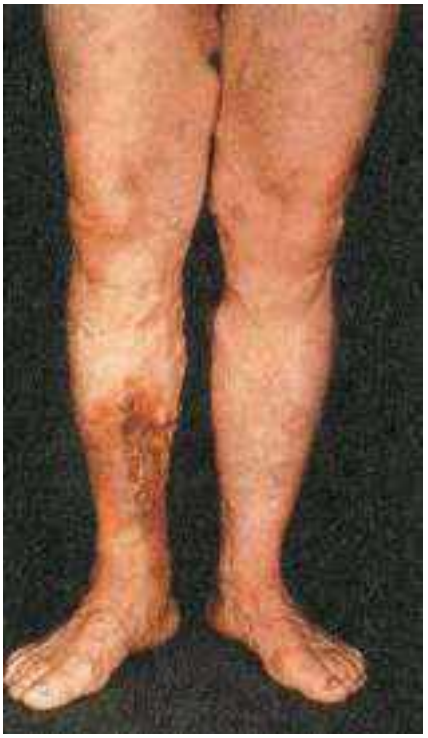


FIGURE 30.15 • Varicose veins of the legs. Severe varicosities of the superficial leg veins have led to stasis dermatitis and secondary ulcerations. (From Rubin R., Strayer D. (Eds.). (2012). *Rubin's pathology: Clinicopathologic foundations of medicine* (6th ed., p. 474). Philadelphia, PA: Lippincott Williams & Wilkins.)

Treatment of venous ulcers includes compression therapy with dressings and inelastic or elastic bandages. Medications that help include aspirin and pentoxifylline. Occasionally skin grafting is required for large or slow-healing venous ulcers. Growth factors (which are administered topically or by perilesional injection) may also be warranted.[31]

Venous Thrombosis

The term *venous thrombosis,* or *thrombophlebitis,* describes the presence of thrombus in a vein and the accompanying inflammatory response in the vessel wall. Thrombi can develop in the superficial or the deep veins. Superficial venous thrombosis can occur on any superficial vein and in the past, was thought to be a benign disease. Recently, SVT has been found to lead to complications such as reoccurrence of SVT, DVT, and pulmonary embolus in 10% of people.[32] DVT most commonly occurs in the lower extremities. DVT of the lower extremity is a serious disorder, complicated by pulmonary embolism, recurrent episodes of DVT, and development of chronic venous insufficiency. Most postoperative thrombi arise in the soleal sinuses or the large veins draining the gastrocnemius muscles.[5] Isolated calf thrombi often are asymptomatic. If left untreated, they may extend to the larger, more proximal veins, with an increased risk of pulmonary emboli of up to 90%.[5]

Disorders of Cardiac Function

32

Jaclyn Conelius

DISORDERS OF THE PERICARDIUM
- Acute Pericarditis
  - *Clinical Manifestations*
  - *Diagnosis*
  - *Treatment*
- Pericardial Effusion and Cardiac Tamponade
  - *Pathogenesis*
  - *Diagnosis*
  - *Treatment*
- Constrictive Pericarditis
  - *Etiology and Clinical Manifestations*
  - *Diagnosis*

CORONARY ARTERY DISEASE
- Coronary Circulation
  - *The Coronary Arteries*
  - *Myocardial Oxygen Supply and Demand*
  - *Assessment of Coronary Blood Flow and Myocardial Perfusion*
  - *Coronary Atherosclerosis and the Pathogenesis of Coronary Artery Disease*
- Acute Coronary Syndrome
  - *Electrocardiographic Changes*
  - *Serum Biomarkers*
  - *Unstable Angina/Non–ST-Segment Elevation Myocardial Infarction*
  - *ST-Segment Elevation Myocardial Infarction*
  - *Management of Acute Coronary Syndrome*
  - *Postinfarction Recovery Period*
  - *Cardiac Rehabilitation*
- Chronic Ischemic Heart Disease
  - *Stable Angina*
  - *Silent Myocardial Ischemia*
  - *Variant (Vasospastic) Angina*
  - *Chest Pain with Normal Coronary Angiography*
  - *Ischemic Cardiomyopathy*
  - *Diagnosis*
  - *Treatment*

CARDIOMYOPATHIES
- Primary Cardiomyopathies
  - *Genetic Cardiomyopathies*
  - *Mixed (Genetic and Nongenetic) Cardiomyopathies*
  - *Acquired Cardiomyopathies*
- Secondary Cardiomyopathies

INFECTIOUS AND IMMUNOLOGIC DISORDERS
- Infective Endocarditis
  - *Etiology and Pathogenesis*
  - *Clinical Manifestations*
  - *Diagnosis*
  - *Treatment*
- Rheumatic Heart Disease
  - *Pathogenesis*
  - *Clinical Manifestations*
  - *Diagnosis*
  - *Treatment and Prevention*

VALVULAR HEART DISEASE
- Hemodynamic Derangements
- Mitral Valve Disorders
  - *Mitral Valve Stenosis*
  - *Mitral Valve Regurgitation*
  - *Mitral Valve Prolapse*
- Aortic Valve Disorders
  - *Aortic Valve Stenosis*
  - *Aortic Valve Regurgitation*

HEART DISEASE IN INFANTS AND CHILDREN
- Embryonic Development of the Heart
- Fetal and Perinatal Circulation
- Congenital Heart Defects
  - *Pathophysiology*
  - *Manifestations and Treatment*
  - *Types of Defects*
  - *Adults with Congenital Heart Disease*
- Kawasaki Disease
  - *Pathogenesis*
  - *Clinical Manifestations*
  - *Diagnosis and Treatment*

CORONARY ARTERY DISEASE

After completing this section of the chapter, you should be able to meet the following objectives:

- Describe blood flow in the coronary circulation and relate it to the determinants of myocardial oxygen supply and demand.
- Define the term *acute coronary syndrome* and distinguish among chronic stable angina, unstable angina, non–ST-segment elevation myocardial infarction, and ST-segment elevation myocardial infarction in terms of pathology, symptomatology, ECG changes, and serum cardiac markers.
- Define the treatment goal for acute coronary syndrome.

The term *coronary artery disease* (CAD) describes heart disease caused by impaired coronary blood flow. In most cases, CAD is caused by atherosclerosis, which affects not only the coronary arteries but arteries in other areas of the body. Diseases of the coronary arteries can cause myocardial ischemia and angina, myocardial infarction or heart attack, cardiac arrhythmias, conduction defects, heart failure, and sudden death. Each year, more than 1.6 million Americans have new or recurrent myocardial infarctions; one third of those die within the first 24 hours, and many of those who survive suffer significant morbidity.[6]

Major risk factors for CAD include cigarette smoking, elevated blood pressure, elevated serum total and low-density lipoprotein (LDL) cholesterol, low serum high-density lipoprotein (HDL) cholesterol, diabetes, advancing age, abdominal obesity, and physical inactivity.[2] Individuals with diabetes and the metabolic syndrome are at particularly increased risk for development of CVD and have significant morbidity from the disease.[2]

Coronary Circulation

The Coronary Arteries

The two main coronary arteries, the left and the right, arise from the coronary sinus just above the aortic valve[7] (Fig. 32.3). The left coronary artery supplies blood flow to the anterior and left lateral portions of the LV. The *left main coronary artery* then divides into the left anterior descending and circumflex branches. The *left anterior descending artery* passes down through the groove between the two ventricles, giving off diagonal branches, which supply the LV, and perforating branches, which supply the anterior portion of the interventricular septum and the anterior papillary muscle of the LV. The *circumflex branch* of the left coronary artery passes to the left and moves posteriorly in the groove that separates the left atrium and ventricle, giving off branches that supply the left lateral wall of the LV. The *right coronary artery* lies in the right atrioventricular groove, and its branches supply most of the right ventricle and the posterior part of the LV in 80% to 90% of people. The right coronary artery usually moves to the back of the heart, where it forms the *posterior descending artery*, which normally supplies the posterior portion of the heart, interventricular septum, sinoatrial (SA) and atrioventricular (AV) nodes, and posterior papillary muscle. By convention, the coronary artery that supplies the posterior third of the septum (either the right coronary artery or the left circumflex) is called *dominant*. In a right dominant circulation, present in approximately four fifths of people, the left circumflex perfuses the lateral wall of the

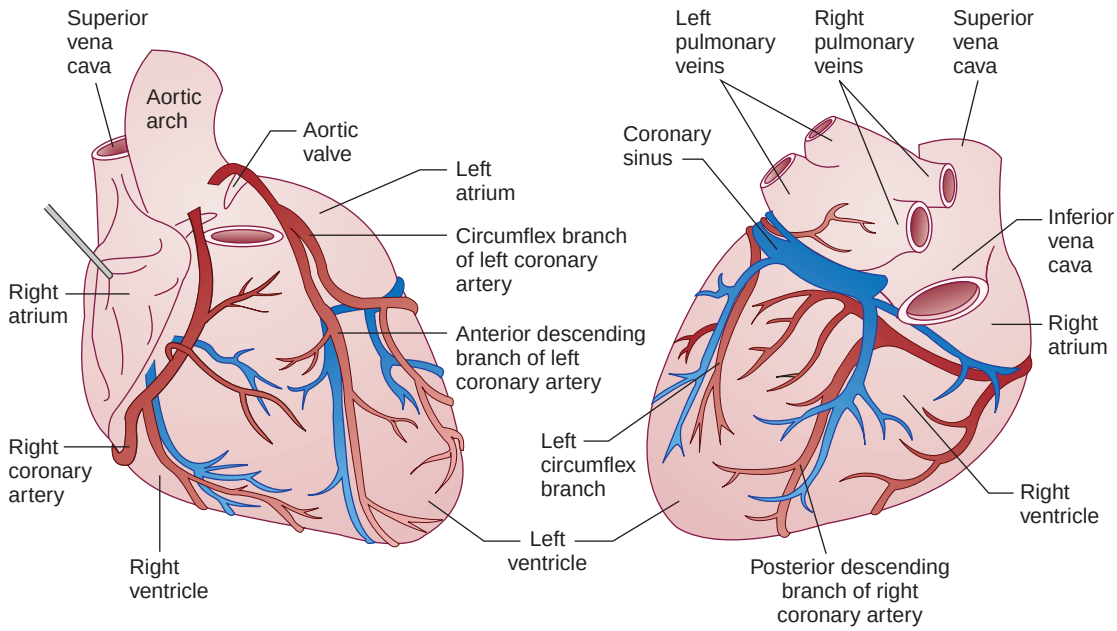


FIGURE 32.3 • Coronary arteries and some of the coronary sinus veins.

LV, and the right coronary artery supplies the entire right ventricular free wall and the posterior third of the septum.[7] Thus, occlusion of the right as well as the left coronary artery can cause left ventricular damage.

The large epicardial coronary arteries lie on the surface of the heart, with smaller intramyocardial arteries branching and penetrating the myocardium before merging with a network or plexus of subendocardial vessels. Although there are no connections between the large coronary arteries, there are anastomotic channels that join the small arteries. With gradual occlusion of the larger vessels, the smaller collateral vessels increase in size and provide alternative channels for blood flow. One of the reasons CAD does not produce symptoms until it is far advanced is that the collateral channels develop at the same time the atherosclerotic changes are occurring.

Blood flow in the coronary arteries is controlled largely by physical, neural, and metabolic factors. The openings for the coronary arteries originate in the root of the aorta just outside the aortic valve. Thus, the main factor responsible for perfusion of the coronary arteries is the aortic blood pressure, which is generated by the heart itself. Myocardial blood flow, in turn, is largely regulated by the metabolic activity of the myocardium and autoregulatory mechanisms that control vessel dilation. In addition to generating the aortic pressure that moves blood through the coronary vessels, the contracting heart muscle influences its own blood supply by compressing the intramyocardial and subendocardial blood vessels during systole. The autonomic nervous system exerts its effects on coronary blood flow through changes in heart rate, cardiac contractility, and blood pressure.

Coronary blood flow is largely regulated by the need of the cardiac muscle for oxygen. Even under normal resting conditions, the heart extracts and uses 70% of oxygen in blood flowing through the coronary arteries. Because there is little oxygen reserve in the blood, the coronary arteries must increase their flow to meet the metabolic needs of the myocardium during periods of increased activity. The normal resting blood flow through the coronary arteries averages approximately 225 mL/minute or about 4% to 5% of the total cardiac output.[7] During strenuous exercise, coronary flow may increase four- to fivefold to meet the energy requirements of the heart.

One of the major determinants of coronary blood flow is the metabolic activity of the heart. Numerous agents, referred to as *metabolites*, are thought to act as mediators for the vasodilation that accompanies increased cardiac work. These substances, which include potassium ions, lactic acid, carbon dioxide, and adenosine, are released from working myocardial cells. Of these substances, adenosine appears to have the greatest vasodilator effect and is perhaps the critical mediator of local blood flow.[7]

The endothelial cells that line blood vessels, including the coronaries, normally form a barrier between the blood and the arterial wall. They also synthesize several substances that, when released, can affect relaxation or constriction of the smooth muscle in the arterial wall. Potent vasodilators produced by the endothelium include nitric oxide, prostacyclin, and endothelium-derived hyperpolarizing factor (EDHF). The most important of these is nitric oxide. Most vasodilators and vasodilating stimuli exert their effects through nitric oxide. Products from aggregating platelets, thrombin, the products of mast cells, and increased shear force, which is responsible for so-called flow-mediated vasodilation, stimulate the synthesis and release of nitric oxide.[7] The endothelium also is the source of vasoconstricting factors, the best known of which are the endothelins.

Myocardial Oxygen Supply and Demand

The coronary circulation supplies the heart muscle with the oxygen and nutrients it needs to pump blood out to the rest of the body. In a person who is resting, 75% of the oxygen in the blood that passes through the myocardium is extracted. As the metabolic needs of the body change, cardiac function and coronary blood flow must adapt to meet these needs. If there is an imbalance in the myocardial oxygen supply and demand, myocardial ischemia and angina, myocardial infarction, or even sudden death can occur.

Myocardial Oxygen Supply. Myocardial oxygen supply is determined by the coronary arteries and capillary inflow and the ability of hemoglobin to transport and deliver oxygen to the heart muscle. Important factors in the transport and delivery of oxygen include the fraction of inspired oxygen in the blood and the number of red blood cells with normally functioning hemoglobin. Even with adequate coronary blood flow, myocardial ischemia can occur in situations of hypoxia, anemia, or carbon monoxide poisoning.[8]

Coronary Atherosclerosis and the Pathogenesis of Coronary Artery Disease

Atherosclerosis is the most common cause of CAD, is slow and progressive, and can begin at a very young age in the United States and other developed countries of the world.

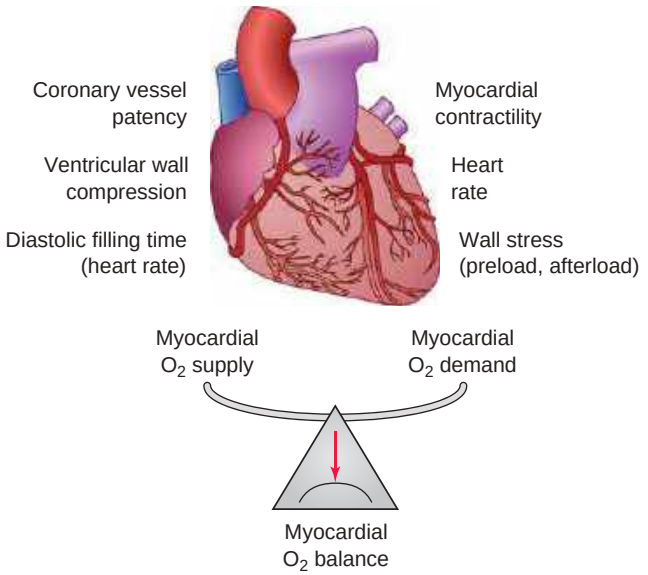


FIGURE 32.4 • Myocardial oxygen (O_2) balance is determined by factors that affect myocardial O_2 supply and myocardial O_2 demand (MVO_2).

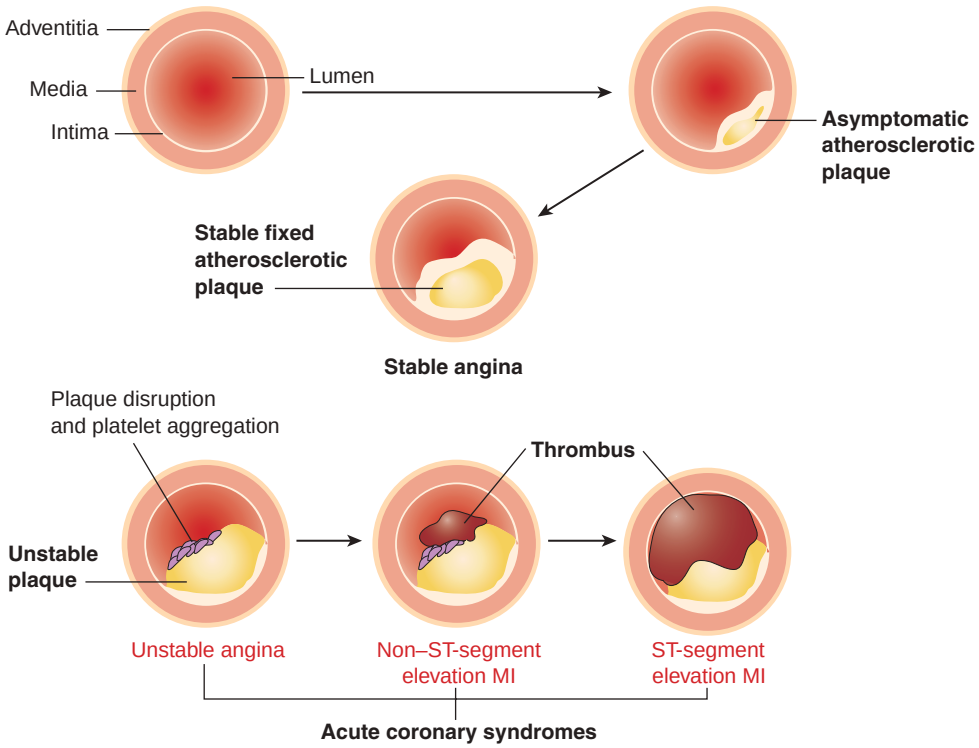


FIGURE 32.5 • Atherosclerotic plaque: stable fixed atherosclerotic plaque in stable angina and unstable plaque with plaque disruption and platelet aggregation in the acute coronary syndromes.

Atherosclerosis can affect one or all three of the major epicardial coronary arteries and their branches. Clinically significant lesions may be located anywhere in these vessels, but tend to predominate in the first several centimeters of the left anterior descending and left circumflex or the entire length of the right coronary artery.[15] Sometimes the major secondary branches also are involved.

Coronary artery disease is commonly divided into two types of disorders: the acute coronary syndrome and chronic ischemic heart disease. The a*cute coronary syndrome* (ACS) represents a spectrum of acute ischemic heart diseases ranging from unstable angina to myocardial infarction resulting from disruption of an atherosclerotic plaque. *Chronic ischemic heart disease* is characterized by recurrent and transient episodes of myocardial ischemia and stable angina that result from narrowing of a coronary artery lumen due to atherosclerosis and/or vasospasm.

Stable versus Unstable Plaque. There are two types of atherosclerotic lesions:

- Fixed or stable plaque, which obstructs blood flow
- Unstable/vulnerable plaque or high-risk plaque, which can rupture and cause platelet adhesion and thrombus formation

The fixed or stable plaque is commonly implicated in stable angina and the unstable plaque in unstable angina and myocardial infarction. In most cases, the myocardial ischemia underlying unstable angina, acute myocardial infarction, stroke, and, in many cases, sudden cardiac death (SCD) is precipitated by abrupt plaque changes, followed by thrombosis. The major determinants of plaque vulnerability to disruption include the size of the lipid-rich core, the stability and thickness of its fibrous cap, the presence of inflammation, and the lack of smooth muscle cells[16] (Fig. 32.5). Plaques with a thin fibrous cap overlaying a large lipid core are at high risk for rupture.

Although plaque disruption may occur spontaneously, it is often triggered by hemodynamic factors such as blood flow characteristics and vessel tension. For example, a sudden surge of sympathetic activity with an increase in blood pressure, heart rate, force of cardiac contraction, and coronary blood flow is thought to increase the risk of plaque disruption. Indeed, many people with myocardial infarction report a trigger event, most often emotional stress or physical activity. Plaque disruption also has a diurnal variation, occurring most frequently during the first hour after arising, suggesting that physiologic factors such as surges in coronary artery tone and blood pressure may promote atherosclerotic plaque disruption and subsequent platelet deposition.[16] It has been suggested that the sympathetic nervous system is activated on arising, resulting in changes in platelet aggregation and fibrinolytic activity that tend to favor thrombosis.

Thrombosis and Vessel Occlusion. Local thrombosis occurring after plaque disruption results from a complex interaction among its lipid core, smooth muscle cells, macrophages, and collagen. The lipid core provides a stimulus for platelet aggregation and thrombus formation.[17] Both smooth muscle and foam cells in the lipid core contribute to the expression of tissue factor in unstable plaques. Once exposed to blood, tissue factor initiates the extrinsic coagulation pathway, resulting in the local generation of thrombin and deposition of fibrin.[18]

Platelets play an important role in linking plaque disruption to acute CAD. As a part of the response to plaque disruption,

platelets adhere to the endothelium and release substances (*i.e.,* adenosine diphosphate [ADP], thromboxane A_2, and thrombin) that promote further aggregation of platelets and thrombus formation. The platelet membrane, which contains glycoprotein receptors that bind fibrinogen and link platelets together, contributes to thrombus formation. Platelet adhesion and aggregation occurs in several steps. First, release of ADP, thromboxane A_2, and thrombin initiates the aggregation process. Second, glycoprotein IIb/IIIa receptors on the platelet surface are activated. Third, fibrinogen binds to the activated glycoprotein receptors, forming bridges between adjacent platelets.

There are two types of thrombi formed as a result of plaque disruption—white platelet-containing thrombi and red fibrin-containing thrombi. The thrombi in unstable angina have been characterized as grayish white and presumably platelet rich.[19] Red thrombi, which develop with vessel occlusion in myocardial infarction, are rich in fibrin and red blood cells superimposed on the platelet component and completely obstruct blood flow.

Acute Coronary Syndrome

Acute coronary syndrome includes unstable angina, non–ST-segment elevation (non–Q-wave) myocardial infarction, and ST-segment elevation (Q-wave) myocardial infarction. Persons without ST-segment elevation on ECG are those in whom thrombotic coronary occlusion is subtotal or intermittent, whereas those with ST-segment elevation are usually found to have complete coronary occlusion on angiography, and many ultimately have Q-wave myocardial infarction.

KEY POINTS

CORONARY ARTERY DISEASE

- The term *coronary artery disease* refers to disorders of myocardial blood flow due to stable or unstable coronary atherosclerotic plaques.
- Unstable atherosclerotic plaques tend to fissure or rupture, causing platelet aggregation and potential for thrombus formation with production of a spectrum of acute coronary syndromes of increasing severity, ranging from unstable angina, to non–ST-segment elevation myocardial infarction, to ST-segment elevation myocardial infarction.
- Stable atherosclerotic plaques produce fixed obstruction of coronary blood flow with myocardial ischemia occurring during periods of increased metabolic need, such as in stable angina.

Electrocardiographic Changes

The classic ECG changes that occur with ACS involve T-wave inversion, ST-segment elevation, and development of an abnormal Q wave[20]. The changes that occur may not be present immediately after the onset of symptoms and vary considerably depending on the duration of the ischemic event (acute versus evolving), its extent (subendocardial versus transmural), and its location (anterior versus inferior posterior). Because these changes usually occur over time and are seen on the ECG leads that view the involved area of the myocardium, provision for continuous and serial 12-lead ECG monitoring is indicated. The repolarization phase of the action potential (T-wave and ST segment on the ECG) is usually the first to be involved during myocardial ischemia and injury. As the involved area becomes ischemic, myocardial repolarization is altered, causing changes in the T wave. This is usually represented by T-wave inversion, although a hyperacute T-wave elevation may occur as the earliest sign of infarction. ST-segment changes occur with ischemic myocardial injury and depending on what leads are involved can indicate the lesion of interest. Normally, the ST segment of the ECG is nearly isoelectric (*e.g.,* flat along the baseline) because all healthy myocardial cells attain the same potential during early repolarization. Acute severe ischemia reduces the resting membrane potential and shortens the duration of the action potential in the ischemic area. These changes create a voltage difference between the normal and ischemic areas of the myocardium that leads to a so-called current of injury between these regions. It is these currents of injury that are represented on the surface ECG as a deviation of the ST segment. When the acute injury is transmural, the overall ST vector is shifted in the direction of the outer epicardium, resulting in ST-segment elevation.[20] With Q-wave infarction, abnormal Q waves and R wave loss develop because there is no depolarizing current conduction from the necrotic tissue. When the injury is confined primarily to the subendocardium, the overall ST segment is shifted toward the inner ventricular layer, resulting in an overall depression and not elevation of the ST segment.

Serum Biomarkers

Even though cardiac biomarkers aid clinicians in diagnosing unstable angina/non–ST-segment elevation myocardial infarction (UA/NSTEMI) in approximately one third of people, awaiting results should delay reperfusion treatment for ST-segment elevation myocardial infarction (STEMI). The 12-lead ECG should initiate reperfusion treatment since this therapy is time sensitive. Serum biomarkers for ACS include cardiac-specific troponin I (TnI) and troponin T (TnT) and creatine kinase MB (CK-MB). As the myocardial cells become necrotic, their intracellular contents begin to diffuse into the surrounding interstitium and then into the blood. The rate at which the enzymes appear in the blood depends on their intracellular location, their molecular weight, and local blood flow. For example, they may appear at an earlier-than-predicted time in patients who have undergone successful reperfusion therapy.

The *troponin assays* have high specificity for myocardial tissue and have become the primary biomarker tests for the diagnosis of myocardial infarction. The troponin complex, which is part of the actin filament, consists of three subunits

(*i.e.,* troponin C [TnC], TnT, and TnI) that regulate calcium-mediated actin–myosin contractile process in striated muscle. TnI and TnT, which are present in cardiac muscle, begin to rise within 3 hours after the onset of myocardial infarction and may remain elevated for 7 to 10 days after the event. This is especially advantageous in the late diagnosis of myocardial infarction.[21]

Creatine kinase is an intracellular enzyme found in muscle cells. There are three isoenzymes of CK, with the MB isoenzyme being highly specific for injury to myocardial tissue. Serum levels of CK-MB exceed normal ranges within 4 to 8 hours of myocardial injury and decline to normal within 2 to 3 days.[21]

When comparing troponin and CK-MB, the troponin level identifies necrosis in cardiac muscles earlier then CK-MB. Clinicians examining cardiac biomarkers should focus on troponin levels, rather than CK-MB levels, for diagnosis and establishing the success of reperfusion.[21]

Unstable Angina/Non–ST-Segment Elevation Myocardial Infarction

UA/NSTEMI is considered to be a clinical syndrome of myocardial ischemia ranging from stable angina to myocardial infarction.[19] Typically, UA and NSTEMI differ in whether the ischemia is severe enough to cause sufficient myocardial damage to release detectable quantities of serum cardiac markers. Persons who have no evidence of serum markers for myocardial damage are considered to have UA, whereas a diagnosis of NSTEMI is indicated if a serum marker of myocardial injury is present.

The pathophysiology of UA/NSTEMI can be divided into five phases:

1. The development of the unstable plaque that ruptures or plaque erosion with superimposed nonocclusive thrombosis
2. An obstruction such as spasm, constriction, dysfunction, or adrenergic stimuli
3. Severe narrowing of the coronary lumen
4. Inflammation
5. Any physiological state causing ischemia related to decreased oxygen supply such as fever or hypotension[22]

Inflammation can play a prominent role in plaque instability, with inflammatory cells releasing cytokines that cause the fibrous cap to become thinner and more vulnerable to rupture or erosion. The pain associated with UA/NSTEMI has a persistent and severe course and is characterized by at least one of three features:

1. It occurs at rest (or with minimal exertion), usually lasting more than 20 minutes (if not interrupted by nitroglycerin).
2. It is severe and described as frank pain and of new onset (*i.e.,* within 1 month).
3. It is more severe, prolonged, or frequent than previously experienced.[22]

Risk stratification of people presenting with UA/NSTEMI is important because the outcome can range from excellent, with little change in treatment, to NSTEMI or death, requiring aggressive treatment. UA/NSTEMI is classified by severity based on clinical history, ECG pattern, and serum biomarkers. UA/NSTEMI is classified as:

- Class I (new onset severe angina)
- Class II (angina at rest within the past month, but not within the last 48 hours)
- Class III (angina at rest within 48 hours)

The ECG pattern in UA/NSTEMI demonstrates ST-segment depression (or transient ST-segment elevation) and T-wave changes. The degree of ST-segment deviation has been shown to be an important measure of ischemia and prognosis.

ST-Segment Elevation Myocardial Infarction

Acute STEMI, also known as *heart attack*, is characterized by the ischemic death of myocardial tissue associated with atherosclerotic disease of the coronary arteries. The area of infarction is determined by the coronary artery that is affected and by its distribution of blood flow (Fig. 32.6). Approximately 30% to 40% of infarcts affect the right coronary artery, 40% to 50% affect the left anterior descending artery, and the remaining 15% to 20% affect the left circumflex artery.[21]

Pathophysiology. The extent of the infarct depends on the location and extent of occlusion, amount of heart tissue supplied by the vessel, duration of the occlusion, metabolic needs of the affected tissue, extent of collateral circulation, and other factors such as heart rate, blood pressure, and cardiac rhythm. An infarct may involve the endocardium, myocardium, epicardium, or a combination of these. Transmural infarcts involve the full thickness of the ventricular wall and most commonly occur when there is obstruction of a single artery (Fig. 32.7). Subendocardial infarcts involve the inner one third to one half of the ventricular wall and occur more frequently in the presence of severely narrowed but still patent arteries. Most infarcts are transmural, involving the free wall of the LV and the interventricular septum.

The principal biochemical consequence of myocardial infarction is the conversion from aerobic to anaerobic metabolism with inadequate production of energy to sustain normal myocardial function. As a result, a striking loss of contractile function occurs within 60 seconds of onset.[21] Changes in cell structure (*i.e.,* glycogen depletion and mitochondrial swelling) develop within several minutes. These early changes are reversible if blood flow is restored. Although gross tissue changes are not apparent for hours after onset of myocardial infarction, the ischemic area ceases to function within a matter of minutes, and irreversible damage to cells occurs in approximately 40 minutes. Irreversible myocardial cell death (necrosis) occurs after 20 to 40 minutes of severe ischemia.[21] Microvascular injury occurs in approximately 1 hour and follows irreversible cell injury. If the infarct is large enough, it depresses overall left ventricular function and pump failure ensues.

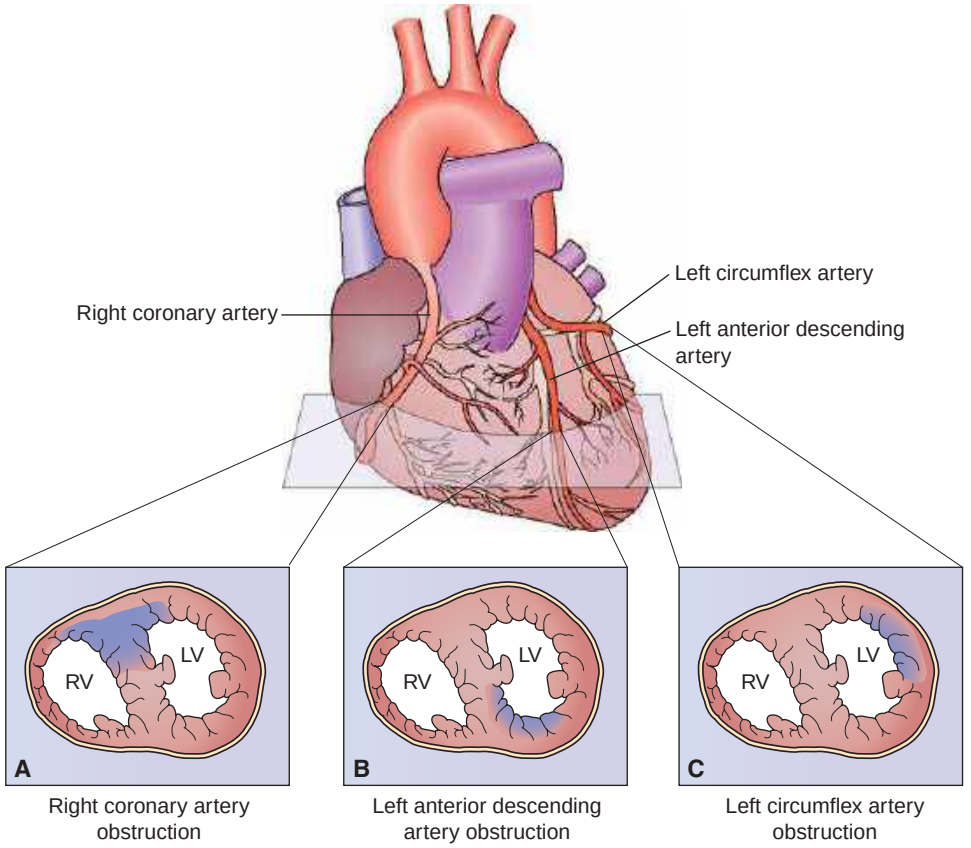


FIGURE 32.6 • Areas of the heart affected by occlusion of the (**A**) right coronary artery, (**B**) left anterior descending coronary artery, and (**C**) left circumflex coronary artery. (RV, right ventricle; LV, left ventricle.)

Multiple dynamic structural changes maintain cardiac function in persons with STEMI. Both the infarcted and noninfarcted areas of the ventricle undergo progressive changes in size, shape, and thickness, comprising early wall thinning, healing, hypertrophy, and dilation, collectively termed *ventricular remodeling*. As the nonfunctioning muscle in the infarcted area becomes thin and dilated, the muscle in the surrounding, noninfarcted area becomes thicker as it undergoes adaptive hypertrophy so it can take over the work of the muscle in the infarcted zone. However, the adaptive effect of remodeling may be overwhelmed by aneurysm formation or depression of myocardial function, causing further impairment of ventricular function.[21]

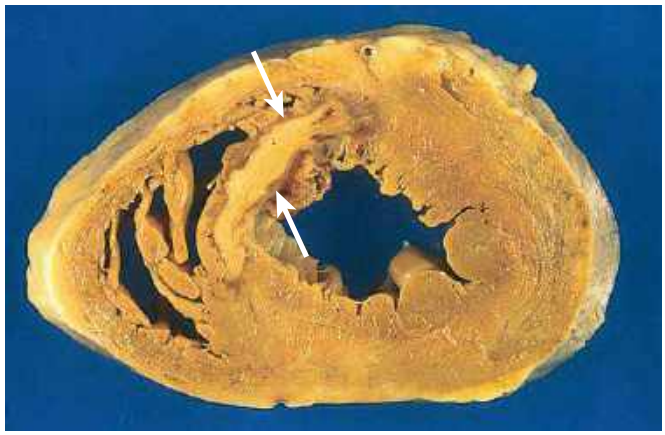


FIGURE 32.7 • Acute myocardial infarct. A cross section of the ventricles of a man who died a few days after the onset of severe chest pain shows a transmural infarct in the posterior and septal regions of the LV. The necrotic myocardium is soft, yellowish, and sharply demarcated. (From Rubin R., Strayer D. E. (Eds.) (2012). *Rubin's pathology: Clinicopathologic foundations of medicine* (6th ed., p. 502). Philadelphia, PA: Lippincott Williams & Wilkins.)

Clinical Manifestations. STEMI may occur as an abrupt-onset event or as a progression from UA/NSTEMI. The onset of STEMI usually is abrupt, with pain as the significant symptom. The pain typically is severe and crushing, often described as being constricting, suffocating, or like "something sitting on my chest." It usually is substernal, radiating to the left arm, neck, or jaw, although it may be experienced in other areas of the chest. Unlike that of angina, the pain associated with STEMI is more prolonged and not relieved by rest or nitroglycerin, and narcotics frequently are required. Some persons may not describe it as "pain," but as "discomfort." Women often experience atypical ischemic-type chest discomfort, whereas the elderly may complain of shortness of breath more frequently than chest pain.[6]

Gastrointestinal complaints are common with STEMI. There may be a sensation of epigastric distress; nausea and vomiting may occur. These symptoms are thought to be related to the severity of the pain and vagal stimulation. The epigastric distress may be mistaken for indigestion, and the person may seek relief with antacids or other home remedies, which only delays getting medical attention. Complaints of fatigue

and weakness, especially of the arms and legs, are common. Pain and sympathetic stimulation combine to give rise to tachycardia, anxiety, restlessness, and feelings of impending doom. A productive cough may be present with frothy, pink sputum. The skin often is pale, cool, and moist. Impairment of myocardial function may lead to hypotension and shock.

Sudden death from STEMI is death that occurs within 1 hour of symptom onset.[21] It usually is attributed to fatal arrhythmias, which may occur without evidence of infarction. Early hospitalization after onset of symptoms greatly improves the chances of averting sudden death because appropriate resuscitation facilities are immediately available when the ventricular arrhythmia occurs.

Management of Acute Coronary Syndrome

Because the specific diagnosis of STEMI often is difficult to make at the time of entry into the health care system, the immediate management of UA/NSTEMI and STEMI is generally the same. The prognosis in STEMI is largely related to the occurrences of two general complications—arrhythmias and mechanical complications (pump failure). The majority of deaths from STEMI are due to the sudden development of ventricular arrhythmias. Therefore, the major elements in management of people with STEMI include

- Recognition of symptoms and prompt seeking of medical care
- Prompt deployment of an emergency medical team capable of resuscitation procedures, including defibrillation
- Expeditious transport to a hospital equipped for managing arrhythmias and providing advanced cardiac life support
- Expeditious implementation of reperfusion therapy within 60 to 90 minutes[21]

People who experience signs and symptoms of STEMI often delay seeking treatment, despite current public information regarding the benefits of early treatment. People who delay seeking treatment in the hospital include older adults, women, African Americans, people of low socioeconomic status, people with a history of angina and/or diabetes, and people who consult a relative and/or physician.

Emergency department goals for management of ACS include identification of people who are candidates for reperfusion therapy. The history and physical examination should be conducted thoroughly, but efficiently, so as not to delay reperfusion therapy. Prior episodes of CVD, including ACS, coronary bypass surgery, or PCI, should be ascertained. Evaluation of the person's chief complaint, typically chest pain, along with other associated symptoms is essential in differentiating ACS from other diagnoses.

For any person presenting to the emergency department with symptoms of ACS, a 12-lead ECG should be obtained and read by a physician within 10 minutes of arrival at the emergency department. The typical ECG changes may not be present immediately after the onset of symptoms, except as arrhythmias. Diagnostic ECG tracings (*i.e.,* ST-segment elevation, prolongation of the Q wave, and inversion of the T wave) may be difficult to identify in people with STEMI who present with chest pain. Therefore serial ECG tracings should be obtained. Some added difficulties include premature ventricular contractions, which are common arrhythmias after myocardial infarction. The occurrence of other arrhythmias and conduction defects depends on the areas of the heart and conduction pathways that are included in the infarct. A new bundle branch block, particularly left bundle branch block, also serves as a criterion for STEMI and indicates a need for rapid reperfusion.

Commonly indicated treatment regimens include administration of oxygen, aspirin, nitrates, pain medications, antiplatelet and anticoagulant therapy, and β-adrenergic blocking agents (beta-blockers).[21] People with ECG evidence of infarction should receive immediate reperfusion therapy with a thrombolytic agent or PCI within 60 to 90 minutes.[21] The importance of intensive insulin control to maintain normal blood glucose (80 to 110 mg/dL) in people who are critically ill has been supported by multiple studies. Current ACC/AHA guidelines recommend the maintenance of strict glucose control during STEMI.[6]

Pain relief is a major objective in the treatment of STEMI. Control of pain in STEMI is accomplished through a combination of oxygen, nitrates, analgesics (*e.g.,* morphine), and β-adrenergic blocking agents. The administration of oxygen augments the oxygen content of inspired air and increases the oxygen saturation of hemoglobin. Arterial oxygen levels may fall precipitously after STEMI, and oxygen administration helps to maintain the oxygen content of the blood perfusing the coronary circulation. In people with severe heart failure from STEMI, continuous positive-pressure ventilation or endotracheal intubation and support with mechanical ventilation may be necessary.

Nitroglycerin is given because of its vasodilating effect and ability to relieve coronary pain. The vasodilating effects of the drug decrease venous return (*i.e.,* reduce preload) and arterial blood pressure (*i.e.,* reduce afterload), thereby reducing oxygen consumption. Nitroglycerin may also limit infarction size and is most effective if given within 4 hours of symptom onset. Nitroglycerin usually is administered sublingually initially, after which the need for intravenous infusion is assessed. The use of intravenous nitroglycerin may be indicated for treatment of ongoing ischemic pain, control of hypertension, or management of pulmonary congestion. Nitroglycerin should not be administered to patients with severe hypotension or to patients who have received a phosphodiesterase inhibitor for erectile dysfunction within the previous 24 hours.

Although a number of analgesic agents have been used to treat the pain of STEMI, morphine is usually the drug of choice.[21] It usually is indicated if chest pain is unrelieved with oxygen and nitrates. The reduction in anxiety that accompanies the administration of morphine contributes to a decrease in restlessness and autonomic nervous system activity, with a subsequent decrease in the metabolic demands of the heart. It is commonly given intravenously because of the rapid onset of action and because the intravenous route does not elevate enzyme levels. The intravenous route also bypasses the

variable rate of absorption of subcutaneous or intramuscular sites, which often are underperfused because of the decrease in cardiac output that occurs after infarction.

β-Adrenergic blocking drugs act as antagonists that block β-receptor–mediated functions of the sympathetic nervous system and thus decrease myocardial oxygen demand by reducing heart rate and cardiac contractility, and systemic arterial blood pressure. The lengthening of diastole caused by the slower heart rate may enhance myocardial perfusion, especially to the subendocardium. Beta-blockers also alter resting myocardial membrane potentials and may decrease life-threatening ventricular arrhythmias. Because sympathetic nervous system activity increases the metabolic demands of the myocardium, oral or intravenous beta-blockers are usually administered within the first few hours after the onset of STEMI. They should not be given in STEMI caused by cocaine use because it could accentuate coronary spasm. Other relative contraindications to beta-blockers include symptomatic bradycardia, hypotension, moderate-to-severe left ventricular failure, shock, or second- or third-degree heart block.

Platelets play a major role in the thrombotic response to atherosclerotic plaque disruption; therefore, inhibition of platelet aggregation is an important aspect in the early treatment of both UA/NSTEMI and STEMI. Aspirin (*i.e.,* acetylsalicylic acid) is the preferred antiplatelet agent for preventing platelet aggregation in persons with ACS. Aspirin, which acts by inhibiting synthesis of the prostaglandin thromboxane A_2, is thought to promote reperfusion and reduce the likelihood of rethrombosis. The actions of aspirin are related to the presence of the acetyl group, which irreversibly acetylates the critical platelet enzyme, cyclooxygenase, which is required for thromboxane A_2 synthesis. Because the action is irreversible, the effect of aspirin on platelet function lasts for the lifetime of the platelet—approximately 8 to 10 days. For patients who are unable to take aspirin because of hypersensitivity or gastrointestinal intolerance, clopidogrel may be prescribed. Clopidogrel is a thienopyridine derivative that reduces platelet aggregation by inhibiting the ADP pathway in platelets. Unlike aspirin, it has no effect on prostaglandin synthesis. Results of several studies have resulted in recommendations by the AHA for the use of clopidogrel along with aspirin for persons with UA/NSTEMI and for preprocedural loading and long-term therapy for persons undergoing PCI. Antithrombin agents are also used in the treatment of people with ACS. Anticoagulation therapy, which targets the coagulation pathway and formation of the fibrin clot, involves the use of unfractionated and low–molecular-weight heparin. The rationale for the use of antithrombin therapy in patients with STEMI is the prevention of deep vein thrombosis, pulmonary emboli, and cerebral embolization.

Angiotensin-converting-enzyme (ACE) inhibitors are often used during the early and convalescent phases of STEMI, demonstrating a benefit in terms of decreased mortality rate. ACE inhibitors increase cardiac output and stroke volume and reduce systemic pulmonary vascular resistance as well as pulmonary capillary wedge pressure. This, in turn, reduces LV dysfunction and decreased SCD. The greatest benefit is in those people with previous infarctions, heart failure, and tachycardia. ACE inhibitors usually are started within the first 24 hours, after fibrinolytic therapy has been completed. Therapy with ACE inhibitors is usually begun with low-dose oral administration and increased steadily to full dose.[6] Although the use of ACE inhibitors in short-term therapy for patients with UA/NSTEMI does not appear to have benefits, long-term use is helpful in preventing recurrent ischemic episodes.[19]

Reperfusion Strategies. The term *reperfusion* refers to reestablishment of blood flow through use of pharmacologic agents (fibrinolytic therapy), PCI, or coronary artery bypass grafting (CABG). All people presenting with STEMI should be assessed for reperfusion therapy as soon as possible on entry into the health care system. Time since onset of symptoms, risk of STEMI, possible risks associated with fibrinolytic therapy, and time required for transport to a skilled PCI laboratory should all be considered.

Early reperfusion (within 15 to 20 minutes) after onset of occlusion can prevent necrosis and improve myocardial perfusion in the infarct zone. Reperfusion after a longer interval can salvage some of the myocardial cells that would have died owing to longer periods of ischemia. It also may prevent the microvascular injury that occurs over a longer period. Even though much of the viable myocardium existing at the time of reflow, or reperfusion, ultimately recovers, critical abnormalities in biochemical function may persist, causing impaired ventricular function. The recovering area of the heart is often referred to as a hibernating myocardium. Because myocardial function is lost before cell death occurs, a hibernating myocardium may not be capable of sustaining life, and persons with large areas of dysfunctional myocardium may require life support until the stunned regions regain their function.[21]

Fibrinolytic Therapy. Fibrinolytic drugs dissolve blood and platelet clots and are used to reduce mortality, limit infarct size, encourage infarct healing and myocardial remodeling, and reduce the potential for life-threatening arrhythmias. These agents interact with plasminogen to generate plasmin, which lyses fibrin clots and digests clotting factors V and VIII, prothrombin, and fibrinogen. The fibrinolytic agents include streptokinase, alteplase, reteplase, and tenecteplase-tPA. The best results occur if treatment is initiated within 30 minutes of symptom onset. The magnitude of the benefit declines after this period, but it is possible that some benefit can be achieved for up to 12 hours after the onset of pain. The person must be a low-risk candidate for complications caused by bleeding, with no intracranial hemorrhage or significant trauma within the last 3 months. The primary complication of fibrinolytic therapy is intracranial hemorrhage, which usually occurs within the first 24 hours of treatment.[22]

Percutaneous Coronary Intervention. PCI is indicated as an early invasive procedure for patients with UA/NSTEMI who have no serious comorbidity and who have lesions amenable to PCI.[23] PCI includes percutaneous transluminal coronary angioplasty (PTCA), stent implantation, atherectomy, and thrombectomy. The goal in PCI is to perform the procedure

within 90 minutes of the patient's first medical contact (the "door-to-balloon interval").[23]

Balloon PTCA involves dilation of a stenotic atherosclerotic plaque with an inflatable balloon (Fig. 32.8). The procedure is similar to cardiac catheterization for coronary angiography, in that the double-lumen balloon dilation catheter is introduced percutaneously into the femoral or brachial artery and advanced under fluoroscopic view into the stenotic area of the affected coronary vessel. There it is used to expand the coronary arterial lumen by stretching and tearing the atherosclerotic plaque and, to a lesser extent, by distributing the plaque along its longitudinal axis. This procedure is often used in conjunction with the placement of stents, but now rarely used alone. Acute complications of PTCA include thrombosis and vessel dissection; longer-term complications involve restenosis of the dilated vessel.

The use of *coronary stenting* has been shown to improve short- and long-term outcomes compared with PTCA alone. People undergoing stent procedures are treated with antiplatelet and anticoagulant drugs to prevent thrombosis, which is a major risk of the procedure.[24] The bare metal self-expanding wire mesh stents that were used initially had high thrombosis rates and have largely been replaced by balloon-expandable stents. Radiation brachytherapy was used to improve restenosis of the stents and involved the use of localized intracoronary artery radiation for reduction of in-stent restenosis. The procedure is credited with inhibiting cell proliferation and vascular lesion formation and preventing constrictive arterial remodeling. This procedure has limitations including the need for radiation therapy and has been shown to be inferior to drug-eluting stents.[24] Bare metal stents are being used in 20% to 30% of the people who require PCI.

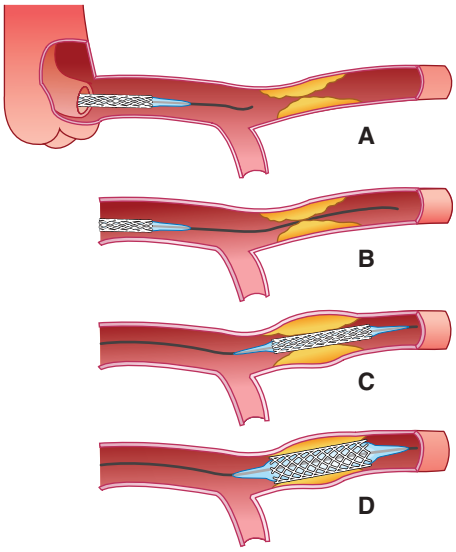


FIGURE 32.8 • Balloon-expandable stent insertion. (**A**) Insertion of a guide catheter with a collapsed balloon-expandable stent mounted over a guide wire into a coronary artery. (**B**) Advancement of guide wire across the coronary lesion. (**C**) Positioning of the balloon-expandable stent across the lesion. (**D**) Balloon inflation with expansion of the stent. Once the stent is expanded, the balloon system is removed.

Drug-eluting stents (with sirolimus, paclitaxel, zotarolimus, and everolimus) are also being used to suppress local neointimal proliferation that causes restenosis of the coronary artery.[24] Recent clinical trials found that long term use up to 1 year of aspirin and clopidogrel is recommended to prevent restenosis. *Atherectomy* (*i.e.,* cutting of the atherosclerotic plaque with a high-speed circular blade from within the vessel) is a mechanical technique to remove atherosclerotic tissue during angioplasty. Laser angioplasty devices are also used. However, with the availability of stents, these procedures are used less frequently than in the past. Thrombectomy (removal of the thrombus) involves the use of a special catheter device to fracture the thrombus into small pieces and then pull the fracture fragments into the catheter tip so they can be propelled proximally and removed.

Coronary Artery Bypass Grafting. CABG is one of the most common surgeries performed in the world, providing relief of angina, improvement in exercise tolerance, and prolongation of life. The procedure involves revascularization of the affected myocardium by placing a saphenous vein graft between the aorta and the affected coronary artery distal to the site of occlusion or by using the internal mammary artery as a means of revascularizing the left anterior descending artery or its branches (Fig. 32.9). One to five distal anastomoses commonly are done.

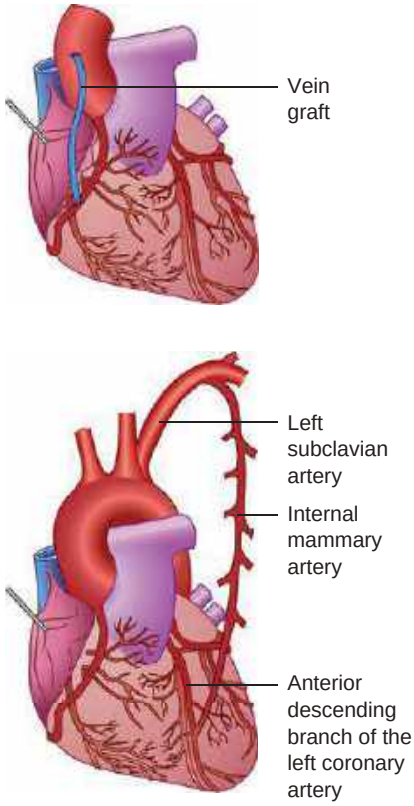


FIGURE 32.9 • Coronary artery revascularization. (**Top**) Saphenous vein bypass graft. The vein segment is sutured to the ascending aorta and the right coronary artery at a point distal to the occluding lesion. (**Bottom**) Mammary artery bypass graft. The mammary artery is anastomosed to the anterior descending left coronary artery, bypassing the obstructing lesion.

Chronic Ischemic Heart Disease

Myocardial ischemia occurs when the ability of the coronary arteries to supply blood is inadequate to meet the metabolic demands of the heart. Limitations in coronary blood flow most commonly are the result of atherosclerosis, but vasospasm may serve as an initiating or contributing factor.[29] There are several major types of chronic ischemic coronary artery disease: chronic stable angina, silent myocardial ischemia, variant or vasospastic angina, chest pain with normal coronary angiography, and ischemic cardiomyopathy.

Stable Angina

Chronic stable angina is associated with a fixed coronary obstruction that produces a disparity between coronary blood flow and metabolic demands of the myocardium. Stable angina is the initial manifestation of ischemic heart disease in approximately half of persons with CAD.[30] Although most people with stable angina have atherosclerotic heart disease, angina does not develop in a considerable number of people with advanced coronary atherosclerosis. This probably is because of their sedentary lifestyle, the development of adequate collateral circulation, or the inability of these people to perceive pain.

Angina pectoris usually is precipitated by situations that increase the work demands of the heart, such as physical exertion, exposure to cold, and emotional stress. The pain typically is described as a constricting, squeezing, or suffocating sensation. It usually is steady, increasing in intensity only at the onset and end of the attack. The pain of angina commonly is located in the precordial or substernal area of the chest; it is similar to myocardial infarction in that it may radiate to the left shoulder, jaw, arm, or other areas of the chest (Fig. 32.13). In some people, the arm or shoulder pain may be confused with arthritis; in others, epigastric pain is confused with indigestion.

Typically, chronic stable angina is provoked by exertion or emotional stress and relieved within minutes by rest or the use of nitroglycerin. A delay of more than 5 to 10 minutes before relief is obtained suggests that the symptoms are not due to ischemia or that they are due to severe ischemia. Angina that occurs at rest, is of new onset, or is increasing in intensity or duration denotes an increased risk for myocardial infarction and should be evaluated immediately using the criteria for ACS.

Silent Myocardial Ischemia

Silent myocardial ischemia occurs in the absence of anginal pain. The factors that cause silent myocardial ischemia appear to be the same as those responsible for angina: impaired blood flow from the effects of coronary atherosclerosis or vasospasm. The reason for the painless episodes of ischemia is unclear. The episodes may be shorter and involve less myocardial tissue than those pro- ducing pain. Another explanation is that persons with silent angina have defects in pain threshold or pain transmission, or autonomic neuropathy with sensory denervation. There is evidence of an increased incidence of silent myocardial isch- emia in persons with diabetes mellitus, probably the result of autonomic neuropathy, which is a common complication of diabetes.[29] Silent STEMI comprises a significant proportion of all STEMIs in older adults.

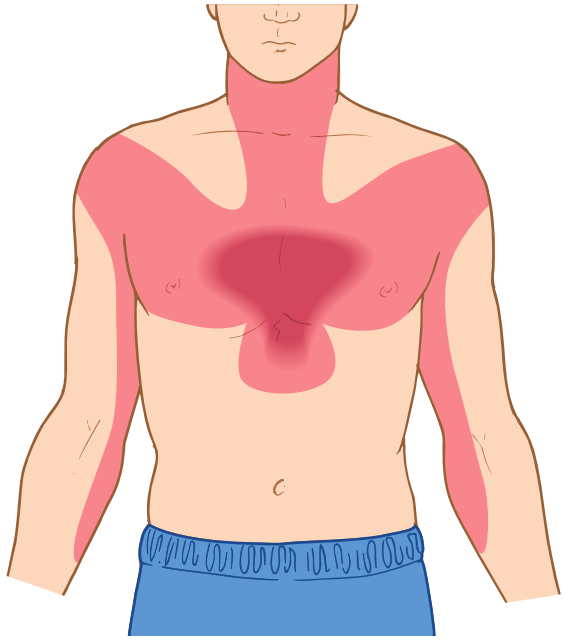


FIGURE 32.13 • Areas of pain due to angina.

Variant (Vasospastic) Angina

Variant angina is also known as *vasospastic* or *Prinzmetal angina*. The causes of variant angina are not completely understood, but a combination of pathologic processes may be responsible. It has been suggested that it may result from endothelial dysfunction, hyperactive sympathetic nervous system responses, defects in the handling of calcium by vascular smooth muscle, or from an alteration in nitric oxide production.[29] In some people it is associated with hypercontractility of vascular smooth muscle and is associated with migraine headaches or Raynaud phenomenon.

Unlike stable angina that occurs with exertion or stress, variant angina usually occurs during rest or with minimal exercise and frequently occurs nocturnally (between midnight and 8 AM). Arrhythmias often occur when the pain is severe, and most people are aware of their presence during an attack. People with variant angina who have serious arrhythmias during spontaneous episodes of pain are at a higher risk of sudden death.

Heart Failure and Circulatory Shock

34

Jaclyn Conelius

HEART FAILURE
- Pathophysiology of Heart Failure
  - *Control of Cardiac Performance and Output*
  - *Systolic versus Diastolic Dysfunction*
  - *Right versus Left Ventricular Dysfunction*
  - *High-Output versus Low-Output Failure*
  - *Compensatory Mechanisms*
- Acute Heart Failure Syndromes
- Clinical Manifestations of Heart Failure
  - *Respiratory Manifestations*
  - *Fatigue, Weakness, and Mental Confusion*
  - *Fluid Retention and Edema*
  - *Cachexia and Malnutrition*
  - *Cyanosis*
  - *Arrhythmias and Sudden Cardiac Death*
- Diagnosis and Treatment
  - *Diagnosis*
  - *Treatment*

CIRCULATORY FAILURE (SHOCK)
- Pathophysiology of Circulatory Shock
- Cardiogenic Shock
  - *Pathophysiology*
  - *Clinical Manifestations*
  - *Treatment*
- Hypovolemic Shock
  - *Pathophysiology*
  - *Clinical Manifestations*
  - *Treatment*
- Distributive Shock
  - *Neurogenic Shock*
  - *Anaphylactic Shock*
  - *Sepsis and Septic Shock*
- Obstructive Shock
- Complications of Shock
  - *Acute Lung Injury/Acute Respiratory Distress Syndrome*
  - *Acute Renal Failure*
  - *Gastrointestinal Complications*
  - *Disseminated Intravascular Coagulation*
  - *Multiple Organ Dysfunction Syndrome*

HEART FAILURE IN CHILDREN AND OLDER ADULTS
- Heart Failure in Infants and Children
  - *Clinical Manifestations*
  - *Diagnosis and Treatment*
- Heart Failure in Older Adults
  - *Clinical Manifestations*
  - *Diagnosis and Treatment*

Adequate perfusion of body tissues depends on the pumping ability of the heart, a vascular system that transports blood to the cells and back to the heart, sufficient blood to fill the circulatory system, and tissues that are able to extract and use oxygen and nutrients from the blood. Heart failure and circulatory shock are separate conditions that reflect failure of the circulatory system. Both conditions exhibit common compensatory mechanisms even though they differ in terms of pathogenesis and causes.

HEART FAILURE

After completing this section of the chapter, you should be able to meet the following objectives:

- Explain how the Frank-Starling mechanism, sympathetic nervous system, renin–angiotensin–aldosterone mechanism, natriuretic peptides, endothelins, and myocardial hypertrophy and remodeling function as adaptive and maladaptive mechanisms in heart failure.
- Differentiate high-output versus low-output heart failure, systolic versus diastolic heart failure, and right-sided versus left-sided heart failure in terms of causes, impact on cardiac function, and major manifestations.
- Differentiate chronic heart failure from acute heart failure syndromes and methods of diagnosis, assessment, and management.

Heart failure has been defined as a complex syndrome resulting from any functional or structural disorder of the heart that results in or increases the risk of developing manifestations of low cardiac output and/or pulmonary or systemic congestion.[1,2] In the United States, heart failure affected an estimated approximately 5 million people in 2007. Heart failure can occur in any age group but primarily affects older adults. Although morbidity and mortality rates from other cardiovascular diseases have decreased over the past several decades,

the incidence of heart failure is increasing at an alarming rate. Approximately 400,000 to 700,000 people are diagnosed with heart failure each year.

The syndrome of heart failure can be produced by any heart condition that reduces the pumping ability of the heart. Among the most common causes of heart failure are coronary artery disease, hypertension, dilated cardiomyopathy, and valvular heart disease.[1] Because many of the processes leading to heart failure are long-standing and progress gradually, heart failure can often be prevented or its progression slowed by early detection and intervention. The importance of these approaches is emphasized by the American College of Cardiology (ACC)/American Heart Association (AHA) guidelines that have incorporated a classification system of heart failure that includes four stages:

1. Stage A—High risk for developing heart failure, but no identified structural abnormalities and no signs of heart failure
2. Stage B—Presence of structural heart disease, but no history of sign and symptoms of heart failure
3. Stage C—Current or prior symptoms of heart failure with structural heart disease
4. Stage D—Advanced structural heart disease and symptoms of heart failure at rest on maximum medical therapy[1]

This staging system recognizes that there are established risk factors and structural abnormalities that are characteristic of the four stages of heart failure. People normally progress from one stage to another unless disease progression is slowed or stopped by treatment.

KEY POINTS

HEART FAILURE

- The function of the heart is to move deoxygenated blood from the venous system through the right heart into the pulmonary circulation, and oxygenated blood from the pulmonary circulation through the left heart and into the arterial circulation.
- Systolic dysfunction represents a decrease in cardiac myocardial contractility and an impaired ability to eject blood from the left ventricle, whereas diastolic dysfunction represents an abnormality in ventricular relaxation and filling.

Pathophysiology of Heart Failure

Cardiac output is the amount of blood that the ventricles eject each minute. The heart has the amazing capacity to adjust its cardiac output to meet the varying needs of the body. During sleep, the cardiac output declines, and during exercise, it increases markedly. The ability to increase cardiac output during increased activity is called the *cardiac reserve*. For example, competitive swimmers and long-distance runners have large cardiac reserves. During exercise, the cardiac output of these athletes rapidly increases to as much as five to six times their resting level.[3] In sharp contrast with healthy athletes, people with heart failure often use their cardiac reserve at rest. For them, just climbing a flight of stairs may cause shortness of breath because they have exceeded their cardiac reserve.

Control of Cardiac Performance and Output

Cardiac output, which is the major determinant of cardiac performance, reflects how often the heart beats each minute (heart rate) and how much blood it pumps with each beat (stroke volume) and can be expressed as the product of the heart rate and stroke volume (*i.e.,* cardiac output = heart rate × stroke volume). The heart rate is regulated by a balance between the activity of the sympathetic nervous system, which produces an increase in heart rate, and the parasympathetic nervous system, which slows it down, whereas the stroke volume is a function of preload, afterload, and myocardial contractility.[3–6]

Preload and Afterload. The work that the heart performs consists mainly of ejecting blood that has returned to the ventricles during diastole into the pulmonary or systemic circulation. It is determined largely by the loading conditions, or what are called the *preload* and *afterload*.

Preload reflects the volume or loading conditions of the ventricle at the end of diastole, just before the onset of systole. It is the volume of blood stretching the heart muscle at the end of diastole and is normally determined by the venous return to the heart. During any given cardiac cycle, the maximum volume of blood filling the ventricle is present at the end of diastole. Known as the *end-diastolic volume,* this volume causes an increase in the length of the myocardial muscle fibers. Within limits, as end-diastolic volume or preload increases, the stroke volume increases in accord with the Frank-Starling mechanism.

Afterload represents the force that the contracting heart muscle must generate to eject blood from the filled heart. The main components of afterload are the systemic (peripheral) vascular resistance and ventricular wall tension. When the systemic vascular resistance is elevated, as with arterial hypertension, an increased left intraventricular pressure must be generated to first open the aortic valve and then move blood out of the ventricle and into the systemic circulation. This increased pressure equates to an increase in ventricular wall stress or tension. As a result, excessive afterload may impair ventricular ejection and increase wall tension.

Myocardial Contractility. Myocardial contractility, also known as *inotropy,* refers to the contractile performance of the heart. It represents the ability of the contractile elements (actin and myosin filaments) of the heart muscle to interact and shorten against a load.[3–6] Contractility increases cardiac output independent of preload and afterload.

The interaction between the actin and myosin filaments during cardiac muscle contraction (*i.e.,* cross-bridge attachment and detachment) requires the use of energy supplied by the breakdown of adenosine triphosphate (ATP) and the presence of calcium ions (Ca^{++}). ATP provides the energy needed for cross-bridge formation during cardiac muscle contraction and for cross-bridge detachment during muscle relaxation.

As with skeletal muscle, when an action potential passes over the cardiac muscle fiber, the impulse spreads to the interior of the muscle fiber along the membranes of the transverse (T) tubules. The T tubule action potentials in turn act to cause release of Ca^{++} from the sarcoplasmic reticulum (Fig. 34.1). These Ca^{++} ions diffuse into the myofibrils and catalyze the chemical reactions that promote sliding of the actin and myosin filaments along one another to produce muscle shortening. In addition to the Ca^{++} ions released from the sarcoplasmic reticulum, a large quantity of extracellular Ca^{++} also diffuses into the sarcoplasm through voltage-dependent L-type Ca^{++} channels in T tubules at the time of the action potential. Without the extra Ca^{++} that enters through the L-type Ca^{++} channels, the strength of the cardiac contraction would be considerably weaker. Opening of the L-type Ca^{++} channels is facilitated by the second messenger cyclic adenosine monophosphate (cAMP), the formation of which is coupled to β-adrenergic receptors. The catecholamines (norepinephrine and epinephrine) exert their inotropic effects by binding to these receptors. The L-type calcium channel also contains several other types of drug receptors. The dihydropyridine Ca^{++} channel blocking drugs (*e.g.*, nifedipine) exert their effects by binding to one site, while diltiazem and verapamil appear to bind to closely related but not identical receptors in another region. Blockade of the Ca^{++} channels in cardiac muscle by these drugs results in a reduction in contractility throughout the heart and a decrease in sinus node pacemaker rate and in atrioventricular node conduction velocity.

Another mechanism that can modulate inotropy is the sodium ion (Na^+)/Ca^{++} exchange pump and the ATPase-dependent Ca^{++} pump on the myocardial cell membrane (see Fig. 34.1). These pumps transport Ca^{++} out of the cell, thereby preventing the cell from becoming overloaded with Ca^{++}. If Ca^{++} extrusion is inhibited, the rise in intracellular Ca^{++} can increase inotropy. Digitalis and related cardiac glycosides are inotropic agents that exert their effects by inhibiting the Na^+/potassium ion (K^+)–ATPase pump, which increases

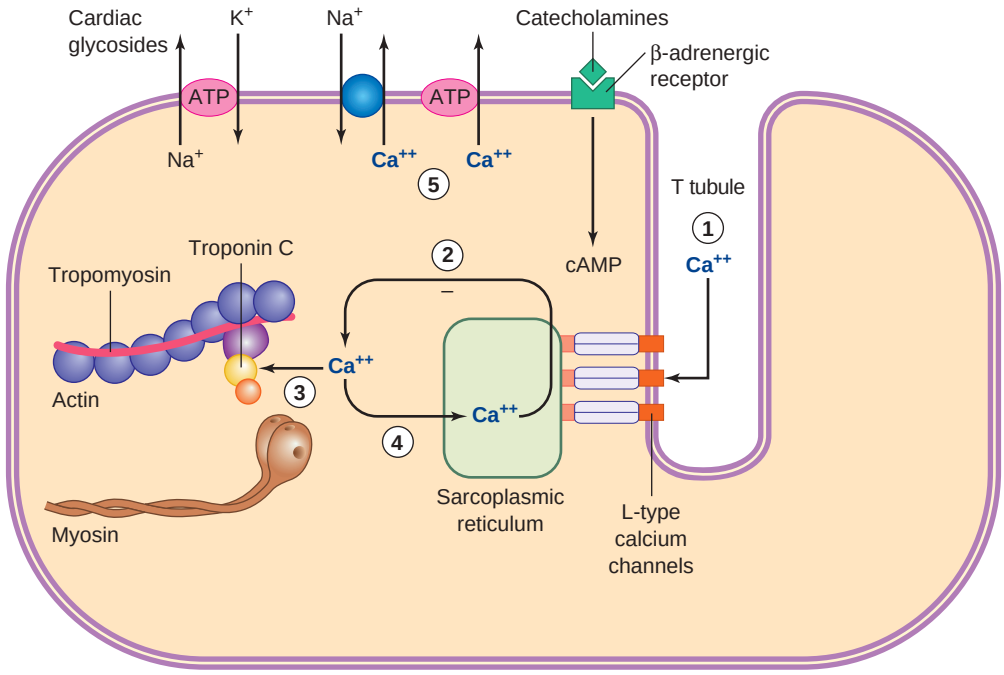


FIGURE 34.1 • Schematic representation of the role of calcium ions (Ca^{++}) in cardiac excitation–contraction coupling. The influx (site 1) of extracellular Ca^{++} through the L-type Ca^{++} channels in the T tubules during excitation triggers (site 2) release of Ca^{++} by the sarcoplasmic reticulum. This Ca^{++} binds to troponin C (site 3). The Ca^{++}–troponin complex interacts with tropomyosin to unblock active sites on the actin and myosin filaments, allowing cross-bridge attachment and contraction of the myofibrils (systole). Relaxation (diastole) occurs as a result of calcium reuptake by the sarcoplasmic reticulum (site 4) and extrusion of intracellular Ca^{++} by the sodium Na^+/Ca^{++} exchange transporter or, to a lesser extent, by the Ca^{++} ATPase pump (site 5). Mechanisms that raise systolic Ca^{++} increase the level of developed force (inotropy). Binding of catecholamines to β-adrenergic receptors increases Ca^{++} entry by phosphorylation of the Ca^{++} channels through a cyclic adenosine monophosphate (cAMP)-dependent second messenger mechanism. The cardiac glycosides increase intracellular Ca^{++} by inhibiting the Na^+/K^+–ATPase pump. The elevated intracellular Na^+ reverses the Na^+/Ca^{++} exchange transporter (site 5), so less Ca^{++} is removed from the cell. (Modified from Klabunde R. E. (2005). *Cardiovascular physiology concepts* (p. 46). Philadelphia, PA: Lippincott Williams & Wilkins.)

intracellular Na^+; this in turn leads to an increase in intracellular Ca^{++} through the Na^+/Ca^{++} exchange pump.

Right versus Left Ventricular Dysfunction

Heart failure has been classified according to the side of the heart (right ventricular or left ventricular) that is primarily affected (Fig. 34.2). Although the initial event that leads to heart failure may be primarily right or left ventricular in origin, long-term heart failure usually involves both sides. The pathophysiologic changes that occur in the myocardium itself, including the compensatory responses in conditions like myocardial infarction, are not significantly different between right and left ventricular dysfunction and are not addressed in detail in this section.

Daily measurement of weight can be used as a means of assessing fluid accumulation in a person with chronic heart failure. As a rule, a weight gain of more than 2 lb (0.90 kg) in 24 hours or 5 lb (2.27 kg) in 1 week is considered a sign of worsening failure.[9]

Right-sided heart failure also produces congestion of the viscera. As venous distention progresses, blood backs up in the hepatic veins that drain into the inferior vena cava, and the liver becomes engorged. This may cause hepatomegaly and right upper quadrant pain. In severe and prolonged right-sided failure, liver function is impaired and hepatic cells may die. Congestion of the portal circulation also may lead to engorgement of the spleen and the development of ascites.

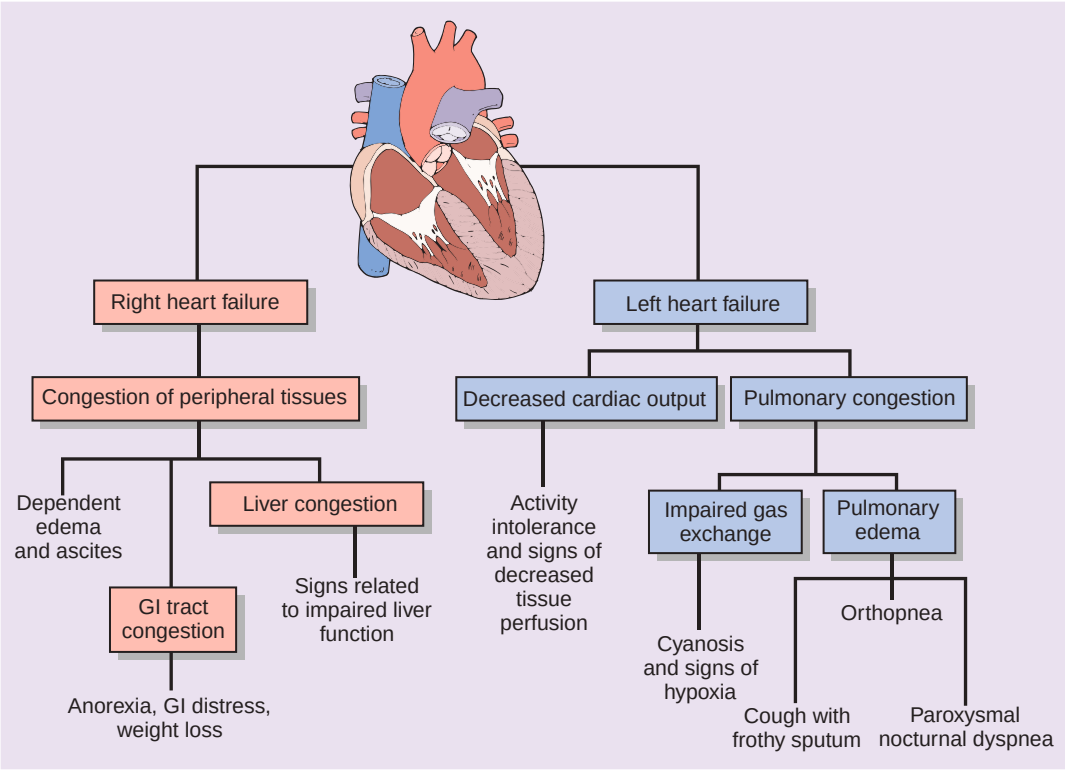


FIGURE 34.2 • Manifestations of left- and right-sided heart failure. (GI, gastrointestinal.)

Right Ventricular Dysfunction. Right-sided heart failure impairs the ability to move deoxygenated blood from the systemic circulation into the pulmonary circulation. Consequently, when the right ventricle fails, there is a reduction in the amount of blood moved forward into the pulmonary circulation and then into the left side of the heart, ultimately causing a reduction of left ventricular cardiac output. Also, if the right ventricle does not move the blood forward, there is accumulation or congestion of blood into the systemic venous system. This causes an increase in right ventricular end-diastolic, right atrial, and systemic venous pressures. A major effect of right-sided heart failure is the development of peripheral edema (see Fig. 34.2). Because of the effects of gravity, the edema is most pronounced in the dependent parts of the body. When the person is in the upright position, edema is seen in the lower extremities; when the person is supine, the edema is seen in the area over the sacrum. The accumulation of edema fluid is evidenced by a gain in weight (*i.e.,* 1 pint [568 mL] of accumulated fluid results in a 1 lb [0.45 kg] weight gain).

Congestion of the gastrointestinal tract may interfere with digestion and absorption of nutrients, causing anorexia and abdominal dis- comfort. The jugular veins, which are above the level of the heart, are normally not visible in the standing position or when sitting with the head at higher than a 30-degree angle. In severe right-sided failure, the external jugular veins become distended and can be visualized when the person is sitting up or standing.

The causes of right ventricular dysfunction include conditions that impede blood flow into the lungs or compromise the pumping effectiveness of the right ventricle. Left ventricular failure is the most common cause of right ventricular failure. Sustained pulmonary hypertension also causes right ventricular dysfunction and failure. Pulmonary hypertension occurs in people with chronic pulmonary disease, severe pneumonia, pulmonary embolus, or aortic or mitral stenosis. Other common causes include stenosis or regurgitation of the tricus- pid or pulmonic valves, right ventricular infarction, and cardiomyopathy.

Left Ventricular Dysfunction. Left-sided heart failure impairs the movement of blood from the low-pressure pulmonary circulation into the high-pressure arterial side of the systemic circulation. With impairment of left heart function, there is a decrease in cardiac output to the systemic circulation. Blood accumulates in the left ventricle, left atrium, and pulmonary circulation, which causes an elevation in pulmonary venous pressure (see Fig. 34.2). When the pressure in the pulmonary capillaries (normally approximately 10 mm Hg) exceeds the capillary osmotic pressure (normally approximately 25 mm Hg), there is a shift of intravascular fluid into the interstitium of the lung and development of pulmonary edema (Fig. 34.3).

The most common causes of left ventricular dysfunction are hypertension and acute myocardial infarction. Left ventricular heart failure and pulmonary congestion can develop very rapidly in people with acute myocardial infarction. Even when the infarcted area is small, there may be a surrounding area of ischemic tissue. This may result in large areas of ventricular wall hypokinesis or akinesis and rapid onset of pulmonary congestion and edema. Stenosis or regurgitation of the aortic or mitral valve also creates the level of left-sided backflow that results in pulmonary congestion. As pulmonary pressure rises as a result of congestion, it may progress to produce right-sided heart failure.

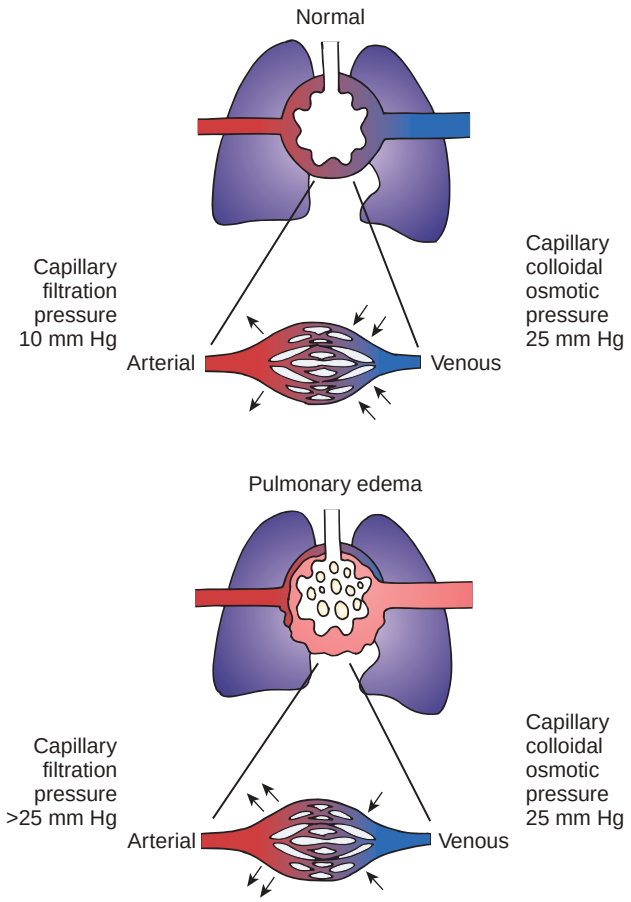


FIGURE 34.3 • Mechanism of respiratory symptoms in left-sided heart failure. In the normal exchange of fluid in the pulmonary capillaries (**top**), the capillary filtration pressure that moves fluid out of the capillary into the lung is less than the capillary colloidal osmotic pressure that pulls fluid back into the capillary. Development of pulmonary edema (**bottom**) occurs when the capillary filtration pressure exceeds the capillary colloidal osmotic pressure that pulls fluid back into the capillary.

Compensatory Mechanisms

In heart failure, the cardiac reserve is largely maintained through compensatory or adaptive responses such as the Frank-Starling mechanism, activation of neurohumoral influences such as the sympathetic nervous system reflexes, the renin–angiotensin–aldosterone mechanism, NPs, locally produced vasoactive substances, and myocardial hypertrophy and remodeling[9] (Fig. 34.4).

The first of these adaptations occurs rapidly over minutes to hours of myocardial dysfunction and may be adequate to maintain the overall pumping performance of the heart at relatively normal levels. Myocardial hypertrophy and remodeling occur slowly over months to years and play an important role in the long-term adaptation to hemodynamic overload. In the failing heart, early decreases in cardiac function may go unnoticed because these compensatory mechanisms maintain the cardiac output. However, these mechanisms contribute not only to the adaptation of the failing heart but also to the pathophysiology of heart failure.[9]

Frank-Starling Mechanism. The Frank-Starling mechanism operates through an increase in preload (Fig. 34.5). With increased diastolic filling, there is increased stretching of the myocardial fibers and more optimal approximation of the heads on the thick myosin filaments to the troponin binding sites on the thin actin filaments, with a resultant increase in the force of the next contraction. In the normally functioning heart, the Frank-Starling mechanism serves to match the outputs of the two ventricles. As illustrated in Figure 34.5, there is no one single Frank-Starling curve.[4] An increase in contractility, or inotropy, will increase cardiac output at any end-diastolic volume, causing the curve to move up and to the left, whereas a decrease in inotropy will cause the curve to move down and to the right. In heart failure, inotropy is decreased compared with normal. Thus, the stroke volume will not be as high as with normal inotropy, regardless of the increase in preload.

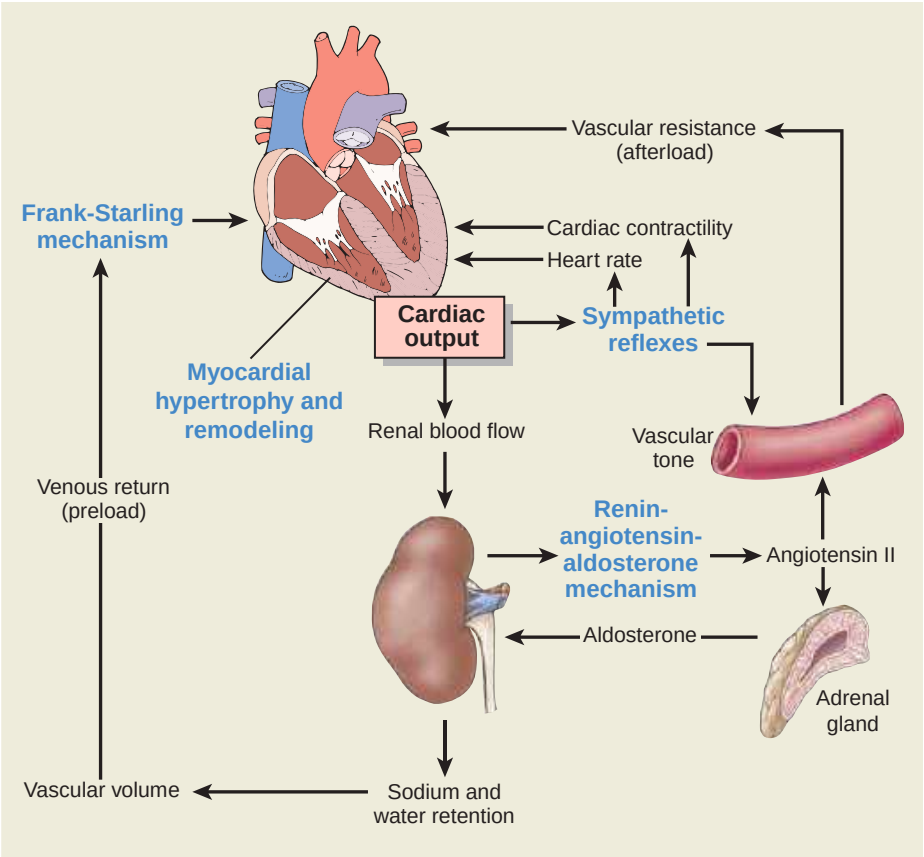


FIGURE 34.4 • Compensatory mechanisms in heart failure. The Frank-Starling mechanism, sympathetic reflexes, renin–angiotensin–aldosterone mechanism, and myocardial hypertrophy function in maintaining cardiac output for the failing heart.

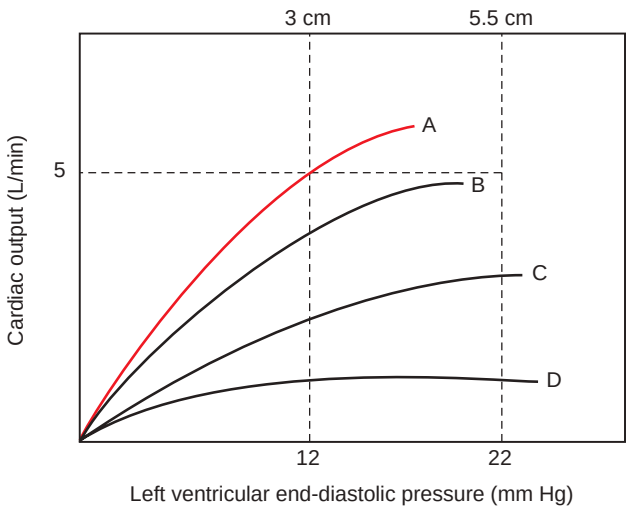


FIGURE 34.5 • Left ventricular function curves. *Curve A:* Normal function curve, with a normal cardiac output and optimal left ventricular end-diastolic (LVED) filling pressure. *Curve B:* Compensated heart failure with normal cardiac output at higher LVED pressures. *Curve C:* Decompensated heart failure with a decrease in cardiac output and elevated LVED, with eventual elevation of pulmonary capillary pressure and development of pulmonary congestion. *Curve D:* Cardiogenic shock, with an extreme decrease in cardiac output and marked increase in LVED pressures.

In heart failure, a decrease in cardiac output and renal blood flow leads to increased sodium and water retention, a resultant increase in vascular volume and venous return to the heart, and an increase in ventricular end-diastolic volume. Within limits, as preload and ventricular end-diastolic volume increase, there is a resultant increase in cardiac output. Although this may preserve the resting cardiac output, the resulting chronic elevation of left ventricular end-diastolic pressure is transmitted to the atria and the pulmonary circulation, causing pulmonary congestion.

An increase in muscle stretch, as occurs with the Frank-Starling mechanism, also causes an increase in ventricular wall tension with a resultant increase in myocardial oxygen consumption. Because increased wall tension increases myocardial oxygen requirements, it can produce ischemia and contribute to further impairment of inotropy, moving the Frank-Starling curve farther down and to the right (see Fig. 34.5). In this situation, the increase in preload is no longer contributing to compensation but rather causing heart failure to worsen. The use of diuretics in people with heart failure helps to reduce vascular volume and ventricular filling, thereby unloading the heart and reducing ventricular wall tension.

Sympathetic Nervous System Activity. Stimulation of the sympathetic nervous system plays an important role in the compensatory response to decreased cardiac output and stroke volume.[9,11] Both cardiac sympathetic tone and catecholamine (epinephrine and norepinephrine) levels are elevated during the late stages of most forms of heart failure. By direct stimulation of heart rate and cardiac contractility, regulation of vascular tone, and enhancement of renal sodium and water retention, the sympathetic nervous system initially helps to maintain perfusion of the various body organs. In people who progress to more severe heart failure, blood is diverted to the more critical cerebral and coronary circulations.

Although the sympathetic nervous system response is meant to augment blood pressure and cardiac output and is the most immediate compensatory mechanism, it can become maladaptive. An increase in sympathetic activity by stimulation of the β-adrenergic receptors of the heart leads to tachycardia, vasoconstriction, and cardiac arrhythmias. Acutely, tachycardia significantly increases the workload of the heart, thus increasing myocardial oxygen demand and leading to cardiac ischemia, myocyte damage, and decreased contractility (inotropy).[11] Cardiac ischemia and cardiomyopathy both contribute to worsening of heart failure. By promoting arrhythmias, the catecholamines released with sympathetic nervous system stimulation also may contribute to the high rate of sudden death seen with heart failure.

There is evidence that prolonged sympathetic stimulation may also lead to desensitization of β-adrenergic receptors without affecting α-adrenergic receptors.[4] Even though circulating norepinephrine levels are increased in people with heart failure, the lack of functioning β-adrenergic receptors in relation to α-adrenergic receptors may lead to vasoconstriction and an increase in systemic vascular resistance. An increase in systemic vascular resistance causes an increase in cardiac afterload and ventricular wall stress, thus increasing myocardial oxygen consumption. Other effects include decreased renal perfusion and additional augmentation of the renin–angiotensin–aldosterone system, as well as decreased blood flow to skin, muscle, and abdominal organs.[11]

Renin–Angiotensin–Aldosterone Mechanism. One of the most important effects of lowered cardiac output in heart failure is a reduction in renal blood flow and glomerular filtration rate, which leads to sodium and water retention. With decreased renal blood flow, there is a progressive increase in renin secretion by the kidneys with parallel increases in circulating levels of angiotensin II. The increased concentration of angiotensin II contributes directly to a generalized and excessive vasoconstriction, as well as facilitating norepinephrine release and inhibiting reuptake of norepinephrine by the sympathetic nervous system.[11]

Angiotensin II also provides a powerful stimulus for aldosterone production by the adrenal cortex. Aldosterone increases tubular reabsorption of sodium, with an accompanying increase in water retention. Because aldosterone is metabolized in the liver, its levels are further increased when heart failure causes liver congestion. Angiotensin II also increases the level of antidiuretic hormone (ADH), which serves as a vasoconstrictor and inhibitor of water excretion. In heart failure, the progressive accumulation of fluid leads to ventricular dilation and increased wall tension. The increased oxygen demand that accompanies increased wall tension eventually outweighs the compensatory Frank-Starling mechanism, reducing inotropy and progressing heart failure.

In addition to their individual effects on sodium and water balance, angiotensin II and aldosterone are also involved in regulating the inflammatory and reparative processes that follow tissue injury.[12] In this capacity, they stimulate inflammatory cytokine production (*e.g.,* tumor necrosis factor [TNF] and interleukin-6), attract inflammatory cells (*e.g.,* neutrophils and macrophages), activate macrophages at sites of injury and repair, and stimulate the growth of fibroblasts and synthesis of collagen fibers. Fibroblast and collagen deposition results in ventricular hypertrophy and myocardial wall fibrosis, which decreases compliance (*i.e.,* increases stiffness), ultimately causing inappropriate remodeling of the heart and progression of both systolic and diastolic ventricular dysfunction.[13] Thus, the progression of heart failure may be augmented by aldosterone-mediated effects on both the vasculature and myocardium.

Natriuretic Peptides. The heart muscle produces and secretes a family of related peptide hormones, the cardiac natriuretic hormones or NPs, that have potent diuretic, natriuretic, and vascular smooth muscle effects and also interact with other neurohumoral mechanisms that affect cardiovascular function. Two of the four known NPs most commonly associated with heart failure are atrial natriuretic peptide (ANP) and brain natriuretic peptide (BNP).[14]

As the name indicates, ANP is released from atrial cells in response to atrial stretch, pressure, or fluid overload. BNP is primarily secreted by the ventricles as a response to increased ventricular pressure or fluid overload. In early heart failure, NT-proBNP can be detected as a precursor for BNP in the blood. Although the NPs are not secreted from the same chambers in the heart, they have very similar functions. In response to increased chamber stretch and pressure, they promote rapid and transient natriuresis and diuresis through an increase in the glomerular filtration rate and an inhibition of tubular sodium and water reabsorption.

The NPs also facilitate complex interactions with the neurohormonal system, inhibiting the sympathetic nervous system, the renin–angiotensin–aldosterone system, endothelin inflammatory cytokines, and vasopressin.[14] Suppression of the sympathetic nervous system causes both venous and arterial dilation with consequent reduction in venous return to the heart (decreased preload) and cardiac filling pressures, and a decrease in afterload (arterial vasodilation). Inhibition of angiotensin II and vasopressin by the NPs reduces renal fluid retention. In addition, the NPs directly affect the central nervous system and the brain, inhibiting the secretion of vasopressin and the function of the salt appetite and thirst center.[14]

Circulating levels of both ANP and BNP are reportedly elevated in people with heart failure. BNP and NT-pro BNP levels can be detected through blood work and commercial assays. The concentrations are well correlated with the extent of ventricular dysfunction, increasing up to 30-fold in people with advanced heart disease.[14] Assays of BNP are used clinically in the diagnosis of heart failure and to predict the severity of the condition. Many of the medications used to treat heart failure (*e.g.,* diuretics, such as spironolactone, and the angiotensin-converting enzyme [ACE] inhibitors) reduce BNP concentrations. Therefore, many people with chronic stable heart failure have BNP levels in the normal diagnostic range. However, digoxin and beta-blockers appear to increase BNP levels. There are drugs designed to inhibit degradation on NPs as a potential for therapy.

Endothelins. The endothelins, released from the endothelial cells throughout the circulation, are potent vasoconstrictor peptides. Like angiotensin II, endothelin can also be synthesized and released by a variety of cell types, such as cardiac myocytes. Four endothelin peptides (endothelin-1 [ET-1], ET-2, ET-3, and ET-4) have been identified.[14] However, all of their physiological functions remain unclear. It has been found that the endothelins induce vascular smooth muscle cell proliferation and cardiac myocyte hypertrophy; increase the release of ANP, aldosterone, and catecholamines; and exert antinatriuretic effects on the kidneys. Production of ET-1 is regulated by many factors that are significant for cardiovascular function and have implications for heart failure. For example, it is enhanced by angiotensin II, vasopressin, and norepinephrine and by factors such as shear stress and endothelial stretching.[14] Plasma ET-1 levels also correlate directly with pulmonary vascular resistance, and it is thought that the peptide may play a role in mediating pulmonary hypertension in people with heart failure. There are at least two types of endothelin receptors—type A and type B.[14] Type A receptor is associated with smooth muscle constriction and hypertrophy while Type B receptor is associated with vasodilation. Since ET-1 can act on the heart to cause hypertrophy and sodium and water retention, an endothelin receptor antagonist is now available for use in people with pulmonary arterial hypertension due to severe heart failure.

Inflammatory Mediators. There is ongoing research examining the relationship between inflammatory markers, especially C-reactive protein (CRP), and heart failure. Elevated CRP levels have been associated with adverse consequences in people with heart failure. They have also been shown to be predictive of the development of heart failure in high-risk groups. Of particular interest are the interactions between CRP and mediators, such as angiotensin II and norepinephrine. This inflammatory relationship continues to be examined. However, it is difficult to test since it is not understood how to decrease the inflammatory effect in heart failure.

Myocardial Hypertrophy and Remodeling. The development of myocardial hypertrophy constitutes one of the principal mechanisms by which the heart compensates for an increase in workload.[9] Although ventricular hypertrophy improves the work performance of the heart, it is also an important risk factor for subsequent cardiac morbidity and mortality. Inappropriate hypertrophy and remodeling can result in changes in structure (*i.e.,* muscle mass, chamber dilation) and function (*i.e.,* impaired systolic or diastolic function) that often lead to further pump dysfunction and hemodynamic overload.

Myocardial hypertrophy and remodeling involve a series of complex events at both the molecular and cellular levels. The myocardium is composed of myocytes, or muscle cells, and nonmyocytes. The myocytes are the functional units of cardiac muscle. Their growth is limited by an increment in cell size, as opposed to an increase in cell number. The nonmyocytes include cardiac macrophages, fibroblasts, vascular smooth muscle, and endothelial cells. These cells, which are present in the interstitial space, remain capable of an increase in cell number and provide support for the myocytes. The nonmyocytes also determine many of the inappropriate changes that occur during myocardial hypertrophy. For example, uncontrolled cardiac fibroblast growth is associated with increased synthesis of collagen fibers, myocardial fibrosis, and ventricular wall stiffness. Not only does ventricular wall stiffness increase the workload of the heart, but the fibrosis and remodeling that occur may lead to electrical conduction abnormalities in which the heart contracts in an uncoordinated manner, known as *cardiac dyssynchrony,* causing reduced systolic heart function.[14]

Recent research has focused on the type of hypertrophy that develops in people with heart failure. At the cellular level, cardiac muscle cells respond to stimuli from stress placed on the ventricular wall by pressure and volume overload by initiating several different processes that lead to hypertrophy. These include stimuli that produce the following:

- S*ymmetric hypertrophy* with a proportionate increase in muscle length and width, as occurs in athletes
- *Concentric hypertrophy* with an increase in wall thickness, as occurs in hypertension
- *Eccentric hypertrophy* with a disproportionate increase in muscle length, as occurs in dilated cardiomyopathy[15] (Fig. 34.6)

When the primary stimulus for hypertrophy is *pressure overload,* the increase in wall stress leads to parallel replication of myofibrils, thickening of the individual myocytes, and concentric hypertrophy. Concentric hypertrophy may preserve systolic function for a time, but eventually the work performed by the ventricle exceeds the vascular reserve, predisposing to ischemia. When the primary stimulus is *ventricular volume overload,* the increase in wall stress leads to replication of myofibrils in series, elongation of the cardiac muscle cells, and eccentric hypertrophy. Eccentric hypertrophy leads

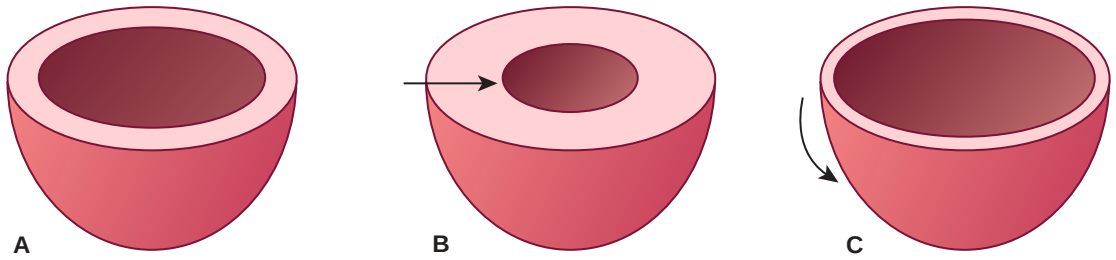


FIGURE 34.6 • Different types of myocardial hypertrophy. (**A**) Normal symmetric hypertrophy with proportionate increases in myocardial wall thickness and length. (**B**) Concentric hypertrophy with a disproportionate increase in wall thickness. (**C**) Eccentric hypertrophy with a disproportionate decrease in wall thickness and ventricular dilation.

to a decrease in ventricular wall thickness with an increase in diastolic volume and wall tension.

Acute Heart Failure Syndromes

The acute heart failure syndromes (AHFS) are "defined as gradual or rapid change in heart failure signs and symptoms resulting in a need for urgent therapy."[14] These symptoms are primarily the result of severe pulmonary edema due to elevated left ventricular filling pressures, with or without a low cardiac output.[14] The syndromes are among the most common disorders seen in emergency departments, and chronic heart failure, often complicated by episodes of acute worsening, is the most common cause of the syndrome.

AHFS are thought to encompass three different types of conditions:

1. Worsening of chronic systolic or diastolic dysfunction that appears to respond to treatment, approximately 80%
2. New-onset acute heart failure that occurs secondary to a precipitating event such as a large myocardial infarction or a sudden increase in blood pressure superimposed on a noncompliant left ventricle
3. Worsening of end-stage/advanced heart failure that is refractory to treatment, with predominantly left ventricular systolic dysfunction associated with a low-output state[16,17]

The difference between new-onset AHFS and AHFS caused by chronic heart failure is in the degree of physiologic response, which is more pronounced in the new-onset AHFS and subtler in chronic heart failure because of the compensatory pathophysiology. For example, with new-onset AHFS, the person will have a stronger sympathetic response with enhanced pulmonary vascular permeability causing rapid and dramatic symptoms of pulmonary edema. Because many compensatory mechanisms operate in people with chronic heart failure, they tolerate higher pulmonary vascular pressures. Chronic changes in neurohormonal regulation lead to stronger activation of the angiotensin–aldosterone system with a resultant volume overload, and venous congestion is more prominent in both the systemic and pulmonary circulations.[16]

Clinical Manifestations of Heart Failure

The manifestations of heart failure depend on the extent and type of cardiac dysfunction that is present and the rapidity with which it develops. A person with previously stable compensated heart failure may develop signs of heart failure for the first time when the condition has advanced to a critical point, such as with a progressive increase in pulmonary hypertension in a person with mitral valve regurgitation. Overt heart failure also may be precipitated by conditions such as infection, emotional stress, uncontrolled hypertension, or fluid overload.[7] Many people with serious underlying heart disease, regardless of whether they have previously experienced heart failure, may be relatively asymptomatic as long they carefully adhere to their treatment regimen. A dietary excess of sodium is a frequent cause of sudden cardiac decompensation.

The manifestations of heart failure reflect the physiologic effects of the impaired pumping ability of the heart, decreased renal blood flow, and activation of the sympathetic compensatory mechanisms. The severity and progression of symptoms depend on the extent and type of dysfunction that is present (systolic versus diastolic, right- versus left-sided). The signs and symptoms include shortness of breath and other respiratory manifestations, fatigue and limited exercise tolerance, fluid retention and edema, cachexia and malnutrition, and cyanosis. People with severe heart failure may exhibit diaphoresis and tachycardia.

Respiratory Manifestations

Shortness of breath due to congestion of the pulmonary circulation is one of the major manifestations of left-sided heart failure. Perceived shortness of breath (*i.e.,* breathlessness) is called *dyspnea*. Dyspnea related to an increase in activity is called *exertional dyspnea. Orthopnea* is shortness of breath that occurs when a person is supine. Gravitational forces cause fluid to become sequestered in the lower legs and feet when the person is standing or sitting. When the person assumes the recumbent position, fluid from the legs and dependent parts of the body is mobilized and redistributed to an already distended pulmonary circulation. *Paroxysmal nocturnal dyspnea* is a sudden attack of dyspnea that occurs during sleep. It disrupts sleep, and the person awakens with a feeling of extreme

suffocation that resolves when he or she sits up. Initially, the experience may be interpreted as awakening from a bad dream.

A subtle and often overlooked symptom of heart failure is a chronic dry, nonproductive cough that becomes worse when the person is lying down. Bronchospasm due to congestion of the bronchial mucosa may cause wheezing and difficulty in breathing. This condition is sometimes referred to as *cardiac asthma*.[7]

Cheyne-Stokes Respiration. Cheyne-Stokes respiration is a pattern of periodic breathing characterized by gradual increase in depth (and sometimes rate) of breathing to a maximum, followed by a decrease resulting in apnea. Although no longer associated solely with heart failure, it is recognized as an independent risk factor for worsening of heart failure. It has been suggested that Cheyne-Stokes respirations may not be just a marker for increasing severity of heart failure but may also aggravate it.[14] During sleep, Cheyne-Stokes breathing causes recurrent awakening and thereby reduces slow-wave and rapid eye movement (REM) sleep. The recurrent cycling of hypoventilation/apnea and hyperventilation may also increase sympathetic activity and predispose to arrhythmias. Nocturnal oxygen has been seen to improve sleep, exercise tolerance, and cognitive function.

Acute Pulmonary Edema. Acute pulmonary edema is the most dramatic symptom of AHFS. It is a life-threatening condition in which capillary fluid moves into the alveoli.[7] The accumulated fluid in the alveoli and airways causes lung stiffness, makes lung expansion more difficult, and impairs the gas exchange function of the lung. With the decreased ability of the lungs to oxygenate the blood, the hemoglobin leaves the pulmonary circulation without being fully oxygenated, resulting in shortness of breath and cyanosis.

The person with severe pulmonary edema is usually seen sitting and gasping for air. The pulse is rapid, the skin is moist and cool, and the lips and nail beds are cyanotic. As the pulmonary edema worsens and oxygen supply to the brain drops, confusion and stupor appear. Dyspnea and air hunger are accompanied by a productive cough with frothy (resembling beaten egg whites) and often blood-tinged sputum—the effect of air mixing with the serum albumin and red blood cells that have moved into the alveoli. The movement of air through the alveolar fluid produces fine crepitant sounds called *crackles,* which can be heard with chest auscultation. As fluid moves into the larger airways, the crackles become louder and coarser.

Fatigue, Weakness, and Mental Confusion

Fatigue and weakness often accompany diminished output from the left ventricle. Cardiac fatigue is different from general fatigue in that it usually is not present in the morning but appears and progresses as activity increases during the day.

In acute or severe left-sided failure, cardiac output may fall to levels that are insufficient for providing the brain with adequate oxygen, and there are indications of mental confusion and disturbed behavior. Confusion, impairment of memory, anxiety, restlessness, and insomnia are common in elderly persons with advanced heart failure, particularly in those with cerebral atherosclerosis. These symptoms may confuse the diagnosis of heart failure in older adults because of their myriad of other causes associated with aging.

Fluid Retention and Edema

Many of the manifestations of heart failure result from the increased capillary pressures (increased hydrostatic pressures) that develop in the peripheral circulation in people with right-sided heart failure and in the pulmonary circulation in people with left-sided heart failure. The increased capillary pressure reflects an overfilling of the vascular system because of increased sodium and water retention and venous congestion, referred to earlier as *backward* failure, resulting from impaired cardiac output.[7,14]

Nocturia is a nightly increase in urine output that occurs relatively early in the course of heart failure. It occurs because of the increased cardiac output, renal blood flow, and glomerular filtration rate that follow the increased blood return to the heart when the person is in a supine position. *Oliguria,* which is a decrease in urine output, is a late sign related to a severely reduced cardiac output and resultant renal failure.

Transudation of fluid into the pleural cavity (hydrothorax) or the peritoneal cavity (ascites) may occur in people with advanced heart failure. Because the pleural veins drain into both the systemic and pulmonary venous beds, hydrothorax is observed more commonly in persons with hypertension involving both venous systems.[7,14] Pleural effusion occurs as the excess fluid in the lung interstitial spaces crosses the visceral pleura, which in turn overwhelms the capacity of the pulmonary lymphatic system. Ascites occurs in people with increased pressure in the hepatic veins and veins draining the peritoneum. It usually reflects right ventricular failure and long-standing elevation of systemic venous pressure in chronic heart failure.[7,14]

Cachexia and Malnutrition

Cardiac cachexia is a condition of malnutrition and tissue wasting that occurs in people with end-stage heart failure. A number of factors probably contribute to its development, including the fatigue and depression that interfere with food intake, congestion of the liver and gastrointestinal structures that impairs digestion and absorption and produces feelings of fullness, and the circulating toxins and mediators released from poorly perfused tissues that impair appetite and contribute to tissue wasting.

Cyanosis

Cyanosis is the bluish discoloration of the skin and mucous membranes caused by excess desaturated hemoglobin in the blood; it often is a late sign of heart failure. Cyanosis may be central, caused by arterial desaturation resulting from impaired pulmonary gas exchange, or peripheral, caused by venous desaturation resulting from extensive extraction of oxygen at the capillary level. Central cyanosis is caused by conditions that impair oxygenation of the arterial blood, such as pulmonary edema, left heart failure, or right-to-left cardiac

shunting. Peripheral cyanosis is caused by conditions such as low-output failure that result in delivery of poorly oxygenated blood to the peripheral tissues, or by conditions such as peripheral vasoconstriction that cause excessive removal of oxygen from the blood. Central cyanosis is best monitored in the lips and mucous membranes because these areas are not subject to conditions, such as a cold environment, that cause peripheral cyanosis. People with right-sided or left-sided heart failure may develop cyanosis especially around the lips and in the peripheral parts of the extremities.

Arrhythmias and Sudden Cardiac Death

Both atrial and ventricular arrhythmias occur in people with heart failure. Atrial fibrillation is the most common arrhythmia. Clinical manifestations associated with atrial fibrillation are related to loss of atrial contraction, tachycardia, irregular heart rate, and symptoms related to a drop in blood pressure.[12,13,18] There is also strong evidence that people with heart failure are at increased risk for sudden cardiac arrest; that is, unwitnessed death or death that occurs within 1 hour of the symptom onset.[12,13,18] In people with ventricular dysfunction, sudden death is caused most commonly by ventricular tachycardia or ventricular fibrillation.[12,13,18]

Diagnosis and Treatment

Diagnosis

Diagnostic methods in heart failure are directed toward establishing the cause of the disorder and determining the extent of the dysfunction. Medical guidelines for diagnosis and treatment are clearly described in the AHA guidelines for heart failure management.[1,2] Because heart failure represents the failure of the heart as a pump and can occur in the course of a number of heart diseases or other systemic disorders, the diagnosis of heart failure often is based on signs and symptoms related to the failing heart itself, such as shortness of breath and fatigue. The functional classification of the New York Heart Association (NYHA) is one guide to classifying the extent of dysfunction.

The NYHA functional classification classifies dysfunction into four classes:[1,2]

1. Class I—People who have known heart disease without symptoms during ordinary activity
2. Class II—People who have heart disease who have slight limitations but not extreme fatigue, palpitations, dyspnea, or angina pain during regular activity
3. Class III—People with heart disease who are comfortable at rest but ordinary activity does result in fatigue, palpitations, dyspnea, and angina pain
4. Class IV—People who have marked progressive cardiac disease and are not comfortable at rest or minimal activity[1,2]

The methods used in the diagnosis of heart failure include risk factor assessment, history and physical examination, laboratory studies, electrocardiography, chest radiography, and echocardiography. The history should include information related to dyspnea, cough, nocturia, generalized fatigue, and other signs and symptoms of heart failure. A complete physical examination includes assessment of heart rate, heart sounds, blood pressure, jugular veins for venous congestion, lungs for signs of pulmonary congestion, and lower extremities for edema. Laboratory tests are used in the diagnosis of anemia and electrolyte imbalances, and to detect signs of chronic liver congestion. Measurements of BNP and NT-proBNP can be useful if the diagnosis of heart failure is uncertain and as risk stratification. The use of serial BNP or NT-proBNP levels has not yet been well established.[2]

Echocardiography plays a key role in assessing right and left ventricular wall motion (normal, akinesis, or hypokinesis), wall thickness, ventricular chamber size, valve function, heart defects, ejection fraction, and pericardial disease.[2] Electrocardiographic findings may indicate atrial or ventricular hypertrophy, underlying disorders of cardiac rhythm, or conduction abnormalities such as right or left bundle branch block. Radionuclide ventriculography and cardiac angiography are recommended if there is reason to suspect coronary artery disease as the underlying cause for heart failure. Chest x-rays provide information about the size and shape of the heart and pulmonary vasculature. The cardiac silhouette can be used to detect cardiac hypertrophy and dilatation. Chest x-rays can indicate the relative severity of the failure by revealing if pulmonary edema is predominantly vascular or interstitial, or has advanced to the alveolar and bronchial stages. Cardiac magnetic resonance imaging (CMRI) and cardiac computed tomography (CCT) are used to document ejection fraction, ventricular preload, and regional wall motion.

Invasive hemodynamic monitoring may be used for assessment in acute, life-threatening episodes of heart failure. These monitoring methods include central venous pressure (CVP), pulmonary artery pressure monitoring, thermodilution measurements of cardiac output, and intra-arterial measurements of blood pressure. CVP reflects the amount of blood returning to the heart. Measurements of CVP are best obtained by a catheter inserted into the right atrium through a peripheral vein, or by the right atrial port (opening) in a pulmonary artery catheter. This pressure is decreased in hypovolemia and increased in right heart failure. The changes that occur in CVP over time usually are more significant than the absolute numeric values obtained during a single reading.

Ventricular volume pressures are obtained by means of a flow-directed, balloon-tipped pulmonary artery catheter. This catheter is introduced through a peripheral or central vein and then advanced into the right atrium. The balloon is then inflated with air, enabling the catheter to float through the right ventricle into the pulmonary artery until it becomes wedged in a small pulmonary vessel (Fig. 34.7). With the balloon inflated, the catheter monitors pulmonary capillary pressures (also called *pulmonary capillary wedge pressure* [PCWP]), which is in direct communication with pressures from the left heart. The pulmonary capillary pressures provide a means of assessing the pumping ability of the left heart.

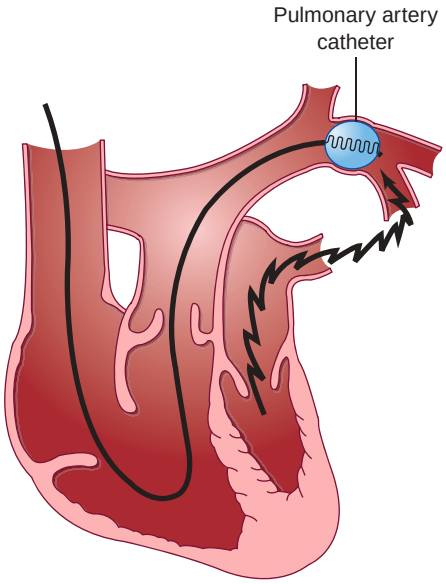


FIGURE 34.7 • Balloon-tipped pulmonary artery catheter positioned in a small pulmonary vessel. The PCWP, which reflects the left ventricular diastolic pressure, is measured with the balloon inflated.

Intra-arterial blood pressure monitoring provides a means for continuous monitoring of blood pressure. It is used in people with acute heart failure when aggressive intravenous medication therapy or a mechanical assist device is required. Measurements are obtained through a small catheter inserted into a peripheral artery, usually the radial artery. The catheter is connected to a pressure transducer, and beat-by-beat measurements of blood pressure are recorded. The monitoring system displays the contour of the pressure waveform and the systolic, diastolic, and mean arterial pressures, along with the heart rate and rhythm.

Remember Mr. Brown from the unit opener case study? He was diagnosed with high blood pressure and hypercholesteremia. His subsequent cardiac catheterization revealed mild ischemic occlusion that did not meet the criteria for a cardiac stent or angioplasty. This result, along with his low ejection fraction of 40%, indicated that the supply of oxygen to his heart muscle was moderately impaired, reducing the force developed by the left ventricle. Therefore, he was diagnosed with ischemic cardiomyopathy and was classified as having Stage B (American Heart Association) and Class II (New York Heart Association) heart failure.

Treatment

The goals of treatment are determined by the rapidity of onset and severity of the heart failure. People with AHFS require urgent therapy directed at stabilizing and correcting the cause of the cardiac dysfunction. For people with chronic heart failure, the goals of treatment are directed toward relieving the symptoms, improving the quality of life, and reducing or eliminating risk factors (*e.g.,* hypertension, diabetes, obesity) with a long-term goal of slowing, halting, or reversing the cardiac dysfunction.[1,2,14,17]

Treatment measures for both acute and chronic heart failure include nonpharmacologic and pharmacologic approaches. Mechanical support devices, including the intra-aortic balloon pump (for acute failure) and the ventricular assist device (VAD), sustain life in people with severe heart failure. Heart transplantation remains the treatment of choice for many people with end-stage heart failure.

Nonpharmacologic Methods. Exercise intolerance is typical in people with chronic heart failure.[19] Consequently, individualized exercise training is important to maximize muscle conditioning. People who are not accustomed to exercise and those with more severe heart failure are started at a lower intensity and shorter sessions than those who are mostly asymptomatic. Sodium and fluid restriction and weight management are important for all people with heart failure; the degree of sodium and fluid restriction is individualized to the severity of heart failure. Counseling, health teaching, and ongoing evaluation programs help people with heart failure to manage and cope with their treatment regimen.

Pharmacologic Treatment. Once heart failure is moderate to severe, pharmacologic in conjunction with nonpharmacologic management is important to prevent and treat acute heart failure and manage chronic heart failure. Evidence-based agents recommended for treatment and management include diuretics, ACE inhibitors or angiotensin II receptor blockers, β-adrenergic blockers, and digoxin.[1,2,14,17,20] The choice of pharmacologic agents is based on symptomatology of the person.

Diuretics are among the most frequently prescribed medications for moderate to severe heart failure.[1,2,20] They promote the excretion of fluid and help to sustain cardiac output and tissue perfusion by reducing preload and allowing the heart to operate at a more optimal part of the Frank-Starling curve. Thiazide and loop diuretics are used. In emergencies, such as acute pulmonary edema, loop diuretics such as furosemide can be administered intravenously. When given as a bolus infusion, intravenous furosemide acts within minutes to increase venous capacitance so that right ventricular output and pulmonary capillary pressures are decreased.

The *ACE inhibitors,* which prevent the conversion of angiotensin I to angiotensin II, have been used effectively in the treatment of chronic heart failure. The renin–angiotensin–aldosterone system is activated early in the course of heart failure and plays an important role in its progression. It results in an increase in angiotensin II, which causes vasoconstriction, unregulated ventricular remodeling, and increased aldosterone production with a subsequent increase in sodium and water retention by the kidneys. ACE inhibitors have been shown to limit these harmful complications. The *angiotensin II receptor blockers* appear to have similar but more limited beneficial effects. They have the advantage of not causing a cough, which is a troublesome side effect of the ACE inhibitors for

many people. Aldosterone has a number of deleterious effects in people with heart failure. *Aldosterone receptor antagonists* may be used in combination with other agents for people with moderately severe to severe heart failure.

β-*Adrenergic receptor blocking drugs* are used to decrease left ventricular dysfunction associated with activation of the sympathetic nervous system. Large clinical trials have shown that long-term therapy with β-adrenergic receptor blocking agents reduces morbidity and mortality in people with chronic heart failure. The mechanism of this benefit remains unclear, but it is likely that chronic elevations of catecholamines and sympathetic nervous system activity cause progressive myocardial damage, leading to a worsening of left ventricular function and a poorer prognosis in people with heart failure. Large, landmark clinical trials of people with stable NYHA class II and III heart failure have demonstrated significant reductions in the overall mortality rate with treatment with various β-adrenergic receptor blocking agents.[21,22]

Digitalis has been a recognized treatment for heart failure for over 200 years. The various forms of digitalis are called *cardiac glycosides*. They improve cardiac function by increasing the force and strength of ventricular contractions. By decreasing sinoatrial node activity and decreasing conduction through the atrioventricular node, they also slow the heart rate and increase diastolic filling time. Although not a diuretic, digitalis promotes urine output by improving cardiac output and renal blood flow. The role of digitalis in the treatment of heart failure has been studied in clinical trials over the past several decades. The results of these studies remain controversial and mixed; there seems to be a growing consensus that although it does not necessarily reduce mortality rates, digitalis can possibly prevent clinical deterioration and hospitalization.

Vasodilator drugs have not been extensively studied as a lone treatment for the management of heart failure but can be effective in symptom management. Agents such as isosorbide dinitrate and hydralazine may be added to other standard medications for patients with chronic heart failure. Vasodilators such as nitroglycerin, nitroprusside, and nesiritide (B-type NP) are used in AHFS to improve left heart performance by decreasing the preload (through vasodilation) or reducing the afterload (through arteriolar dilation), or both.[23,24]

Oxygen Therapy. Oxygen therapy increases the oxygen content of the blood and is most often used in people with acute episodes of heart failure. Continuous positive airway pressure (CPAP) is recommended to reduce the need for endotracheal intubation in patients with AHFS.[25] Because CPAP increases intrathoracic pressure, it also has the potential for decreasing venous return and left ventricular preload, thereby improving the cardiac ejection fraction and stabilizing the hemodynamic status in persons with severe heart failure. Bilevel positive airway pressure (BiPAP), which is like CPAP but also delivers higher pressures during inspiration, is argued by some to be superior to CPAP in that it decreases the respiratory rate and heart rate and improves oxygenation more quickly or more substantially than CPAP.[25]

Cardiac Resynchronization and Implantable Cardioverter–Defibrillators. Some people with heart failure have abnormal intraventricular conduction that results in dyssynchronous and ineffective contractions.[26] Cardiac resynchronization therapy involves the placement of pacing leads into the right and left ventricles as a means of resynchronizing the contraction of the two ventricles. Cardiac resynchronization has been shown to improve ventricular function and blood pressure, improve quality of life, and reduce the risk of death.[23]

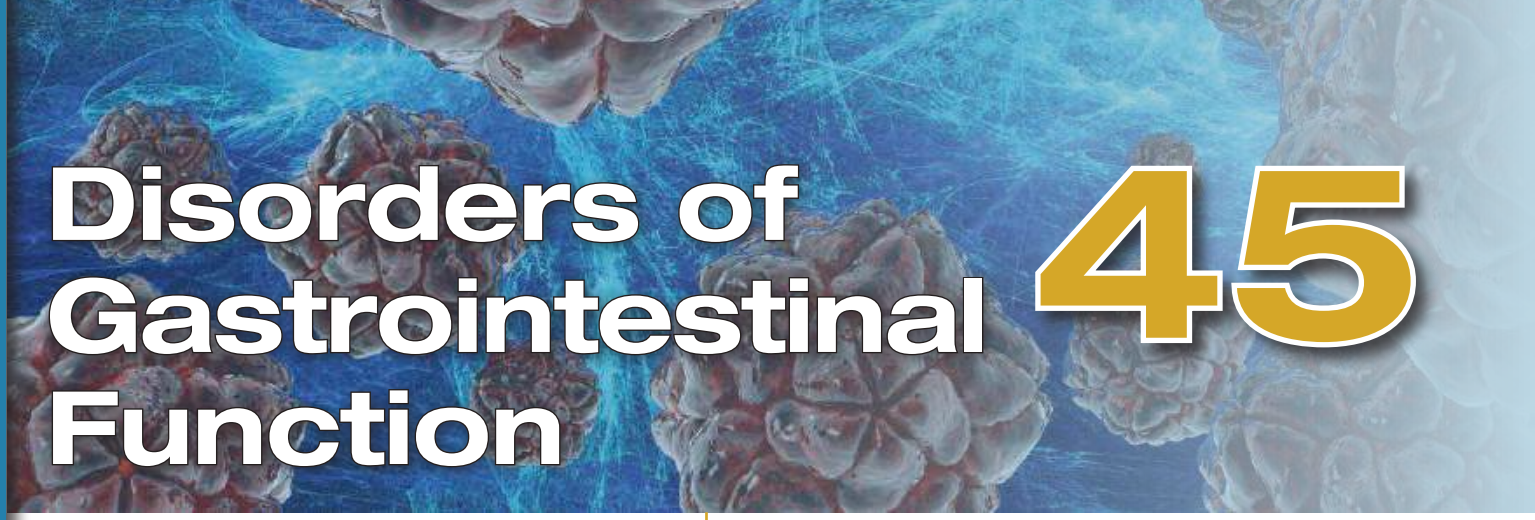
People with heart failure are at significant risk of sudden cardiac death from ventricular fibrillation or ventricular tachycardia. Implantation of a cardioverter–defibrillator is indicated in selected patients with heart failure to prevent sudden cardiac death.[23] A cardioverter–defibrillator is a programmable implanted device that monitors the cardiac rhythm. It has the capacity to pace the heart and deliver electrical shocks to terminate lethal arrhythmias when needed.

Mechanical Support and Heart Transplantation. Refractory heart failure reflects deterioration in cardiac function that is unresponsive to medical or surgical interventions. With improved methods of treatment, more people are reaching a point where a cure is unachievable and death is imminent without mechanical support or heart transplantation.

Since the early 1960s, significant progress has been made in improving the efficacy of *VADs*, which are mechanical pumps used to support ventricular function. VADs are used to decrease the workload of the myocardium while maintaining cardiac output and systemic arterial pressure. This decreases the workload on the ventricle and allows it to rest and recover. In the past, VADs require an invasive open chest procedure for implantation but is not less invasive. They may be used in people who fail or have difficulty being weaned from cardiopulmonary bypass after cardiac surgery, those who develop cardiogenic shock after myocardial infarction, those with end-stage cardiomyopathy, and those who are awaiting cardiac transplantation. Earlier and more aggressive use of VADs as a bridge to transplantation and destination therapy (permanent support) has been shown to increase survival.[23] VADs that allow the patient to be mobile and managed at home are sometimes used for long-term or permanent support for treatment of end-stage heart failure, rather than simply as a bridge to transplantation. VADs can be used to support the function of the left ventricle, right ventricle, or both.[23]

Heart transplantation is the preferred treatment for people with end-stage cardiac failure and otherwise good life expectancy.[2,27] Despite the overall success of heart transplantation, donor availability remains a key problem, and only about 5000 procedures are completed each year with thousands being denied transplantation each year.

Other novel surgical therapies that are being explored include left ventricular remodeling. Left ventricular remodeling is a surgical procedure designed to restore the size and shape of the ventricle and is thought to be a viable surgical alternative to cardiac transplantation for people with severe left ventricular dysfunction.[28]



45 Disorders of Gastrointestinal Function

Zachary Krom

COMMON MANIFESTATIONS OF GI DISORDERS
ANOREXIA, NAUSEA, AND VOMITING
- Anorexia
- Nausea
- Retching and Vomiting

DISORDERS OF THE ESOPHAGUS
- Congenital Anomalies
- Dysphagia
- Esophageal Diverticulum
- Tears (Mallory-Weiss Syndrome)
- Hiatal Hernia
- Gastroesophageal Reflux
  - *Gastroesophageal Reflux Disease*
  - *Gastroesophageal Reflux in Children*
- Cancer of the Esophagus

DISORDERS OF THE STOMACH
- Gastric Mucosal Barrier
- Gastritis
  - *Acute Gastritis*
  - *Chronic Gastritis*
- Peptic Ulcer Disease
  - *Peptic Ulcers*
  - *Zollinger-Ellison Syndrome*
  - *Stress Ulcers*
- Cancer of the Stomach
  - *Etiology and Pathogenesis*
  - *Clinical Manifestations*
  - *Diagnosis and Treatment*

DISORDERS OF THE SMALL AND LARGE INTESTINES
- Irritable Bowel Syndrome
  - *Clinical Manifestations and Diagnosis*
  - *Treatment*
- Inflammatory Bowel Disease
  - *Etiology and Pathogenesis*
  - *Clinical Manifestations*
  - *Crohn Disease*
  - *Ulcerative Colitis*
- Infectious Enterocolitis
  - *Viral Infection*
  - *Bacterial Infection*
  - *Protozoan Infection*
- Diverticular Disease
- Appendicitis
- Alterations in Intestinal Motility
  - *Diarrhea*
  - *Constipation*
  - *Fecal Impaction*
  - *Intestinal Obstruction*
  - *Peritonitis*
- Alterations in Intestinal Absorption
  - *Malabsorption Syndrome*
  - *Celiac Disease*
- Neoplasms
  - *Adenomatous Polyps*
  - *Colorectal Cancer*

Gastrointestinal (GI) disorders do not receive the same publicity in the health-related media as heart disease, cancer, and cerebrovascular disease. In 2009, GI disease was not among the top 15 causes of disease in the United States.[1] However, GI disease has a profound effect on those who suffer from it. According to government reports, digestive diseases rank third in the total economic burden of illness, resulting in considerable human suffering, personal expenditures for treatment, and lost working hours, as well as a drain on the nation's economy. It has been estimated that 60 to 70 million people in the United States have a digestive disease at a cost of 100 billion dollars in medical services.[2] Even more important is the fact that proper nutrition or a change in health practices could prevent or minimize many of these disorders.

Disruption in structure and function can occur at any level of the GI tract, from the esophagus to the colon and rectum. This chapter is divided into four sections:

1. Common manifestations of GI disorders
2. Disorders of the esophagus
3. Disorders of the stomach
4. Disorders of the small and large intestines

Disorders of the hepatobiliary system and exocrine pancreas are discussed in Chapter 46.

COMMON MANIFESTATIONS OF GI DISORDERS ANOREXIA, NAUSEA, AND VOMITING

After completing this section of the chapter, you should be able to meet the following objectives:

- Characterize the relationship among anorexia, nausea, retching, and vomiting.
- Describe the neural structures involved in vomiting and their mediators.

Anorexia, nausea, and vomiting are physiologic responses that are common to many GI disorders. These responses are protective to the extent that they signal the presence of disease and, in the case of vomiting, remove noxious agents from the GI tract. However, they also can contribute to impaired intake or loss of fluids and nutrients.

Anorexia

Anorexia represents a loss of appetite. Several factors influence appetite. One is hunger, which is stimulated by contractions of the empty stomach. The hypothalamus and other associated centers in the brain regulate appetite or the desire for food intake. Smell plays an important role, as evidenced by the fact that appetite can be stimulated or suppressed by the smell of food. Loss of appetite is associated with emotional factors, such as fear, depression, frustration, and anxiety. Many drugs and disease states cause anorexia. For example, in uremia, the accumulation of nitrogenous wastes in the blood contributes to the development of anorexia. Anorexia often is a forerunner of nausea, and most conditions that cause nausea and vomiting also produce anorexia.

Nausea

Nausea is an ill-defined and unpleasant subjective sensation. It is the conscious sensation resulting from stimulation of the medullary vomiting center that often precedes or accompanies vomiting. Nausea usually is preceded by anorexia, and stimuli such as foods and drugs that cause anorexia in small doses usually produce nausea when given in larger doses. A common cause of nausea is distention of the duodenum or upper small intestinal tract. Nausea frequently is accompanied by autonomic nervous system (ANS) manifestations such as watery salivation and vasoconstriction with pallor, sweating, and tachycardia. Nausea also may function as an early warning signal of a pathologic process.

Remember Ms. Rytel who you met at the beginning of this unit? Ms. Rytel has a 36-hour history of vomiting. This is probably due to some adhesions from her multiple surgeries, which may be causing some obstruction. She has had nausea and experienced tachycardia and tachypnea and her color is very pale, which are all ANS manifestations.

Retching and Vomiting

Retching consists of the rhythmic spasmodic movements of the diaphragm, chest wall, and abdominal muscles. It usually precedes or alternates with periods of vomiting. Vomiting or emesis is the sudden and forceful oral expulsion of the contents of the stomach. It usually is preceded by nausea. The contents that are vomited are called *vomitus*. Vomiting, as a basic physiologic protective mechanism, limits the possibility of damage from ingested noxious agents by emptying the contents of the stomach and portions of the small intestine. Nausea and vomiting may represent a total-body response to drug therapy, including overdose, cumulative effects, toxicity, and side effects.

Vomiting involves two functionally distinct medullary centers—the *vomiting center* and the *chemoreceptor trigger zone*.[3] The act of vomiting is thought to be a reflex that is integrated in the vomiting center, which is located in the dorsal portion of the reticular formation of the medulla near the sensory nuclei of the vagus (Fig. 45.1). The chemoreceptor trigger zone is located in a small area on the floor of the fourth ventricle, where it is exposed to both blood and cerebrospinal fluid. It is thought to mediate the emetic effects of bloodborne drugs and toxins.

The act of vomiting consists of taking a deep breath, closing the airways, and producing a strong, forceful contraction of the diaphragm and abdominal muscles along with relaxation of the gastroesophageal sphincter. Respiration ceases during the act of vomiting. Vomiting may be accompanied by dizziness, light-headedness, a decrease in blood pressure, and bradycardia.

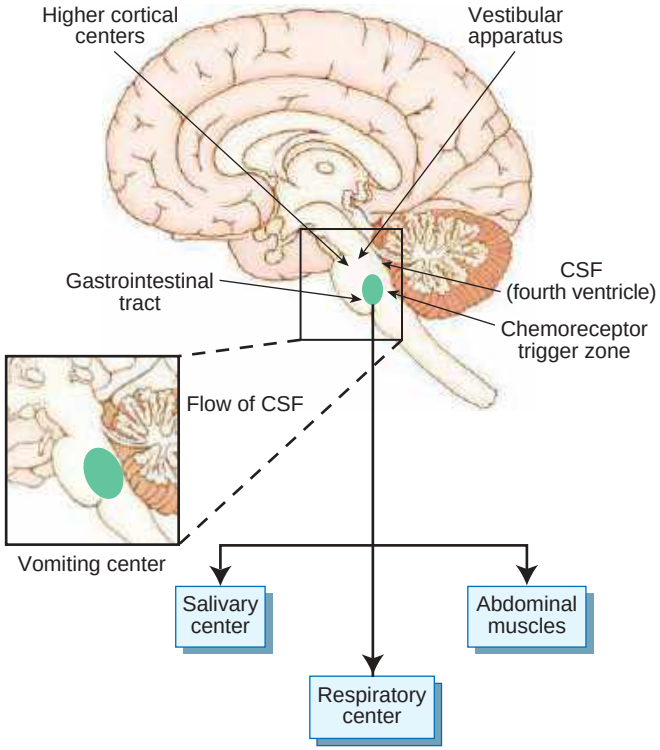


FIGURE 45.1 • Physiologic events involved in vomiting. CSF, cerebrospinal fluid.

The vomiting center receives input from the GI tract and other organs; from the cerebral cortex; from the vestibular apparatus, which is responsible for motion sickness; and from the chemoreceptor trigger zone, which is activated by many drugs and endogenous and exogenous toxins (see Fig. 45.1). Hypoxia exerts a direct effect on the vomiting center, producing nausea and vomiting. This direct effect probably accounts for the vomiting that occurs during periods of decreased cardiac output, shock, environmental hypoxia, and brain ischemia caused by increased intracranial pressure. Inflammation of any of the intra-abdominal organs, including the liver, gallbladder, or urinary tract, can cause vomiting because of the stimulation of the visceral afferent pathways that communicate with the vomiting center. Distention or irritation of the GI tract also causes vomiting through the stimulation of visceral afferent neurons.

Refer back to Ms. Rytel. Ms. Rytel has been vomiting for 36 hours. This is most likely due to distention and irritation of the bowel caused by adhesions, which resulted from multiple surgeries. The irritation is causing persistent stimulation of the visceral afferent pathways, which are linked with the vomiting center.

Several neurotransmitters and receptor subtypes are implicated as neuromediators in nausea and vomiting. Dopamine, serotonin, and opioid receptors are found in the GI tract and in the vomiting center and chemoreceptor trigger zone. Dopamine antagonists, such as prochlorperazine, depress vomiting caused by stimulation of the chemoreceptor trigger zone. Serotonin is believed to be involved in the nausea and emesis associated with cancer chemotherapy and radiation therapy. Serotonin antagonists (*e.g.,* granisetron, ondansetron) are effective in treating the nausea and vomiting associated with these stimuli. Motion sickness appears to be a central nervous system (CNS) response to vestibular stimuli. Norepinephrine and acetylcholine receptors are located in the vestibular center. The acetylcholine receptors are thought to mediate the impulses responsible for exciting the vomiting center. Norepinephrine receptors may have a stabilizing influence that resists motion sickness. Many of the motion sickness drugs (*e.g.,* dimenhydrinate) have a strong CNS anticholinergic effect and act on the receptors in the vomiting center and areas related to the vestibular system.

IN SUMMARY

The signs and symptoms of many GI tract disorders are manifested by anorexia, nausea, and vomiting. Anorexia, or loss of appetite, may occur alone or may accompany nausea and vomiting. Nausea, which is an ill-defined, unpleasant sensation, signals the stimulation of the medullary vomiting center. It often precedes vomiting and frequently is accompanied by autonomic responses, such as salivation and vasoconstriction with pallor, sweating, and tachycardia. The act of vomiting, which is integrated by the vomiting center, involves the forceful oral expulsion of the gastric contents. It is a basic physiologic mechanism that rids the GI tract of noxious agents.

DISORDERS OF THE ESOPHAGUS

After completing this section of the chapter, you should be able to meet the following objectives:

- Define and cite the causes of dysphagia, odynophagia, and achalasia.
- Relate the pathophysiology of gastroesophageal reflux to measures used in the diagnosis and treatment of the disorder in adults and children.
- State the reason for the poor prognosis associated with esophageal cancer.

The esophagus is a tube that connects the oropharynx with the stomach. It lies posterior to the trachea and larynx and extends through the mediastinum, intersecting the diaphragm at the level of the 11th thoracic vertebra.

The esophagus functions primarily as a conduit for passage of food and liquid from the pharynx to the stomach. The walls of the esophagus consist of a mucosal, submucosal, muscularis externa, and adventitial layer, reflecting the general structural organization of the GI tract. The inner mucosal layer contains nonkeratinized stratified epithelium. At the esophageal–stomach junction, the abrasion-resistant epithelium changes abruptly to the simple columnar epithelium of the stomach. The submucosal layer contains mucus-secreting glands that provide the mucin-containing fluids that lubricate the esophageal wall and aid in the passage of food. The muscularis externa layer consists of skeletal muscle in the superior third of the esophagus, a mixture of skeletal and smooth muscle in its middle third, and entirely smooth muscle in its lower third. The outer fibrous adventitial layer of the esophagus is composed entirely of connective tissue, which blends with surrounding structures along its route.

There are sphincters at either end of the esophagus: an upper esophageal sphincter and a lower esophageal sphincter. The upper esophageal, or pharyngoesophageal, sphincter consists of a circular layer of striated muscle, the cricopharyngeal muscle. The lower esophageal, or gastroesophageal, sphincter is an area approximately 3 cm above the junction with the stomach. The gastroesophageal sphincter is a physiologic rather than a true anatomic sphincter. That is, it acts as a valve, but the only structural evidence of a sphincter is a slight thickening of the circular smooth muscle. The smooth muscle in this portion of the esophagus normally remains tonically constricted, creating an intraluminal pressure of about 30 mm Hg, in contrast to the mid-portion of the esophagus, which normally remains relaxed.[4] The lower esophageal sphincter

passes through an opening, or *hiatus*, in the diaphragm as it joins with the stomach, which is located in the abdomen. The portion of the diaphragm that surrounds the lower esophageal sphincter helps to maintain the zone of high pressure needed to prevent reflux of stomach contents.

Congenital Anomalies

Congenital anomalies of the esophagus require early detection and correction because they are incompatible with life. Esophageal atresia (EA) and tracheoesophageal fistula (TEF) are very common congenital anomalies of the esophagus, affecting approximately 1 in 45,000 neonates.[5] In the most common form of EA, representing 85% of cases,[6] the upper esophagus ends in a blind pouch and the TEF is connected to the trachea (Fig. 45.2). This defect now has a survival rate greater than 90% owing largely to early recognition and improved neonatal intensive care units. Infants weighing less than 1500 g have the greatest risk for mortality, especially when combined with a cardiac anomaly.[6]

The newborn infant with EA/TEF typically has frothing and bubbling at the mouth and nose and episodes of coughing, vomiting, cyanosis, and respiratory distress. Feeding exacerbates these manifestations, causes regurgitation, and precipitates aspiration. The inability to pass a catheter into the stomach provides further evidence of the disorder. The infant with isolated TEF may develop respiratory symptoms at a later age.

Treatment of EA and TEF is surgical. Surgical ligation of the TEF and end-to-end anastomosis of the esophagus is performed when possible. Temporary ligation of the TEF and insertion of a gastrostomy tube may be used to delay the need for primary closure in preterm infants and those with more complicated lesions. The main goal of preoperative management is to maintain the airway and prevent lung damage from aspiration of gastric contents. Prone positioning minimizes movement of gastric secretions into a distal fistula, and esophageal suctioning minimizes the risk of aspiration from a blind pouch.

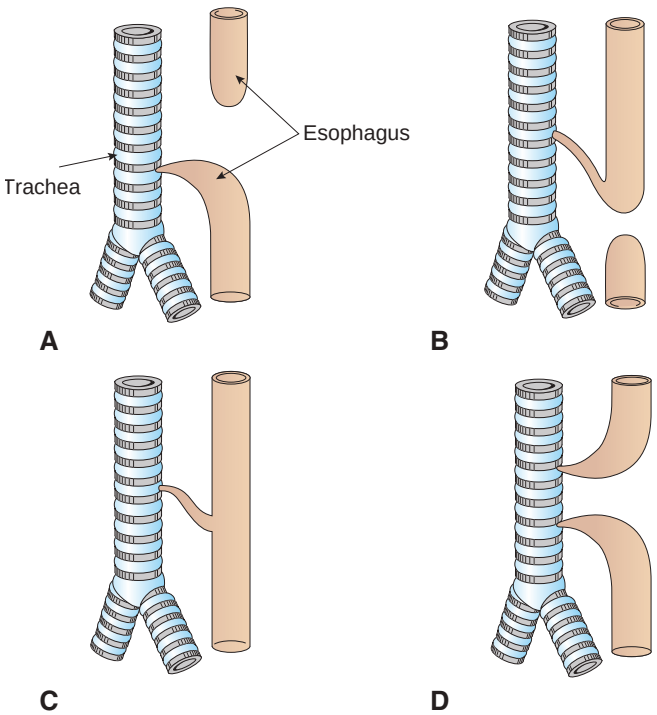


FIGURE 45.2 • Congenital tracheoesophageal fistulas. (**A**) The most common type is a communication between the trachea and the lower portion of the esophagus. The upper segment of the esophagus ends in a blind sac. (**B**) In a few cases, the proximal esophagus communicates with the trachea. (**C**) H type fistula without esophageal atresia, and (**D**) tracheal fistulas to both a proximal esophageal pouch and distal esophagus. (From Rubin R., Strayer D. S. (Eds.) (2012). *Rubin's pathophysiology: Clinicopathologic foundations of medicine* (6th ed., p. 607). Philadelphia, PA: Lippincott Williams & Wilkins.)

Dysphagia

The act of swallowing depends on the coordinated action of the tongue and pharynx. These structures are innervated by cranial nerves V, IX, X, and XII. **Dysphagia** refers to difficulty in swallowing. If swallowing is painful, it is referred to as *odynophagia*. Dysphagia can result from neuromuscular or structural causes. These disorders can produce narrowing of the esophagus, lack of salivary secretion, weakness of the muscular structures that propel the food bolus toward the stomach, or disruption of the neural networks coordinating the swallowing mechanism.[7] An example of a neuromuscular cause involves lesions of the CNS, such as a stroke, which often involve the cranial nerves that control swallowing. Cancer of the esophagus and strictures resulting from scarring, a structural cause, can reduce the size of the esophageal lumen and make swallowing difficult. Scleroderma, an autoimmune disease that causes fibrous replacement of tissues throughout the body and in the GI tract, is another important cause of dysphagia.[8] People with dysphagia usually complain of choking, coughing, or an abnormal sensation of food sticking in the back of the throat or upper chest when they swallow.

In a condition called *achalasia*, the lower esophageal sphincter fails to relax due to a disruption in the input from the enteric neural plexus and the vagus nerve.[9] This results in difficulty passing food into the stomach, and the esophagus above the lower esophageal sphincter becomes enlarged. One or several meals may lodge in the esophagus and pass slowly into the stomach over time. There is danger of aspiration of esophageal contents into the lungs when the person lies down.

Endoscopy, barium esophagoscopy, and videoradiography may be used to determine the site and extent of a swallowing disorder. Esophageal manometry, a procedure in which a small pressure-sensing catheter is inserted into the esophagus, may be done to measure pressures in different parts of the esophagus. Treatment of swallowing disorders depends on the cause and type of altered function that is present. Treatment of dysphagia often involves a multidisciplinary team of health professionals, including a speech pathologist. Mechanical dilation or surgical procedures may be done to enlarge the lower esophageal sphincter in persons with esophageal strictures.

Esophageal Diverticulum

A diverticulum of the esophagus is a herniation of the esophageal wall caused by a weakness of the muscularis layer.[10] An esophageal diverticulum tends to retain food. Complaints that the food stops before it reaches the stomach are common, as are reports of gurgling, belching, coughing, and foul-smelling breath. The trapped food may cause esophagitis and ulceration. Because the condition usually is progressive, correction of the defect requires surgical intervention.

Tears (Mallory-Weiss Syndrome)

Longitudinal tears in the esophagus at the esophagogastric junction that often extend distally are termed *Mallory-Weiss tears*.[11] They are most often encountered in persons with chronic alcoholism after a bout of severe retching or vomiting but may also occur during acute illness with severe vomiting. The presumed pathogenesis is inadequate relaxation of the esophageal sphincter during vomiting, with stretching and tearing of the esophageal junction at the moment of propulsive expulsion of gastric contents. Tears may involve only the mucosa or may penetrate the wall of the esophagus. Infection may lead to inflammatory ulcer or mediastinitis.

Most often bleeding is not severe and does not require surgical intervention. Severe bleeding usually responds to vasoconstrictive medications, transfusions, and balloon compression. Healing is usually prompt, with minimal or no residual effects.

Hiatal Hernia

Hiatal hernia is characterized by a protrusion or herniation of the stomach through the esophageal hiatus of the diaphragm. There are two anatomic patterns of hiatal herniation: axial, or sliding, and nonaxial, or paraesophageal.[12] The sliding hiatal hernia is characterized by a bell-shaped protrusion of the stomach above the diaphragm (Fig. 45.3). Small sliding hiatal hernias are common and considered to be of no significance in asymptomatic people. However, in cases of severe erosive esophagitis where gastroesophageal reflux and a large hiatal hernia coexist, the hernia may retard esophageal acid clearance and contribute to the more severe esophagitis, especially Barrett esophagus (to be discussed). In paraesophageal hiatal hernias, a separate portion of the stomach, usually along the greater portion of the stomach, enters the thorax through a widened opening and then progressively enlarges. In extreme cases, most of the stomach herniates into the thorax. Large paraesophageal hiatal hernias may require surgical treatment.

Gastroesophageal Reflux

The term reflux refers to backward or return movement. In the context of gastroesophageal reflux, it refers to the backward movement of gastric contents into the esophagus, a condition

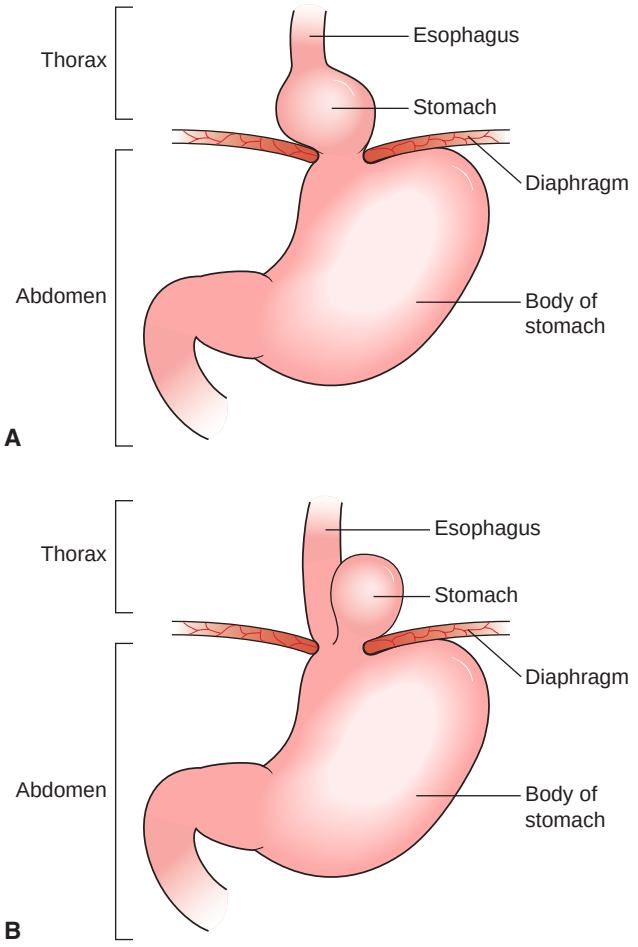


FIGURE 45.3 • Hiatal hernia. (**A**) Sliding hiatal hernia. (**B**) Paraesophageal hiatal hernia.

that causes heartburn or pyrosis. It probably is the most common disorder originating in the GI tract. The associated symptoms usually occur soon after eating, are short lived, and seldom cause more serious problems.

The lower esophageal sphincter regulates the flow of food from the esophagus into the stomach. Both intrinsic and extrinsic mechanisms function in maintaining the antireflux function of the lower esophageal sphincter.[13] The circular muscles of the distal esophagus constitute the intrinsic mechanisms, and the portion of the diaphragm that surrounds the esophagus constitutes the extrinsic mechanism. The oblique muscles of the stomach, located below the lower esophageal sphincter, form a flap that contributes to the antireflux function of the internal sphincter. Relaxation of the lower esophageal sphincter is a brain stem reflex that is mediated by the vagus nerve in response to a number of afferent stimuli. Transient relaxation with reflux is common after meals. Gastric distention and meals high in fat increase the frequency of relaxation. Normally, refluxed material is returned to the stomach by secondary peristaltic waves in the esophagus, and swallowed saliva neutralizes and washes away the refluxed acid.

Gastroesophageal Reflux Disease

According to the Montreal definition, gastroesophageal reflux disease (GERD) is a disorder involving the reflux of stomach contents that causes unfavorable symptoms or complications for the person such as regurgitation and heartburn.[14] It is thought to be associated with transient relaxations of weak or incompetent lower esophageal sphincter. This allows reflux to occur and, in addition, decreased clearance of the refluxed acid from the esophagus after it has occurred. It results in irritant effects of the refluxate.[15] In most cases, reflux occurs during transient relaxation of the esophagus. Delayed gastric emptying also may contribute to reflux by increasing gastric volume and pressure with greater chance for reflux. Esophageal mucosal injury is related to the destructive nature of the refluxate and the amount of time it is in contact with mucosa. Acidic gastric fluids (pH < 4.0) are particularly damaging. The gastroesophageal reflux normally is cleared and neutralized by esophageal peristalsis and salivary bicarbonate. Decreased salivation and salivary buffering capacity may contribute to impaired clearing of acid reflux from the esophagus.

Clinical Manifestations. The most frequent symptom of GERD is heartburn. It frequently is severe, occurring 30 to 60 minutes after eating. It often is made worse by bending at the waist and recumbency and usually is relieved by sitting upright. The severity of heartburn is not indicative of the extent of mucosal injury. Only a small percentage of people who complain of heartburn have mucosal injury. Often, the heartburn occurs during the night. Antacids give prompt, although transient, relief. Other symptoms include belching and chest pain. The pain usually is located in the epigastric or retrosternal area and often radiates to the throat, shoulder, or back. Because of its location, the pain may be confused with angina. The reflux of gastric contents also may produce respiratory symptoms such as asthma, chronic cough, and laryngitis, but it is important to note that the presence of these symptoms is often multifactorial in addition to the diagnosis or GERD.[16] The proposed mechanisms of reflux-associated asthma and chronic cough include microaspiration and macroaspiration, laryngeal injury, and vagal-mediated bronchospasm.

Reflux esophagitis involves mucosal injury to the esophagus, hyperemia, and inflammation. Complications, such as strictures and *Barrett esophagus*, can result from persistent reflux, which produces a cycle of mucosal damage that causes hyperemia, edema, and erosion of the luminal surface. Strictures are caused by a combination of scar tissue, spasm, and edema. They produce narrowing of the esophagus and cause dysphagia when the lumen becomes sufficiently constricted. Barrett esophagus (Fig. 45.4) is characterized by a reparative process in which the squamous mucosa that normally lines the esophagus gradually is replaced by abnormal columnar epithelium resembling that in the stomach or intestines.[17] It is associated with increased risk for development of esophageal adenocarcinoma.

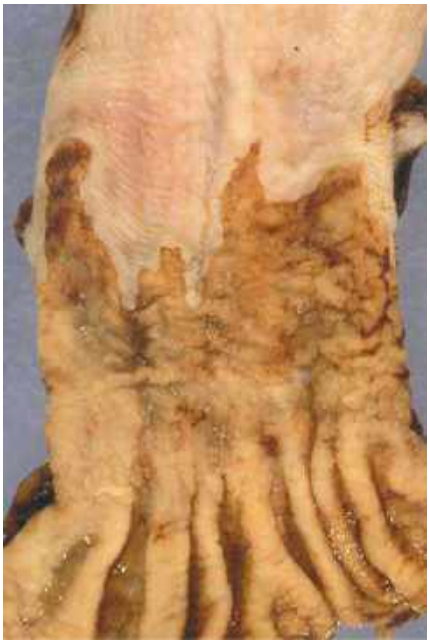


FIGURE 45.4 • The presence of the tan tongues of epithelium interdigitating with the more proximal squamous epithelium is typical of Barrett esophagus. (From Rubin R., Strayer D. S. (Eds.) (2012). *Rubin's pathophysiology: Clinicopathologic foundations of medicine* (6th ed., p. 611). Philadelphia, PA: Lippincott Williams & Wilkins.)

Diagnosis. Diagnosis of gastroesophageal reflux depends primarily on a history of reflux symptomatology and the use of optional diagnostic methods, including acid suppression trials, esophagoscopy, and ambulatory esophageal pH monitoring.[15] Acid suppression trials involve administering a proton pump inhibitor medication for 7 to 14 days to determine if the symptoms are alleviated. Esophagoscopy involves the passage of a flexible fiberoptic endoscope into the esophagus for the purpose of visualizing the lumen of the upper GI tract. It also permits performance of a biopsy, if indicated. For 24-hour pH monitoring, a small tube with a pH electrode is passed through the nose and down into the esophagus. Data from the electrode are recorded in a small, lightweight box worn on a belt around the waist and later are analyzed by computer. The device allows the person to indicate position changes, meals, heartburn, or pain, which then can be correlated with episodes of acid reflux.

Treatment. The treatment of gastroesophageal reflux usually focuses on conservative measures. These measures include avoidance of positions and conditions that increase gastric reflux.[15] Avoidance of large meals and foods that reduce lower esophageal sphincter tone (*e.g.,* caffeine, fats, chocolate), alcohol, and smoking is recommended. It is recommended that meals be eaten sitting up and that the recumbent position be avoided for several hours after a meal. Bending for long periods should be avoided because it tends to increase intra-abdominal pressure and cause gastric reflux. Sleeping with the head elevated helps to prevent reflux during the night. This is

best accomplished by placing blocks under the head of the bed or by using a wedge-shaped bolster to elevate the head and shoulders by at least 6 inches. Weight loss usually is recommended in overweight people.

Antacids or a combination of antacids and alginic acid also are recommended for mild disease. Alginic acid produces a foam when it comes in contact with gastric acid; if reflux occurs, the foam rather than acid rises into the esophagus. Histamine-2 receptor (H_2)–blocking antagonists, which inhibit gastric acid production, is another recommended treatment. The proton pump inhibitors act by inhibiting the gastric proton pump, which regulates the final pathway for acid secretion. These agents may be used for people who continue to have daytime symptoms, recurrent strictures, or large esophageal ulcerations. Surgical treatment may be indicated in some people.

Gastroesophageal Reflux in Children

Gastroesophageal reflux is a common problem in infants and children. The small reservoir capacity of an infant's esophagus coupled with frequent spontaneous reductions in sphincter pressure contributes to reflux. At least one episode of regurgitation a day occurs in as much as half of infants aged 0 to 3 months.[18] By 8 months of age, it becomes less frequent, and it abates by 2 years of age[18] as the child's diet naturally advances and they are able to maintain a more upright posture. Although many infants have minor degrees of reflux, complications can occur in children with more frequent or persistent episodes. The condition occurs more frequently in children with cerebral palsy, Down syndrome, cystic fibrosis, and other neurologic disorders.

In most cases, infants with simple reflux are thriving and healthy, and symptoms resolve between 9 and 24 months of age. Pathologic reflux is classified into three categories:

1. Regurgitation and malnutrition
2. Esophagitis
3. Respiratory problems

Clinical Manifestations. Symptoms of reflux esophagitis include evidence of pain when swallowing, hematemesis, anemia due to esophageal bleeding, heartburn, irritability, and sudden or inconsolable crying. Children with gastroesophageal reflux often express feeding difficulties such as refusal and aversion to certain food textures.[19] Tilting of the head to one side and arching of the back may be noted in children with severe reflux. The head positioning is thought to represent an attempt to protect the airway or reduce the pain-associated reflux. Sometimes regurgitation is associated with dental caries and recurrent otalgia. The ear pain is thought to occur through referral from the vagus nerve in the esophagus to the ear.

A variety of respiratory symptoms are caused by damage to the respiratory mucosa when gastric reflux enters the esophagus. Reflux may cause laryngospasm, apnea, and bradycardia. Asthma may co-occur with GERD in about 50% of asthmatic children.[18] Asthmatic children who are particularly likely to have GERD as a provocative factor are those with symptoms of reflux, those with refractory or steroid-dependent asthma, and those with nocturnal worsening of symptoms.[18]

Diagnosis and Treatment. Diagnosis of gastroesophageal reflux in infants and children often is based on parental and clinical observations. The diagnosis may be confirmed by esophageal pH probe studies, barium fluoroscopic esophagography, and nuclear scintigraphy. In some cases, esophagoscopy may be used to demonstrate reflux and obtain a biopsy.

Various treatment methods are available for infants and children with gastroesophageal reflux. Small, frequent feedings are recommended because of the association between gastric volume and transient relaxation of the esophagus. Thickening an infant's feedings has not been found to decrease the amount of regurgitation occurrences but decrease the volume of reflux.[20] Prone positioning may decrease the likelihood of symptoms, but it has also been associated with increasing the risk for sudden infant death syndrome.[20] In older infants and children, raising the head of the bed and keeping the child upright may help. Medications usually are not added to the treatment regimen until pathologic reflux has been documented by diagnostic testing. Antacids are the most commonly used antireflux therapy and are readily available over the counter. H_2-receptor antagonists and proton pump inhibitors may be used in children with persistent reflux. Prokinetic agents (*e.g.,* metoclopramide, a dopamine-2 and 5-hydroxytryptamine [$5\text{-}HT_3$] receptor antagonist; bethanechol, a cholinergic agonist) are associated with significant side effects, and their use in treatment is not recommended.[20]

Cancer of the Esophagus

Carcinoma of the esophagus accounts for approximately 1% of all diagnosed cancers.[21] It is more common in adults over the age of 65 years. It occurs three times more frequently in men than women, and its occurrence is equal between African Americans and whites.[21]

There are two types of esophageal cancer: squamous cell carcinoma and adenocarcinoma. Most squamous cell esophageal carcinomas are attributable to alcohol and tobacco use. Worldwide, squamous cell carcinomas are the most common type of esophageal cancers, but in the United States, there has been a significant increase in adenocarcinomas.[22] Barrett esophagus and GERD are the two most common risk factors for esophageal adenocarcinoma.[23]

Progressive dysphagia is by far the most frequent complaint of people with esophageal cancer. It is apparent first with ingestion of bulky food, later with soft food, and finally with liquids. Unfortunately, it is a late manifestation of the disease. Unintentional weight loss, anorexia, fatigue, and pain on swallowing also may occur.

Treatment of esophageal cancer depends on tumor stage. Surgical resection provides a means of cure when done in early disease and palliation when done in late disease. Radiation may be used as an alternative to surgery. Chemotherapy may be used before surgery to decrease the size of the tumor, or it

may be used along with irradiation and surgery in an effort to increase survival.[24]

The prognosis for people with cancer of the esophagus, although poor, has improved. Even with modern forms of therapy, however, the long-term survival is limited because, in many cases, the disease has already metastasized by the time the diagnosis is made.

IN SUMMARY

The esophagus is a tube that connects the oropharynx with the stomach; it functions primarily as a conduit for passage of food from the pharynx to the stomach. Although relatively uncommon, congenital anomalies (*i.e.,* EA and TEFs) must be corrected early because they cause aspiration of gastric and oral secretions and are incompatible with life. Dysphagia refers to difficulty in swallowing; it can result from altered nerve function or from disorders that produce narrowing of the esophagus. A diverticulum of the esophagus is an outpouching of the esophageal wall caused by a weakness of the muscularis layer. Longitudinal tears (Mallory-Weiss tears) at the esophagogastric junction can occur with severe bouts of retching or vomiting. They are most often encountered in people with chronic alcoholism, but may also occur during acute illness with severe vomiting. Hiatal hernia is characterized by a protrusion or herniation of the stomach through the esophageal hiatus of the diaphragm. There are two anatomic patterns of herniation: (1) the axial or sliding hiatal hernia, which is the most common type and is characterized by bell-shaped protrusion of the stomach above the diaphragm and (2) the nonaxial or paraesophageal hernia, in which a portion of the stomach enters the thorax through a widened opening.

Gastroesophageal reflux refers to the backward movement of gastric contents into the esophagus, a condition that causes heartburn. Although most persons experience occasional gastroesophageal reflux and heartburn, persistent reflux can result in a cycle of mucosal damage that causes hyperemia, edema, erosion luminal surface, and Barrett esophagus. Reflux can cause respiratory symptoms, including chronic cough, and serve as a potential trigger for asthma. Gastroesophageal reflux is a common problem in infants and children. Reflux commonly corrects itself with age, and symptoms abate in most children by 2 years of age. Although many infants have minor degrees of reflux, some infants and small children have significant reflux that interferes with feeding, causes esophagitis, and results in respiratory symptoms and other complications.

Carcinoma of the esophagus is more common in older adults and occurs more frequently in men than women. There are two types of esophageal cancer: squamous cell carcinoma and adenocarcinoma. Most squamous cell carcinomas are attributable to alcohol and tobacco use. Adenocarcinomas are more closely linked to gastroesophageal reflux and Barrett esophagus.

DISORDERS OF THE STOMACH

After completing this section of the chapter, you should be able to meet the following objectives:

- Differentiate between the causes and manifestations of acute and chronic gastritis.
- Characterize the proposed role of *Helicobacter pylori* in the development of chronic gastritis and peptic ulcer and cite methods for diagnosis and treatment of the infection.
- Cite the etiologic factors in ulcer formation related to Zollinger-Ellison syndrome and stress ulcer.

The stomach is a reservoir for contents entering the digestive tract. It lies in the upper abdomen, anterior to the pancreas, splenic vessels, and left kidney. Anteriorly, the stomach is bounded by the anterior abdominal wall and the left inferior lobe of the liver. While in the stomach, food is churned and mixed with hydrochloric acid and pepsin before being released into the small intestine. Normally, the mucosal surface of the stomach provides a barrier that protects it from the hydrochloric acid and pepsin contained in gastric secretions. Disorders of the stomach include gastritis, peptic ulcer, and gastric carcinoma.

Gastric Mucosal Barrier

The stomach lining usually is impermeable to the acid it secretes, a property that allows the stomach to contain acid and pepsin without having its walls digested. Several factors contribute to the protection of the gastric mucosa, including an exceptionally tight fitting and therefore impermeable epithelial cell surface covering. This is coupled with the tenacious, thick mucous that is secreted by cells, which creates a protective covering for the inner stomach wall that also contains bicarbonates that serve to maintain a neutral pH.[4,25] These mechanisms are collectively referred to as the *gastric mucosal barrier*.

The cells of the gastric epithelia are connected by tight junctions that prevent acid penetration, and they are covered with an impermeable hydrophobic lipid layer that prevents diffusion of ionized water-soluble molecules. Aspirin is able to cross the lipid layer and cause damage to the superficial cells, which can result in acute erosions.[26] Gastric irritation and occult bleeding due to gastric irritation occur in a significant number of persons who take aspirin on a regular basis. Alcohol, which like aspirin is lipid soluble, also disrupts the mucosal barrier. When aspirin and alcohol are taken in combination, the permeability of the gastric mucosal barrier is significantly increased and cellular damage occurs.[27] Bile acids also attack the lipid components of the mucosal barrier and afford the potential for gastric irritation when there is reflux of duodenal contents into the stomach.

Normally, the secretion of hydrochloric acid by the parietal cells of the stomach is accompanied by secretion of bicarbonate ions (HCO_3^-). For every hydrogen ion (H^+) that is secreted, an HCO_3^- is produced, and as long as HCO_3^- production is equal to H^+ secretion, mucosal injury does not occur. Changes in gastric blood flow, as in shock, tend to decrease HCO_3^- production. This is particularly true in situations in which decreased blood flow is accompanied by acidosis. Aspirin and the nonsteroidal anti-inflammatory drugs (NSAIDs), also impair HCO_3^- secretion by way of inhibiting gastric COX-1, a fatty acid enzyme that synthesizes prostaglandins that mediate bicarbonate secretion.[28,29]

The mucus that protects the gastric mucosa is of two types: water insoluble and water soluble.[24] Water-insoluble mucus forms a thin, stable gel that adheres to the gastric mucosal surface and provides protection from the proteolytic (protein-digesting) actions of pepsin. It also forms an unstirred layer that traps bicarbonate, forming an alkaline interface between the luminal contents of the stomach and its mucosal surface. The water-soluble mucus is washed from the mucosal surface and mixes with the luminal contents; its viscid nature makes it a lubricant that prevents mechanical damage to the mucosal surface.

Prostaglandins, chemical messengers derived from cell membrane lipids, play an important role in protecting the gastric mucosa from injury.[28] The prostaglandins are thought to exert their effect through improved mucosal blood flow, decreased acid secretion, increased bicarbonate ion secretion, and enhanced mucus production.

KEY POINTS

DISRUPTION OF THE GASTRIC MUCOSA AND ULCER DEVELOPMENT

- The stomach is protected by a mucosal barrier that prevents gastric secretions and other destructive agents from injuring the epithelial and deeper layers of the stomach wall.
- Two of the major causes of gastric irritation and ulcer formation are aspirin or NSAIDs and infection with *H. pylori*.

Gastritis

Gastritis refers to inflammation of the gastric mucosa. There are many causes of gastritis, most of which can be grouped as either acute or chronic gastritis.

Acute Gastritis

Acute gastritis is characterized by an acute mucosal inflammatory process, usually transient in nature. The inflammation may be accompanied by emesis, pain, and, in severe cases, hemorrhage and ulceration.[30] This erosive form is an important cause of acute GI bleeding. The condition is most commonly associated with local irritants such aspirin or other NSAIDs, alcohol, or bacterial toxins. Oral administration of corticosteroids may also be complicated by acute hemorrhagic gastritis. Any serious illness or trauma that is accompanied by profound physiologic stress that requires substantial medical or surgical treatment renders the gastric mucosa more vulnerable to acute hemorrhagic gastritis because of mucosal injury (discussed under stress ulcers).[12] Uremia, treatment with cancer chemotherapy drugs, and gastric radiation are other causes of acute gastritis.

The complaints of people with acute gastritis vary. People with aspirin-related gastritis can be totally unaware of the condition or may complain only of heartburn or sour stomach. Gastritis associated with excessive alcohol consumption is often a different situation; it often causes transient gastric distress, which may lead to vomiting and, in more severe situations, to bleeding and hematemesis. Gastritis caused by the toxins of infectious organisms, such as the staphylococcal enterotoxins, usually has an abrupt and violent onset, with gastric distress and vomiting ensuing approximately 5 hours after the ingestion of a contaminated food source. Acute gastritis usually is a self-limiting disorder, with complete regeneration and healing occurring within several days of removal of the inciting agent.

Chronic Gastritis

Chronic gastritis is a separate entity from acute gastritis. It is characterized by the absence of grossly visible erosions and the presence of chronic inflammatory changes, leading eventually to atrophy of the glandular epithelium of the stomach. There are types of chronic gastritis: *H. pylori* autoimmune and multifocal atrophic gastritis and chemical gastropathy.[12]

***Helicobacter pylori* Gastritis.** *Helicobacter pylori* infection is the most common cause of chronic gastritis. The prevalence in the United States is associated with socioeconomic status, increased age, Hispanic and African-American ethnicity.[31] *Helicobacter pylori* is present in two thirds of the world's population.[31] It has been suggested that transmission in industrialized countries is largely person to person by vomitus, saliva, or feces, whereas additional transmission routes such as water may be important in developing countries. In industrialized countries, the rate of infection with *H. pylori* has decreased substantially over the past several decades owing to improved sanitation.

Helicobacter pylori gastritis is a chronic inflammatory disease of the antrum and body of the stomach. Chronic infection with *H. pylori* can lead to gastric atrophy and peptic ulcer and is associated with increased risk of gastric adenocarcinoma and the creation of mucosa-associated lymphoid tissue, which can progress to lymphoma.[30]

Pathogenesis. *Helicobacter pylori* are small, curved, or spiral shaped, gram-negative rods (protobacteria) that can colonize the mucus-secreting epithelial cells of the stomach.[30,32] (Fig. 45.5). *Helicobacter pylori* have multiple flagella, which

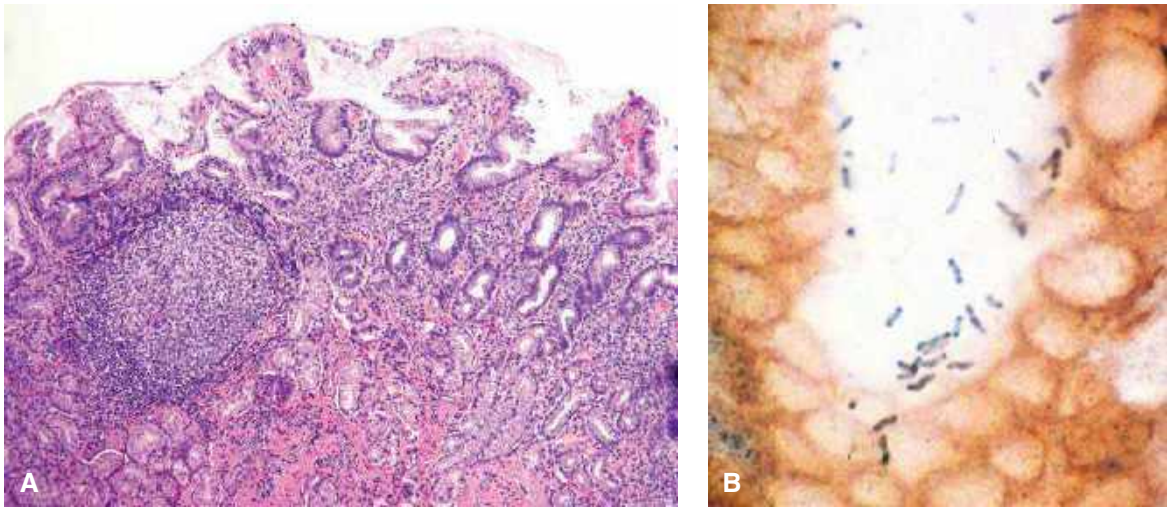


FIGURE 45.5 • *Helicobacter pylori*–associated gastritis. (**A**) The antrum shows an intense lymphocytic and plasma cell infiltrate, which tends to be heaviest in the superficial portions of the lamina propria. (**B**) The microorganisms appear on silver staining as small, curved rods on the surface of the gastric mucosa. (From Rubin R., Strayer D. S. (Eds.) *Rubin's pathophysiology: clinicopathologic foundations of medicine* (6th ed., p. 619). Philadelphia, PA: Lippincott Williams & Wilkins.)

allow them to move through the mucous layer of the stomach, and they secrete urease, which enables them to produce sufficient ammonia to buffer the acidity of their immediate environment. These properties help to explain why the organism is able to survive in the acidic environment of the stomach. *Helicobacter pylori* produce enzymes and toxins that have the capacity to interfere with the local protection of the gastric mucosa against acid, produce intense inflammation, and elicit an immune response. There is increased production of proinflammatory cytokines (IL-6, IL-8) that serve to recruit and activate neutrophils.[33] Several *H. pylori* proteins are immunogenic, and they evoke an intense immune response in the mucosa. Both T and B cells can be seen in the chronic gastritis caused by *H. pylori*. T cells may be responsible for decreasing the constant inflammatory response caused by the cytokines, allowing *H. pylori* to maintain its colonization of the stomach for long periods of time. Although the role of T and B cells in causing epithelial injury has not been established, T-cell–driven activation of B cells may be involved in the pathogenesis of gastric lymphomas.[12]

Why some people with *H. pylori* infection develop clinical disease and others do not is unclear. Scientists are studying the different strains of the bacteria in an attempt to establish whether certain strains are more virulent than others and whether host and environmental factors contribute to the development of clinical disease.[12]

Diagnosis and Treatment. Methods for establishing the presence of *H. pylori* infection include the carbon (C) urea breath test using a radioactive carbon isotope (^{13}C- or ^{14}C-urea), serologic tests, the stool antigen test, and endoscopic biopsy for urease testing.[34] Serologic titers of *H. pylori* antibodies specifically isolate immunoglobulin G and A.

Eradication of *H. pylori* has proved difficult. Treatment requires combination therapy that includes the use of antibiotics such as amoxicillin, tetracycline, aminoglycoside, or bismuth salts in combination with proton pump inhibitors such as lansoprazole and omeprazole.[34] Treatment is usually continued for 10 to 14 days. *Helicobacter pylori* mutate rapidly to develop antibiotic-resistant strains. The combination of two or more antimicrobial agents increases the rates of cure and reduces the risk of resistant strains developing. The proton pump inhibitors have direct antimicrobial properties against *H. pylori*, and by raising the intragastric pH they suppress bacterial growth and optimize antibiotic efficacy. Bismuth has a direct antibacterial effect against *H. pylori*.

Chronic Autoimmune and Multifocal Gastritis. *Autoimmune gastritis*, which accounts for less than 10% of cases of chronic gastritis, is a diffuse form of gastritis that is limited to the body and fundus of the stomach, with a lack or minimal involvement of the antrum.[30] The disorder results from the presence of autoantibodies to components of gastric gland parietal cells and intrinsic factor. Gastric gland and mucosal atrophy lead to a loss of acid production. In the most severe cases, production of intrinsic factor is lost, leading to a vitamin B_{12} deficiency and pernicious anemia. This type of chronic gastritis frequently is associated with other autoimmune disorders such as Hashimoto thyroiditis, Addison disease, and Graves disease.

Multifocal atrophic gastritis is a disorder of uncertain etiology that affects the antrum and adjacent areas of the stomach. It is more common than autoimmune gastritis and is seen more frequently in whites than in other races. It is particularly common in Asia, Scandinavia, and parts of Europe and Latin America.[12] As with autoimmune gastritis, it is associated with

reduced gastric acid secretion, but achlorhydria and pernicious anemia are uncommon.

Chronic autoimmune gastritis and multifocal atrophic gastritis usually cause few symptoms related directly to gastric changes. When severe parietal cell loss occurs in the presence of autoimmune gastritis, hypochlorhydria or achlorhydria and hypergastrinemia are characteristically present. More important is the relationship of chronic gastritis to the development of peptic ulcer and gastric carcinoma. The long-term risk of gastric cancer in people with autoimmune gastritis is miniscule.[30]

Chemical Gastropathy. Chemical gastropathy is a chronic gastric injury resulting from reflux of alkaline duodenal contents, pancreatic secretions, and bile into the stomach. It is most commonly seen in people who have had gastroduodenostomy or gastrojejunostomy surgery. A milder form may occur in people with gastric ulcer, gallbladder disease, or various motility disorders of the distal stomach.

Peptic Ulcer Disease

Peptic ulcer disease is a term used to describe a group of ulcerative disorders that occur in areas of the upper GI tract that are exposed to acid–pepsin secretions. It is related to variety of causes, such as medication use and *H. pylori* infection.[35] Peptic ulcer disease, with its remissions and exacerbations, is a chronic health problem.

Peptic Ulcers

The most common forms of peptic ulcer are duodenal and gastric ulcers. Approximately 10% of the population have or will develop a peptic ulcer.[12] Duodenal ulcers occur five times more commonly than gastric ulcers. The peak age for peptic ulcer has progressively increased in the last 50 years and is now between 30 and 60 years of age for duodenal ulcers, although the disorder can occur in people of any age. Gastric ulcers are more prevalent among middle-aged and older adults. For duodenal ulcers, there is a male predominance, whereas the incidence of gastric ulcers is more equally distributed between men and women.[12]

A peptic ulcer can affect one or all layers of the stomach or duodenum (Fig. 45.6). The ulcer may penetrate only the mucosal surface, or it may extend into the smooth muscle layers. Occasionally, an ulcer penetrates the outer wall of the stomach or duodenum. Spontaneous remissions and exacerbations are common. Healing of the muscularis layer involves replacement with scar tissue. Although the mucosal layers that cover the scarred muscle layer regenerate, the regeneration often is less than perfect, which contributes to repeated episodes of ulceration.

Etiology and Pathogenesis. A variety of risk factors have been shown to have an association with peptic ulcer disease. The two most important are infection with the bacteria *H. pylori* and use of aspirin and other NSAIDs.[35] Both *H. pylori* infection and exposure to NSAIDs have been shown to impair the mechanisms that protect the gastric mucosa from the destructive effects of the corrosive acid that is continually challenging the upper GI tract mucosa, and ulceration reflects a failure of these mechanisms.

The exact mechanism by which *H. pylori* promotes the development of peptic ulcer has not been fully elucidated. *Helicobacter pylori*'s ability to induce inflammation and stimulate the release of cytokines and other mediators of inflammation contributes to mucosal damage. Infection, predominantly in the antrum of the stomach, leads to hypergastrinemia and an increased acid production. Acid injury to the duodenum is thought to promote the development of gastric metaplasia, allowing the organism to colonize these areas and promote the development of duodenal ulcers.

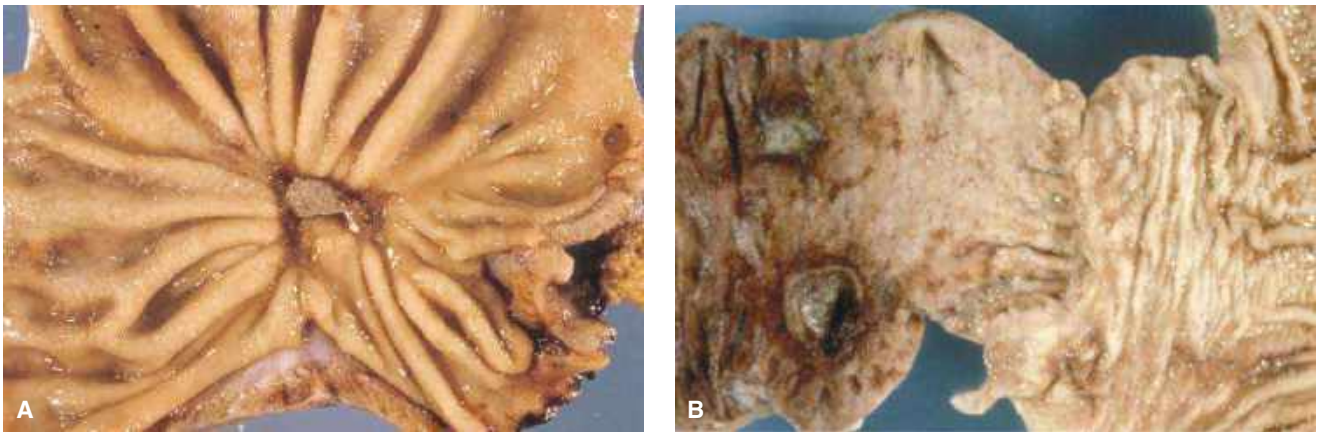


FIGURE 45.6 • Gastric and duodenal ulcers. (**A**) Gastric ulcer—there is a characteristic sharp demarcation from the surrounding mucosa, with radiating gastric folds. The base of the ulcer is gray owing to fibrin deposition. (**B**) Duodenal ulcer—there are two sharply demarcated duodenal ulcers surrounded by inflamed duodenal mucosa. The gastroduodenal junction is in the midportion of the photograph. (From Rubin E., Strayer D. (Eds.) (2012). *Rubin's pathology: Clinicopathologic foundations of medicine*. (6th ed., p. 625). Philadelphia, PA: Lippincott Williams & Wilkins.)

The pathogenesis of NSAID-induced ulcers is thought to involve mucosal injury and inhibition of prostaglandin synthesis.[30] Aspirin appears to be the most ulcerogenic of the NSAIDs. Ulcer development in NSAID users is dose dependent, but some risk occurs even with aspirin doses of 81 mg/day.[36] In contrast to peptic ulcer from other causes, NSAID-induced gastric injury often is without symptoms, and life-threatening complications can occur without warning. There is reportedly less gastric irritation with the newer class of NSAIDs that selectively inhibit cylcooxygenase-2 (COX-2–selective NSAIDs), the principal enzyme involved in prostaglandin synthesis at the site of inflammation, than with the nonselective NSAIDs that also inhibit COX-1, the enzyme involved in prostaglandin production in the gastric mucosa.

Epidemiologic studies have identified independent factors that augment the effect of *H. pylori* infection and NSAID-produced peptic ulcer disease. These factors include advancing age, a prior history of peptic ulcer, multiple NSAID use, and concurrent use of warfarin (an anticoagulant) and corticosteroid drugs. Smoking may augment the risk of peptic ulcer by impairing healing. Alcohol use may cause increased acid production.[35] There is no convincing evidence that dietary factors play a role in development of peptic ulcer. There is increased incidence of peptic ulcer in families. This finding is likely due to familial clustering of *H. pylori* infection, and inherited genetic factors reflecting responses to the organism likely play a secondary role.

Clinical Manifestations. The clinical manifestations of uncomplicated peptic ulcer focus on discomfort and pain. The pain, which is described as burning, gnawing, or cramplike, usually is rhythmic and frequently occurs when the stomach is empty—between meals and at 1 or 2 o'clock in the morning. The pain usually is located over a small area near the midline in the epigastrium near the xiphoid and may radiate below the costal margins, into the back, or, rarely, to the right shoulder. Superficial and deep epigastric tenderness and voluntary muscle guarding may occur with more extensive lesions. An additional characteristic of ulcer pain is periodicity. The pain tends to recur at intervals of weeks or months. During an exacerbation, it occurs daily for a period of several weeks and then remits until the next recurrence. Characteristically, the pain is relieved by food or antacids.

The most common complications of peptic ulcer are hemorrhage, perforation and penetration, and gastric outlet obstruction. Hemorrhage is caused by bleeding from granulation tissue or from erosion of an ulcer into an artery or vein. Acute post hemorrhagic anemia is the second most common secondary diagnosis, when people are admitted to the hospital with peptic ulcer disease.[37] Evidence of bleeding may consist of hematemesis or melena. Bleeding may be sudden, severe, and without warning, or it may be insidious, producing only occult blood in the stool. Up to 20% of people with bleeding ulcers have no antecedent symptoms of pain; this is particularly true in people using NSAIDs. Acute hemorrhage is evidenced by the sudden onset of weakness; dizziness; thirst; cold, moist skin; the desire to defecate; and the passage of loose, tarry, or even red stools and coffee-ground emesis. Signs of circulatory shock develop depending on the amount of blood lost.

Perforation occurs when an ulcer erodes through all the layers of the stomach or duodenum wall. When perforation occurs in older adults, their mortality is significantly increased. With perforation, GI contents enter the peritoneum and cause peritonitis. Radiation of the pain into the back, severe night distress, and inadequate pain relief from eating foods or taking antacids in persons with a long history of peptic ulcer may signify perforation. Penetration is a process similar to perforation, but with penetration the ulcer crater erodes into adjacent organs, including the pancreas, liver, or biliary tree.[35] Typically it has a subtle presentation marked by a gradual increase in severity and frequency of pain.

Outlet obstruction is caused by edema, spasm, or contraction of scar tissue and interference with the free passage of gastric contents through the pylorus or adjacent areas. The presentation of an obstruction is typically insidious, with symptoms of early satiety, feeling of epigastric fullness and heaviness after meals, gastroesophageal reflux, weight loss, and abdominal pain. With severe obstruction, there is vomiting of undigested food.

Diagnosis and Treatment. Diagnostic procedures for peptic ulcer include history taking, laboratory tests, radiologic imaging, and endoscopic examination. The history should include careful attention to aspirin and NSAID use. Peptic ulcer should be differentiated from other causes of epigastric pain. Laboratory findings of hypochromic anemia and occult blood in the stools indicate bleeding. Endoscopy (*i.e.,* gastroscopy and duodenoscopy) can be used to visualize the ulcer area and obtain biopsy specimens to test for *H. pylori* and exclude malignant disease. X-ray studies with a contrast medium such as barium are used to detect the presence of an ulcer crater and to exclude gastric carcinoma.

The treatment of peptic ulcer has changed dramatically over the past several decades and now aims to eradicate the cause and promote a permanent cure for the disease. Pharmacologic treatment focuses on eradicating *H. pylori*, relieving ulcer symptoms, and healing the ulcer crater. Acid-neutralizing, acid-inhibiting drugs and mucosa-protective agents are used to relieve symptoms and promote healing of the ulcer crater. There is no evidence that special diets are beneficial in treating peptic ulcer. Aspirin and NSAID use should be avoided when possible.

There are two pharmacologic methods for reducing gastric acid content. The first involves the neutralization of gastric acid through the use of antacids, and the second a decrease in gastric acid production through the use of H_2-receptor antagonists or proton pump inhibitors. Essentially three types of antacids are used to relieve gastric acidity: calcium carbonate, aluminum hydroxide, and magnesium hydroxide. Many antacids contain a combination of ingredients, such as magnesium-aluminum hydroxide. *Calcium preparations*

are constipating and may cause hypercalcemia and the milk-alkali syndrome. This syndrome is the third leading cause for hypercalcemia hospital admissions.[38] *Magnesium hydroxide* is a potent antacid that also has laxative effects. Approximately 5% to 10% of the magnesium in this preparation is absorbed from the intestine; because magnesium is excreted through the kidneys, this formulation should not be used in persons with renal failure. *Aluminum hydroxide* reacts with hydrochloric acid to form aluminum chloride. It combines with phosphate in the intestine, and prolonged use may lead to phosphate depletion and osteoporosis. Because antacids can decrease the absorption, bioavailability, and renal elimination of a number of drugs, this should be considered when antacids are administered with other medications.

Histamine is the major physiologic mediator for hydrochloric acid secretion. The H_2-receptor antagonists block gastric acid secretion stimulated by histamine, gastrin, and acetylcholine. The absorption of the drug is not altered by the presence or absence of food in the stomach.[39] The volume of gastric secretion and the concentration of pepsin also are reduced. The proton pump inhibitors block the final stage of hydrogen ion secretion by blocking the action of the gastric parietal cell proton pump.

Among the agents that enhance mucosal defenses are sucralfate and prostaglandin analogs. The drug sucralfate, which is a complex salt of sucrose containing aluminum and sulfate, selectively binds to damaged ulcer tissue and serves as a barrier to acid, pepsin, and bile. Sucralfate also can directly absorb bile salts and initiate the secretion of bicarbonate and mucus.[39] The drug is not absorbed systemically. The drug requires an acid pH for activation and should not be administered with antacids or an H_2 antagonist. Misoprostol, a prostaglandin E derivative, promotes ulcer healing by stimulating mucus and bicarbonate secretion and by modestly inhibiting acid secretion. It is the only drug in this class approved by the U.S. Food and Drug Administration (FDA) for clinical use in the prevention of NSAID-induced peptic ulcers. The drug causes dose-dependent diarrhea, and because of its stimulant effect on the uterus, it is contraindicated in women of childbearing age.

The current surgical management of peptic ulcer disease is largely limited to treatment of complications. When surgery is needed, it usually is performed using minimally invasive methods. With bleeding ulcers, hemostasis often can be achieved by endoscopic methods, and endoscopic balloon dilation often is effective in relieving outflow obstruction.

Zollinger-Ellison Syndrome

The Zollinger-Ellison syndrome is a rare condition caused by a gastrin-secreting tumor (gastrinoma). In persons with this disorder, increased gastric acid secretion results in GERD or severe peptic ulcer disease.[40] The tumors may be single or multiple; duodenal tumors account for 50% to 88% of this type of gastrinoma.[41] Approximately 50% of gastrin-producing tumors are malignant.[42] The increased gastric secretions cause symptoms related to peptic ulcer. Diarrhea may result from hypersecretion or from the inactivation of intestinal lipase and impaired fat digestion that occur with a decrease in intestinal pH.

Hypergastrinemia may also occur in an autosomal dominant disorder called the *multiple endocrine neoplasia type 1 (MEN 1) syndrome*, which is characterized by multiple endocrine neoplasms. The syndrome is characterized by hyperparathyroidism and multiple endocrine tumors, including gastrinomas. Approximately 20% to 25% of gastrinomas are due to MEN 1.[40]

The diagnosis of the Zollinger-Ellison syndrome is based on elevated serum gastrin and basal gastric acid levels and elimination of the MEN 1 syndrome as a cause of the disorder. Computed tomography (CT), abdominal ultrasonography, and selective angiography are used to localize the tumor and determine if metastatic disease is present.

Treatment of Zollinger-Ellison syndrome involves control of gastric acid secretion by proton pump inhibitors and treatment of the malignant neoplasm.[42] Surgical removal is indicated when the tumor is malignant and has not metastasized.

Stress Ulcers

A stress ulcer refers to GI ulcerations that develop in relation to major physiologic stress.[12] People at high risk for development of stress ulcers include those with large–surface-area burns (Curling's ulcer),[43] trauma, sepsis, acute respiratory distress syndrome, severe liver failure, and major surgical procedures. These lesions occur most often in the fundus and body of the stomach and are thought to result from ischemia to the mucosal tissue and alterations in the gastric mucosal barrier.[43]Another form of stress ulcer, called *Cushing ulcer*, consists of gastric, duodenal, and esophageal ulcers arising in persons with intracranial injury, operations, or tumors. They are thought to be caused by hypersecretion of gastric acid resulting from stimulation of vagal nuclei by increased intracranial pressure.

People admitted to hospital intensive care units are at particular risk for development of stress ulcers.[44] Proton pump inhibitors are the first line of medications used in the prevention of stress ulcers.[43]

Cancer of the Stomach

According to the International Agency for Research in Cancer, in 2008, gastric carcinoma was the fourth most common type of cancer in the world. Half of the global cases are reported in Eastern Asia. Less than 30% of all cases occur in developed countries, and the global incidence of occurrence in males over females is 2 to 1.[45] In 2010, 21,000 new diagnoses of stomach cancer were predicted in the United States and almost 11,000 people were predicted to die because of the disease.[46]

Etiology and Pathogenesis

Factors thought to increase the risk of gastric cancer include genetic factors, carcinogenic factors in the diet (*e.g., N*-nitroso compounds and benzopyrene found in smoked and preserved foods), autoimmune gastritis, and gastric adenomas or polyps.

The incidence of stomach cancer in the United States has significantly decreased since 1930, presumably because of improved storage of food with decreased consumption of salted, smoked, and preserved foods.[46] Chronic infection with *H. pylori* appears to serve as a cofactor in some types of gastric carcinomas. The bacterial infection causes gastritis, followed by atrophy, intestinal metaplasia, and carcinoma. This sequence of cellular events depends on both the presence of the bacterial proteins and the host immune response, with the latter being influenced by the host genetic background. In addition to genetics, the likelihood of developing gastric cancer from an *H. pylori* infection is related to strain of *H. pylori* infection, environmental factors, and the duration of infection.[47] Autoimmune gastritis, like *H. pylori* infection, increases the risk of gastric cancer, presumably due to chronic inflammation and intestinal metaplasia.[48]

Between 50% and 60% of gastric cancers occur in the pyloric region or adjacent to the antrum. Compared with a benign ulcer, which has smooth margins and is concentrically shaped, gastric cancers tend to be larger, are irregularly shaped, and have irregular margins.

Clinical Manifestations

Unfortunately, stomach cancers often are asymptomatic until late in their course. Symptoms, when they do occur, usually are vague and include indigestion, anorexia, weight loss, vague epigastric pain, vomiting, and an abdominal mass. Because these symptoms are essentially nonspecific, early detection is difficult.

Diagnosis and Treatment

Diagnosis of gastric cancer is accomplished by a variety of techniques, including barium x-ray studies, endoscopic studies with biopsy, and cytologic studies (*e.g.,* Papanicolaou smear) of gastric secretions.[49] Cytologic studies can prove particularly useful as routine screening tests for persons with atrophic gastritis or gastric polyps. CT and endoscopic ultrasonography often are used to delineate the spread of a diagnosed stomach cancer.

Depending on the location and extent of the lesion, surgery in the form of radical subtotal gastrectomy usually is the treatment of choice. Irradiation and chemotherapy have not proved particularly useful as primary treatment modalities in stomach cancer. These methods usually are used for palliative purposes or to control metastatic spread of the disease.

IN SUMMARY

Disorders of the stomach include gastritis, peptic ulcer, and cancer of the stomach. Gastritis refers to inflammation of the gastric mucosa. Acute gastritis refers to a transient inflammation of the gastric mucosa; it is associated most commonly with local irritants such as bacterial endotoxins, caffeine, alcohol, and aspirin. Chronic gastritis is characterized by the absence of grossly visible erosions and the presence of chronic inflammatory changes leading eventually to atrophy of the glandular epithelium of the stomach. There are three main types of chronic gastritis: *H. pylori* gastritis, autoimmune gastritis and multifocal atrophic gastritis, and chemical gastropathy. *Helicobacter pylori* is an "S"-shaped bacterium that colonizes the mucus-secreting epithelial cells of the stomach. Infection increases the risk of chronic gastritis, peptic ulcer, gastric carcinoma, and low-grade B-cell lymphoma. Treatment of *H. pylori* infection involves the use of multidrug therapy aimed at increasing the pH of gastric secretions and antimicrobial agents designed to eradicate the organism.

Peptic ulcer is a term used to describe a group of ulcerative disorders that occur in areas of the upper GI tract that are exposed to acid–pepsin secretions, most commonly the duodenum and stomach. There are two main causes of peptic ulcer: *H. pylori* infection and aspirin or NSAID use. The treatment of peptic ulcer focuses on eradication of *H. pylori*, avoidance of gastric irritation from NSAIDs, and conventional pharmacologic treatment directed at symptom relief and ulcer healing.

The Zollinger-Ellison syndrome is a rare condition caused by a gastrin-secreting tumor in which gastric acid secretion reaches such levels that ulceration becomes inevitable. Stress ulcers, also called *Curling ulcers*, occur in relation to major physiologic stresses such as burns and trauma and are thought to result from ischemia, tissue acidosis, and bile salts entering the stomach in critically ill persons with decreased GI tract motility. Another form of stress ulcer, Cushing ulcer, occurs in persons with intracranial trauma or surgery and is thought to be caused by hypersecretion of gastric acid resulting from stimulation of vagal nuclei by increased intracranial pressure.

Although the incidence of cancer of the stomach has declined over the past 50 years in the United States, it remains the leading cause of death worldwide. Because there are few early symptoms with this form of cancer, the disease often is far advanced at the time of diagnosis.

DISORDERS OF THE SMALL AND LARGE INTESTINES

After completing this section of the chapter, you should be able to meet the following objectives:

- Compare the characteristics of Crohn disease and ulcerative colitis.
- Describe the pathogenesis of the symptoms associated with appendicitis.

There are many similarities in conditions that disrupt the integrity and function of the small and large intestines. The walls of the small and large intestines consist of five layers:

1. An inner mucosal layer, which lines the lumen of the intestine
2. A submucosal layer
3. A circular muscularis layer
4. A layer of longitudinal muscle fibers
5. An outer serosal layer

Among the conditions that cause altered intestinal function are irritable bowel syndrome (IBS), inflammatory bowel disease (IBD), diverticulitis, appendicitis, disorders of bowel motility (*i.e.,* diarrhea, constipation, and bowel obstruction), malabsorption syndrome, and cancers of the colon and rectum.

Irritable Bowel Syndrome

The term *irritable bowel syndrome* is used to describe a functional GI disorder characterized by a variable combination of chronic and recurrent intestinal symptoms not explained by structural or biochemical abnormalities. There is evidence to suggest that 10% to 15% of the U.S. population have the disorder and one in four people worldwide.[50]

Irritable bowel disease is characterized by persistent or recurrent symptoms of abdominal pain; altered bowel function; and varying complaints of flatulence, bloating, nausea and anorexia, constipation or diarrhea, and anxiety or depression. A hallmark of IBS is abdominal pain that is relieved by defecation and associated with a change in consistency or frequency of stools. Abdominal pain usually is intermittent, cramping, and in the lower abdomen. It does not usually occur at night or interfere with sleep. The condition is believed to result from dysregulation of intestinal motor activity and central neural functions modulated by the CNS.[51] People with IBS tend to experience increased motility and abnormal intestinal contractions in response to psychological and physiologic stresses. The role that psychological factors play in the disease is uncertain. Although changes in intestinal activity are normal responses to stress, these responses appear to be exaggerated in persons with IBS. Women tend to be affected more often than men. Menarche often is associated with onset of the disorder. Women frequently notice an exacerbation of symptoms during the premenstrual period, suggesting a hormonal component.

Clinical Manifestations and Diagnosis

Because IBS lacks anatomic or physiologic markers, diagnosis is usually based on signs and symptoms of abdominal pain or discomfort, bloating, and constipation or diarrhea, or alternating bouts of constipation and diarrhea. A commonly used set of diagnostic criteria require continuous or recurrent symptoms of at least 12 weeks' duration (which may be nonconsecutive) of abdominal discomfort or pain in the preceding 12 months, with two of three accompanying features: relief with defecation, onset associated with a change in bowel frequency, and onset associated with a change in form (appearance) of stool.[52]

Other symptoms that support the diagnosis of IBS include abnormal stool frequency (more than three times per day or less than three times per week), abnormal stool form (lumpy/hard or loose/watery), abnormal stool passage (straining, urgency, or feeling of incomplete evacuation), passage of mucus, and bloating or feeling of abdominal distention.[52] A history of lactose intolerance should be considered because intolerance to lactose and other sugars may be a precipitating factor in some people. The acute onset of symptoms raises the likelihood of organic disease, as does weight loss, anemia, fever, occult blood in the stool, nighttime symptoms, or signs and symptoms of malabsorption. These signs and symptoms require additional investigation of differential diagnoses.[53]

Treatment

The treatment of IBS focuses on methods of stress management, particularly those related to symptom production. Reassurance is important. Usually, no special diet is indicated, although adequate fiber intake usually is recommended. Avoidance of offending dietary substances by following specific elimination diets that omit such foods as fatty and gas-producing foods, alcohol, and caffeine-containing beverages may be beneficial.[54] Various pharmacologic agents, including antispasmodic and anticholinergic drugs, have been used with varying success in treatment of the disorder. Alosetron, a 5-HT_3 antagonist, was the first specific drug to be approved by the FDA for the treatment of irritable bowel disease. It acts by reducing intestinal secretion, decreasing visceral afferent nerve activity (thereby reducing abdominal pain), and reducing intestinal motility. The drug, which was indicated for treatment of women with the severe diarrheal form of the disease, was removed from the market in late 2000 because of serious side effects involving ischemic colitis and severe constipation and then reintroduced in 2002 under a restricted prescribing program.[55]

Inflammatory Bowel Disease

The term *inflammatory bowel disease* is used to designate two related inflammatory intestinal disorders: Crohn disease and ulcerative colitis. The worldwide prevalence of IBD is 396 in 10,000 persons.[56] Although the two diseases differ sufficiently to be distinguishable, they have many features in common. Both diseases produce inflammation of the bowel, both lack confirming evidence of a proven causative agent, both have a pattern of familial occurrence, and both can be accompanied by systemic manifestations. Crohn disease most commonly affects the distal small intestine and proximal colon, but can affect any area of the GI tract from the esophagus to the anus, whereas ulcerative colitis is confined to the colon and rectum (Fig. 45.7). The distinguishing characteristics of Crohn disease and ulcerative colitis are summarized in Table 45.1.

Etiology and Pathogenesis

A remarkable feature of the GI tract is that the mucosal immune system is always ready to respond against ingested pathogens but is unresponsive to the normal intestinal microflora. According to the currently accepted hypothesis, this normal

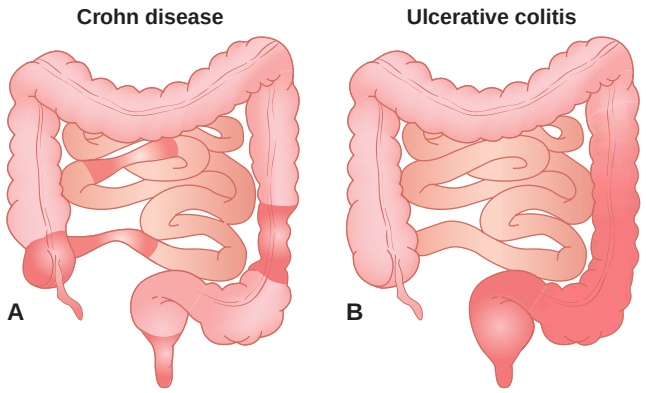


FIGURE 45.7 • Distribution patterns of disease with (**A**) skip lesions in Crohn disease and (**B**) continuous involvement of the colon, beginning with the rectum, in ulcerative colitis.

state of homeostasis is disrupted in IBD, leading to unregulated and exaggerated immune responses. The question remains if the response is an appropriate defense mechanism to a pathogen or is the immune system responding in an inappropriate manner. Thus, as in many other autoimmune disorders, the pathogenesis of Crohn disease and ulcerative colitis involves a failure of immune regulation, genetic predisposition, and an environmental trigger, especially microbial flora.[57]

Genetic Susceptibility. The genetic basis of IBD has long been suspected. First-degree relatives of people diagnosed with IBD have a 30 to 100 times greater incidence of IBD.[58] For Crohn disease, a recent study found a concordance rate of 27% in monozygotic twins and 2% in dizygotic twins.[59] With ulcerative colitis, this genetic component was found to be weaker, but still present. These associations clearly indicate that genetic susceptibility plays an important role in the development of IBD. However, classic Mendelian inheritance patterns are not seen, and IBD therefore cannot be attributed to a single gene. Many candidate genes are known to be associated with, and likely to contribute to, the development of IBD. These include the human leukocyte antigen (HLA) associations. Accumulating evidence also suggests that both Crohn disease and ulcerative colitis are associated with profound disorders of mucosal immunity. The IBD1 locus on chromosome 16 has recently been shown to contribute to Crohn disease.[60] The product of the implicated gene, NOD2 (so named because the coded protein has a nucleotide oligomerization domain) activates the nuclear factor kappa beta (NFκβ) transcription factor. The NOD2 protein is expressed in many types of leukocytes as well as epithelial cells and is thought to act as an intracellular receptor for lipopolysaccharides on microbes. On binding microbial products, it may trigger the NFκβ pathway, which leads to the production of cytokines and other proteins involved in the innate immune defense against microorganisms. The NOD2 mutations that are associated with Crohn disease may reduce the activity of the protein, resulting in persistence of intracellular microbes and prolonged immune responses. Another region studied extensively is IBD3 on chromosome 6. This is the area that includes the HLA complex that has been linked to Crohn disease and ulcerative colitis. Another area linked specifically to Crohn disease is on chromosome 5q (IBD5). This area is rich in genes encoding several cytokines that may contribute to the disease.

Role of Environmental Factors. Animal studies have definitively established the importance of the gut flora in IBD. The sites affected by IBD, the distal ileum and the colon, are awash with bacteria. Although it is unlikely that IBD is caused by microbes, it seems likely that microbes may provide the antigen trigger for an unregulated immune response.

Another environmental factor related to IBD is smoking.[57] Crohn disease is more commonly associated with people who currently smoke while ulcerative colitis is associated with people who have never smoked or have quit. The relationship between nicotine and IBD is thought to be due to coagulopathies occurring in the intestine or as a result of an immune response.

TABLE 45.1 DIFFERENTIATING CHARACTERISTICS OF CROHN DISEASE AND ULCERATIVE COLITIS

CHARACTERISTIC	CROHN DISEASE	ULCERATIVE COLITIS
Types of inflammation	Granulomatous	Ulcerative and exudative
Level of involvement	Primarily submucosal	Primarily mucosal
Extent of involvement	Skip lesions	Continuous
Areas of involvement	Primarily ileum, secondarily colon	Primarily rectum and left colon
Diarrhea	Common	Common
Rectal bleeding	Rare	Common
Fistulas	Common	Rare
Strictures	Common	Rare
Perianal abscesses	Common	Rare
Development of cancer	Uncommon	Relatively common

Clinical Manifestations

The clinical manifestations of both Crohn disease and ulcerative colitis are ultimately the result of activation of inflammatory cells with elaboration of inflammatory mediators that cause nonspecific tissue damage. Both diseases are characterized by remissions and exacerbations of diarrhea, fecal urgency, and weight loss. Acute complications, such as intestinal obstruction, may develop during periods of fulminant disease (Fig. 45.8).

A number of systemic manifestations have been identified in people with Crohn disease and ulcerative colitis. These include axial arthritis affecting the spine and sacroiliac joints and oligoarticular arthritis affecting the large joints of the arms and legs; inflammatory conditions of the eye, usually uveitis; skin lesions, especially erythema nodosum; stomatitis; and autoimmune anemia, hypercoagulability of blood, and sclerosing cholangitis. Occasionally, these systemic manifestations may herald the recurrence of intestinal disease. In children, growth retardation may occur, particularly if the symptoms are prolonged and nutrient intake has been poor.

Crohn Disease

Crohn disease is a recurrent, granulomatous type of inflammatory response that can affect any area of the GI tract. The terminal ileum or cecum is the most common portion of the bowel where inflammation occurs.[58,61] It is a slowly progressive, relentless, and often disabling disease. The disease usually strikes people in their twenties or thirties, with women being affected slightly more often than men.

A characteristic feature of Crohn disease is the sharply demarcated, granulomatous lesions that are surrounded by normal-appearing mucosal tissue. When the lesions are multiple, they often are referred to as *skip lesions* because they are interspersed between what appear to be normal segments of the bowel. All the layers of the bowel are involved, with the submucosal layer affected to the greatest extent. The surface of the inflamed bowel usually has a characteristic "cobblestone" appearance resulting from the fissures and crevices that develop, surrounded by areas of submucosal edema.[12,61] (Fig. 45.9). There usually is a relative sparing of the smooth muscle layers of the bowel, with marked inflammatory and fibrotic changes of the submucosal layer. The bowel wall, after a time, often becomes thickened and inflexible; its appearance has been likened to a lead pipe or rubber hose. The adjacent mesentery may become inflamed, and the regional lymph nodes and channels may become enlarged.

Clinical Manifestations. The clinical course of Crohn disease is variable; often, there are periods of exacerbations and remissions, with symptoms being related to the location of the lesions. The principal symptoms, which are dependent upon the area of the GI system that is affected, include diarrhea, abdominal pain, weight loss, fluid and electrolyte disorders, malaise, and low-grade fever.[61] Because Crohn disease affects the submucosal layer to a greater extent than the mucosal layer, there is less bloody diarrhea than with ulcerative colitis. Ulceration of the perianal skin is common, largely because of the severity of the diarrhea. The absorptive surface of the intestine may be disrupted; nutritional deficiencies may occur, related to the specific segment of the intestine involved. When Crohn disease occurs in childhood, one of its major manifestations may be retardation of growth and significant malnutrition.[62]

Complications of Crohn disease include fistula formation, abdominal abscess formation, and intestinal obstruction. Fistulas are tubelike passages that form connections between different sites in the GI tract. They also may develop between other sites, including the bladder, vagina, urethra, and skin. Perineal fistulas that originate in the ileum are relatively

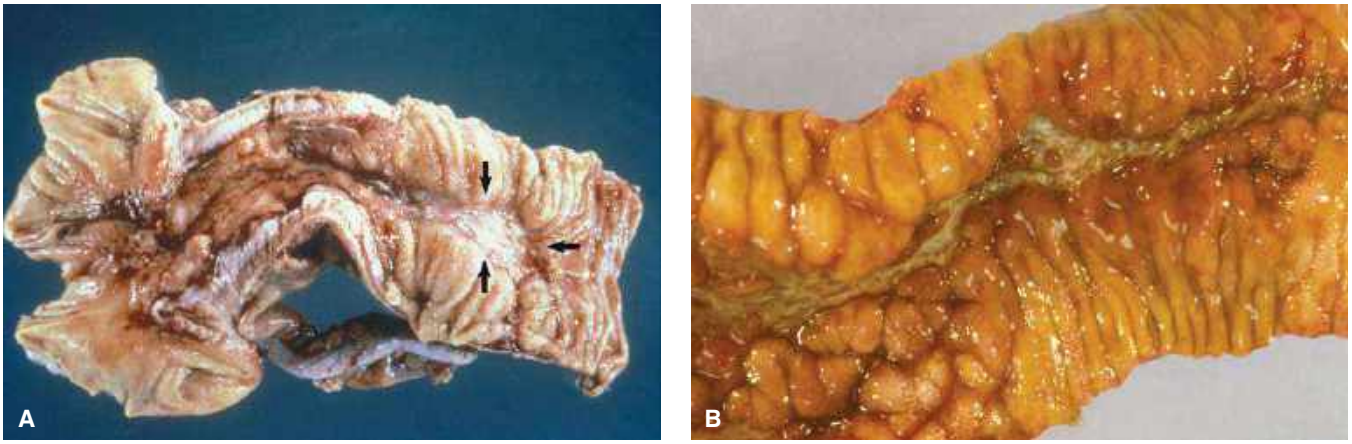


FIGURE 45.8 • Crohn disease. (**A**) The terminal ileum shows striking thickening of the wall of the distal portion with distortion of the ileocecal valve. A longitudinal ulcer is also depicted (see *arrows*). (**B**) The mucosal surface of the colon displays a "cobblestone" appearance owing to the presence of linear ulcerations and edema and inflammation of the intervening tissue. (From Rubin E., Strayer D. (Eds.) (2012). *Rubin's pathology: Clinicopathologic foundations of medicine* (6th ed., p. 654). Philadelphia, PA: Lippincott Williams & Wilkins.)

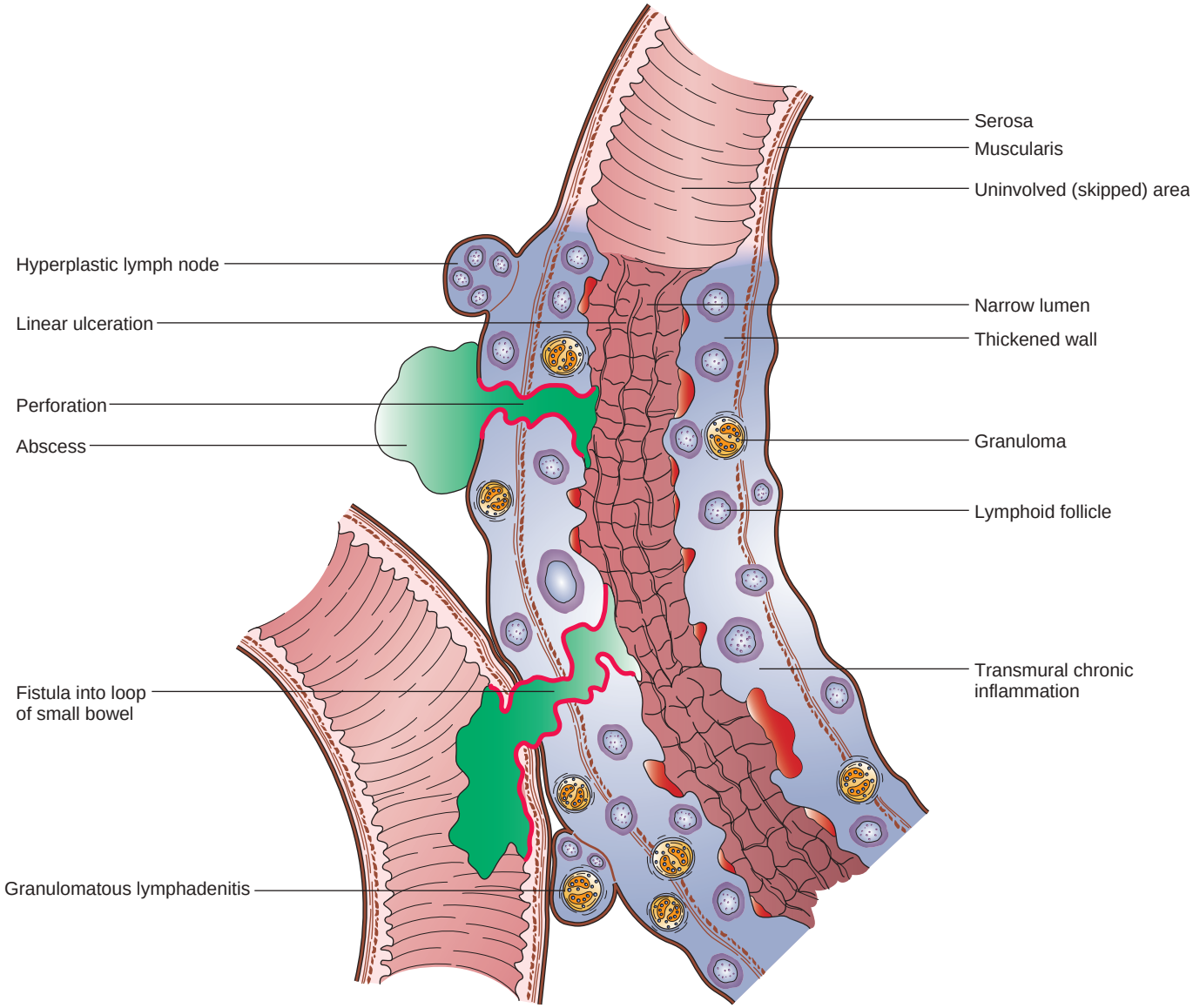


FIGURE 45.9 • Crohn disease. The major components of Crohn disease in the small intestine. (From Rubin E., Strayer D. (Eds.) (2012). *Rubin's pathology: Clinicopathologic foundations of medicine* (6th ed., p. 655). Philadelphia, PA: Lippincott Williams & Wilkins.)

common.[56] Fistulas between segments of the GI tract may lead to malabsorption, syndromes of bacterial overgrowth, and diarrhea. They also can become infected and cause abscess formation.

Diagnosis. The diagnosis of Crohn disease requires a thorough history and physical examination. Sigmoidoscopy is used for direct visualization of the affected areas and to obtain biopsies. Measures are taken to exclude infectious agents as the cause of the disorder. This usually is accomplished by the use of stool cultures and examination of fresh stool specimens for ova and parasites. In people suspected of having Crohn disease, radiographic contrast studies provide a means for determining the extent of involvement of the small bowel and establishing the presence and nature of fistulas. CT scans may be used to detect an inflammatory mass or abscess.

Treatment. Treatment methods focus on terminating the inflammatory response and promoting healing, maintaining adequate nutrition, and preventing and treating complications. Several medications have been successful in suppressing the inflammatory reaction, including the corticosteroids, sulfasalazine, metronidazole, azathioprine, 6-mercaptopurine, methotrexate, and infliximab. Surgical resection of damaged bowel, drainage of abscesses, or repair of fistula tracts may be necessary.

Sulfasalazine is a topically active agent that has a variety of anti-inflammatory effects. The beneficial effects of

sulfasalazine are attributable to one component of the drug, 5-aminosalicylic acid (5-ASA). Agents containing 5-ASA affect multiple sites in the arachidonic acid pathway critical to the pathogenesis of inflammation. Sulfasalazine contains 5-ASA with sulfapyridine linked to an azo bond. The drug is poorly absorbed from the intestine, and the azo linkage is broken down by the bacterial flora in the ileum and colon to release 5-ASA. Metronidazole is an antibiotic used to treat bacterial overgrowth in the small intestine. A recent meta-analysis found two thiopurine drugs, azathioprine and 6-mercaptopurine, to be effective in reducing the reoccurrence of Crohn disease.[63] The use of methotrexate is another option for clinicians to choose instead of the thiopurine drugs, although the studies regarding its use are limited.[57] Infliximab is a monoclonal antibody that targets the destruction of tumor necrosis factor (TNF), a mediator of the inflammatory response, whose expression is increased in inflammatory processes such as Crohn disease.[61] It is the first drug approved specifically for Crohn disease and is used in the treatment of people with active moderate-to-severe Crohn disease who have had an inadequate response to corticosteroids or other immune modulators. Although infliximab is currently the only anti-TNF agent approved for treatment of persons with IBD, controlled studies of other anti-TNF and immunomodulating agents such as thalidomide, adalimumab, and certolizumab Pegol are ongoing.[61]

Nutritional deficiencies are common in Crohn disease because of diarrhea, steatorrhea, and other malabsorption problems. A nutritious diet that is high in calories, vitamins, and proteins is recommended. Because fats often aggravate the diarrhea, it is recommended that they be avoided. Elemental diets, which are nutritionally balanced but residue free and bulk free, may be given during the acute phase of the illness. These diets are largely absorbed in the jejunum and allow the inflamed bowel to rest. Total parenteral nutrition (*i.e.,* parenteral hyperalimentation) consists of intravenous administration of hypertonic glucose solutions to which amino acids and fats may be added. This form of nutritional therapy may be needed when food cannot be absorbed from the intestine. Because of the hypertonicity of these solutions, they must be administered through a large-diameter central vein.

Ulcerative Colitis

Ulcerative colitis is a nonspecific inflammatory condition of the colon. The disease is more common in the United States and Western countries. The disease may arise at any age, with a peak incidence between ages 15 and 25 years.[58] Unlike Crohn disease, which can affect various sites in the GI tract, ulcerative colitis is confined to the rectum and colon. The disease usually begins in the rectum and spreads proximally, affecting primarily the mucosal layer, although it can extend into the submucosal layer. The length of proximal extension varies. It may involve the rectum alone (ulcerative proctitis), the rectum and sigmoid colon (proctosigmoiditis), or the entire colon (pancolitis). The inflammatory process tends to be confluent and continuous instead of skipping areas, as it does in Crohn disease.

Characteristic of the disease are the lesions that form in the crypts of Lieberkühn in the base of the mucosal layer. The inflammatory process leads to the formation of pinpoint mucosal hemorrhages, which in time suppurate and develop into *crypt abscesses*. These inflammatory lesions may become necrotic and ulcerate. Although the ulcerations usually are superficial, they often extend, causing large denuded areas (Fig. 45.10). As a result of the inflammatory process, the mucosal layer often develops tonguelike projections that resemble polyps and therefore are called *pseudopolyps*. The bowel wall thickens in response to repeated episodes of colitis.

Clinical Manifestations. Ulcerative colitis typically presents as a relapsing disorder marked by attacks of diarrhea. The diarrhea may persist for days, weeks, or months and then subside, only to recur after an asymptomatic interval of several months to years or even decades. Because ulcerative colitis affects the mucosal layer of the bowel, the stools typically contain blood and mucus. Nocturnal diarrhea usually occurs when daytime symptoms are severe. There may be mild abdominal cramping and fecal incontinence. Anorexia, weakness, and fatigability are common.

Based on clinical and endoscopic findings, the disease is characterized by how much of the colon is affected and the extent of the inflammation. Severity is defined as mild, moderate, severe, or fulminant.[58] The most common form of the disease is the mild form, in which the person has less than four stools daily, with or without blood, no systemic signs of toxicity, and a normal erythrocyte sedimentation rate (ESR). People with moderate disease have more than four stools daily, but have minimal signs of toxicity. Severe disease is manifested by more than six bloody stools daily, and evidence of toxicity as demonstrated by fever, tachycardia, anemia, and elevated ESR (Fig. 45.11). People with fulminant disease have features that include more than 10 bowel move-

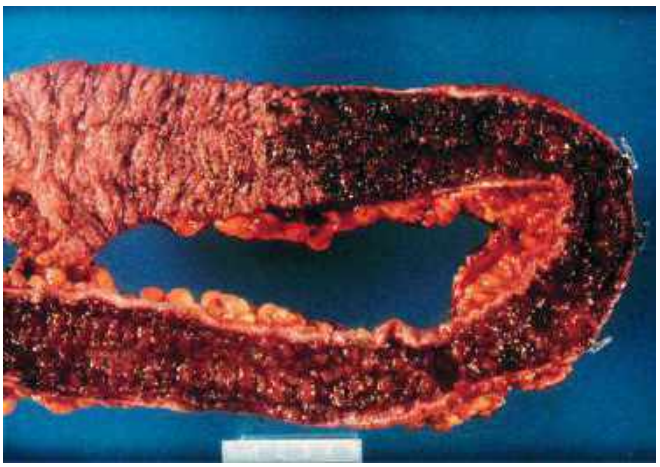


FIGURE 45.10 • Ulcerative colitis. Prominent erythema and ulceration of the colon begin in the ascending colon and are most severe in the rectosigmoid area. (From Rubin R., Strayer D. S. (Eds.) (2012). *Rubin's pathophysiology: Clinicopathologic foundations of medicine* (6th ed., p. 656). Philadelphia, PA: Lippincott Williams & Wilkins.)

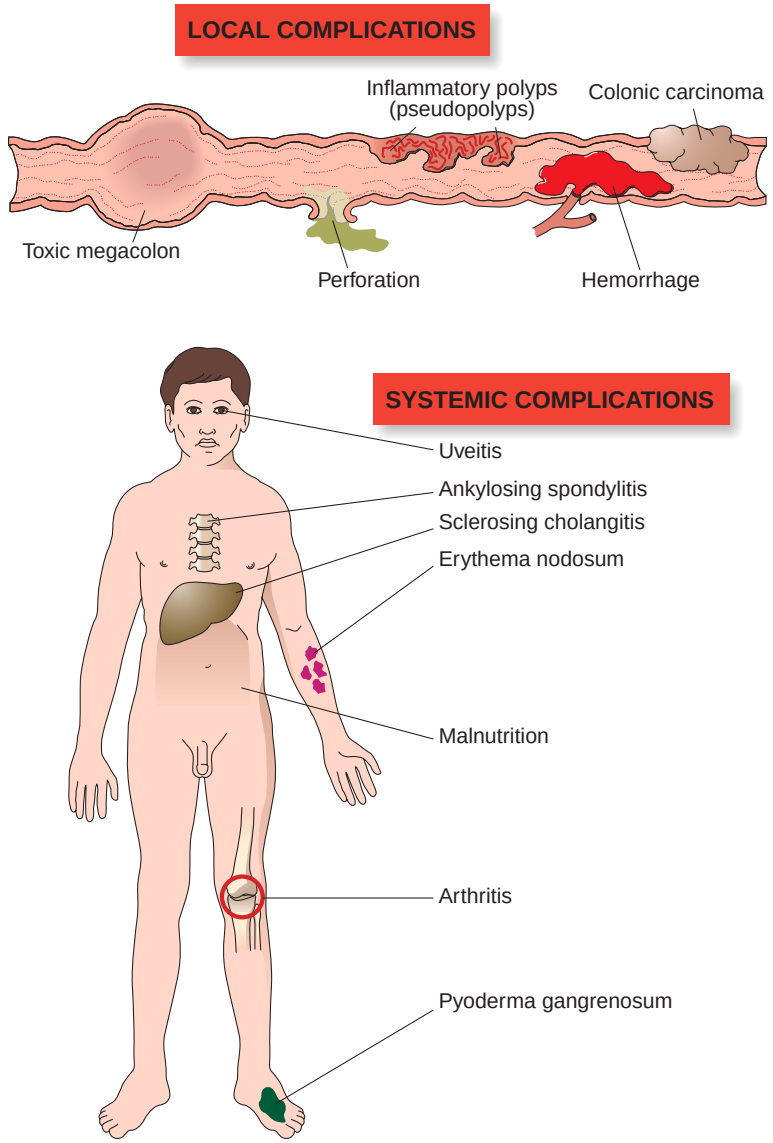


FIGURE 45.11 • Complications of ulcerative colitis. (From Rubin R., Strayer D. S. (Eds.) (2012). *Rubin's pathophysiology: Clinicopathologic foundations of medicine* (6th ed., p. 658). Philadelphia, PA: Lippincott Williams & Wilkins.)

ments a day, continuous bleeding, fever and other signs of toxicity, abdominal tenderness and distention, need for blood transfusions, and colonic dilation on abdominal radiographs. These people are at risk for development of toxic megacolon, which is characterized by dilation of the colon and signs of systemic toxicity. It results from extension of the inflammatory response, with involvement of neural and vascular components of the bowel.

Diagnosis and Treatment. Diagnosis of ulcerative colitis is based on history and physical examination. The diagnosis usually is confirmed by sigmoidoscopy, colonoscopy, biopsy, and by negative stool examinations for infectious or other causes. Colonoscopy should not be performed on people with severe disease because of the danger of perforation, but may be performed after demonstrated improvement to determine the extent of disease and need for subsequent cancer surveillance.

Treatment depends on the extent of the disease and severity of symptoms. It includes measures to control the acute manifestations of the disease and prevent recurrence. Some people with mild-to-moderate symptoms are able to control their symptoms simply by avoiding caffeine, lactose (milk), highly spiced foods, and gas-forming foods. Fiber supplements may be used to decrease diarrhea and rectal symptoms. Surgical treatment (*i.e.,* removal of the rectum and entire colon) with the creation of an ileostomy or ileoanal anastomosis may be required for people who do not respond to medications and conservative methods of treatment.

The medications used in treatment of ulcerative colitis are similar to those used in the treatment of Crohn disease. They include the nonabsorbable 5-ASA compounds (*e.g.,* mesalamine, olsalazine).[64] The corticosteroids are used selectively to lessen the acute inflammatory response. Many of these medications can be administered rectally by suppository

or enema. Immunomodulating drugs and anti-TNF therapies may be used to treat persons with severe colitis.

Cancer of the colon is one of the feared long-term complications of ulcerative colitis. Ulcerative colitis is characterized by deoxyribonucleic acid (DNA) damage with microsatellite instability in mucosa cells. More recently, genomic instability was detected in nondysplastic areas of people with ulcerative colitis, suggesting that these people have DNA repair deficiency and genomic instability throughout the intestinal tract.[65] In a meta-analysis focusing on studies involving people with ulcerative colitis, the cumulative risk for people to have colorectal cancer was 1.6% by 10 years, 8.6% by 20 years, and 18.4% by 30 years.[66] All people with the diagnosis should receive a colonoscopy for screening purposes within 8 years after they begin to have symptoms. The frequency of surveillance colonoscopies is often every 1 to 3 years and is dependent upon the results of the examinations and biopsies obtained.[67]

Infectious Enterocolitis

A number of microbial agents, including viruses, bacteria, and protozoa, can infect the GI tract, causing diarrhea and sometimes ulcerative and inflammatory changes in the small or large intestine. Infectious enterocolitis is a global problem, causing more than 12,000 deaths per day among children in developing countries. Although far less common in industrialized countries, these disorders still have infection rates second only to the common cold. Most infections are spread by the oral–fecal route, often through contaminated water or food.

Viral Infection

Most viral infections affect the superficial epithelium of the small intestine, destroying these cells and disrupting their absorptive function. Repopulation of the small intestinal villi with immature enterocytes and preservation of crypt secretory cells leads to net secretion of water and electrolytes compounded by incomplete absorption of nutrients and osmotic diarrhea. Symptomatic disease is caused by several distinct viruses, including the rotavirus, which most commonly affects children 6 to 24 months of age; the norovirus (or Norwalk), which is responsible for the majority of nonbacterial food-borne epidemic gastroenteritis in all age groups; and enteric adenoviruses, which primarily affect children younger than 24 months.[68]

Rotavirus. Worldwide, rotavirus is the leading cause of severe diarrhea and is estimated to cause 527,000 children under the age of 5 to die each year.[69] Prior to 2006, the disease was responsible for 400,000 doctor visits and 20 to 60 deaths in children below the age of 5 in the United States.[70] In 2006, RotaTeq, a live oral vaccine for rotavirus was approved by the FDA. A different live vaccine was approved in 1998 but was withdrawn from the market less than a year later when several infants developed intussusception after receiving the vaccine.

The disease tended to be most severe in children 3 to 24 months of age. Infants younger than 3 months of age are relatively protected by transplacental antibodies and possibly by breast-feeding. The virus spreads via a fecal–oral route, and outbreaks are common in children in day care centers. The virus is shed before and for days after clinical illness. Very few infectious virions are needed to cause disease in a susceptible host.

Rotavirus infection typically begins after an incubation period of 1 to 3 days, with mild-to-moderate fever and vomiting, followed by onset of frequent watery stools.[68] The fever and vomiting usually disappear on about the second day, but the diarrhea continues for 5 to 7 days. Dehydration may develop rapidly, particularly in infants. Treatment is largely supportive. Avoiding and treating dehydration are the main goals.

Bacterial Infection

Infectious enterocolitis can be caused by a number of bacteria. There are several pathogenic mechanisms for bacterial enterocolitis: ingestion of preformed toxins that are present in contaminated food; infection by toxigenic organisms that proliferate in the gut lumen and produce an enterotoxin; and infection by enteroinvasive organisms, which proliferate in the lumen and invade and destroy mucosal epithelial cells. The pathogenic effects of bacterial infections depend on the ability of the organism to adhere to the mucosal epithelial cells, elaborate enterotoxins, and then invade the mucosal epithelial cells.

In general, bacterial infections produce more severe effects than viral infections. The complications of bacterial enterocolitis result from massive fluid loss or destruction of intestinal mucosa and include dehydration, sepsis, and perforation. Among the organisms that cause bacterial enterocolitis are *Staphylococcus aureus* (toxins associated with "food poisoning"), *Escherichia coli*, *Shigella* species, *Salmonella*, and *Campylobacter*.[71] Two particularly serious forms of bacterial enterocolitis are caused by *Clostridium difficile* and *E. coli* O157:H7.

***Clostridium difficile* Colitis.** *Clostridium difficile* colitis is associated with antibiotic therapy.[72–74] *Clostridium difficile* is a gram-positive, spore-forming bacillus that is part of the normal flora in 1% to 3% of humans.[51] The spores are resistant to the acid environment of the stomach and convert to vegetative forms in the colon. Treatment with broad-spectrum antibiotics predisposes to disruption of the normal protective bacterial flora of the colon, leading to colonization by *C. difficile* along with the release of toxins that cause mucosal damage and inflammation. Almost any antibiotic may cause *C. difficile* colitis, but broad-spectrum antibiotics with activity against gram-negative enteric bacteria are the most frequent agents. After antibiotic therapy has made the bowel susceptible to infection, colonization by *C. difficile* occurs by the oral–fecal route. *Clostridium difficile* infection usually is acquired in the hospital, where the organism is commonly encountered.

In general, *C. difficile* is noninvasive. Development of *C. difficile* colitis and diarrhea requires an alteration in the normal gut flora, acquisition and germination of the spores, overgrowth of *C. difficile*, and toxin production. The toxins bind to and damage the intestinal mucosa, causing hemorrhage, inflammation, and necrosis. The toxins also interfere with protein synthesis, attract inflammatory cells, increase capillary permeability, and stimulate intestinal peristalsis. The infection commonly manifests with diarrhea that is mild to moderate and sometimes is accompanied by lower abdominal cramping. Typically symptoms begin within 4 to 9 days after an antibiotic treatment has been started and, in most cases, systemic manifestations are absent, and the symptoms subside after the antibiotic has been discontinued.[75]

A more severe form of colitis, *pseudomembranous colitis*, is characterized by an adherent inflammatory membrane overlying the areas of mucosal injury. It is a life-threatening form of the disease. People with the disease are acutely ill, with lethargy, fever, tachycardia, abdominal pain and distention, and dehydration. The smooth muscle tone of the colon may be lost, resulting in toxic dilation of the colon. Prompt therapy is needed to prevent perforation of the bowel.

The diagnosis of *C. difficile*–associated diarrhea requires a careful history, with particular emphasis on antibiotic use. Diagnostic findings include a history of antibiotic use and laboratory tests that confirm the presence of *C. difficile* toxins in the stool. Treatment includes the immediate discontinuation of antibiotic therapy. Specific treatment aimed at eradicating *C. difficile* is used when symptoms are severe or persistent. Metronidazole is the drug of choice, with vancomycin being reserved for people who cannot tolerate metronidazole, do not respond to the drug, or have severe symptoms. Metronidazole can be given intravenously or oral. When given oral, it is absorbed from the upper GI tract and may cause side effects, such as nausea. Vancomycin can be given orally or via an enema. It is poorly absorbed systemically, and its actions are limited to the GI tract, resulting in a smaller number of side effects.[76]

***Escherichia coli* O157:H7 Infection.** *Escherichia coli* O157:H7 has become recognized as an important cause of epidemic and sporadic colitis.[77] *Escherichia coli* O157:H7 is a strain of *E. coli* found in the feces and contaminated milk of healthy dairy and beef cattle, but it also has been found in contaminated pork, poultry, and lamb. Infection usually is by food-borne transmission, often by ingesting undercooked hamburger. The organism also can be transferred to nonmeat products such as fruits and vegetables. Transmission has also been reported in persons swimming in a fecally contaminated lake as well as among visitors to farms and petting zoos, where children are in direct contact with animals. Person-to-person transmission may occur, particularly in nursing homes, day care settings, and hospitals. The very young and the very old are particularly at risk for the infection and its complications.

The infection may cause no symptoms or cause a variety of manifestations, including acute, nonbloody diarrhea; hemorrhagic colitis; hemolytic uremic syndrome (HUS); and thrombotic thrombocytopenic purpura. The infection often presents with abdominal cramping and watery diarrhea and subsequently may progress to bloody diarrhea. The diarrhea commonly lasts 5 to 10 days.[78]

Most strains of *E. coli* are harmless. However, enterohemorrhagic *E. coli* can release *Shigella*-like toxins that attach to and damage the mucosal lining of the intestine. Subsequently, the *Shigella*-like toxins gain access to the circulatory system and travel in the plasma and on the surface of platelets and monocytes. The *Shigella*-like toxins bind to high-affinity galactose-containing receptors in the membranes of glomerular, cerebral, or microvascular endothelial cells; renal mesangial and tubular cells; and monocytes and platelets.[79] Two complications of the infection, HUS and thrombotic thrombocytopenic purpura, reflect the effects of the *Shigella*-like toxins. The HUS is characterized by hemolytic anemia, thrombocytopenia, and renal failure. It occurs predominantly in infants and young children and is the most common cause of acute renal failure in children. A recent study found that HUS patients had a mortality rate of 4.6%.[80] Thrombotic thrombocytopenic purpura is manifested by thrombocytopenia, renal failure, fever, and neurologic manifestations. It often is regarded as the severe end of the disease that leads to HUS plus neurologic problems.

No specific therapy is available for *E. coli* O157:H7 infection. Treatment is largely symptomatic and directed toward treating the effects of complications. The use of antibiotics or antimotility/antidiarrheal agents in the early stages of diarrhea has been shown to increase the risk of HUS because the gut is exposed to a greater amount of toxins for a longer time.

Because of the seriousness of the infection and its complications, education of the public about techniques for decreasing primary transmission of the infection from animal sources is important. Undercooked meats and unpasteurized milk are sources of transmission. Food handlers and consumers should be aware of the proper methods for handling uncooked meat to prevent cross-contamination of other foods. Particular attention should be paid to hygiene in day care centers and nursing homes, where the spread of infection to the very young and very old may result in severe complications.

Protozoan Infection

Amebiasis refers to an infection by *Entamoeba histolytica* involving the colon and occasionally the liver.[30] Humans are the only known reservoir for *E. histolytica*, which reproduce in the colon and pass in the feces. Although *E. histolytica* infection occurs worldwide, it is more common and more severe in tropical and subtropical areas, where crowding and poor sanitation prevail. Intestinal amebiasis ranges from completely asymptomatic infection to serious dysenteric disease.

Entamoeba histolytica has two distinct stages: the trophozoites (ameboid form) and cysts.[81] The trophozoites thrive in the colon and feed on bacteria and human cells. They may colonize any portion of the large bowel, but the area of maximum disease is usually the cecum. Persons with symptomatic disease pass both cysts and trophozoites in their feces, but quickly die when exposed to air outside of the body. Only the

cysts are infectious because they survive gastric acidity, which destroys the trophozoites. Once established, the trophozoites invade the crypts of colonic glands and burrow down into the submucosa; the organism then fans out to create a flask-shaped ulcer with a narrow neck and broad base. *Entamoeba histolytica* that have invaded into the submucosal veins of the colon enter the portal vein and embolize to the liver to produce solitary and, less often, multiple discrete hepatic abscesses.[82]

Some people have an acute onset of diarrhea as early as 8 days (commonly 2 to 4 weeks) after infection.[83] Others may be asymptomatic or have only mild intestinal symptoms for months or several years before either intestinal symptoms or liver abscesses appear. Manifestations include abdominal discomfort, tenderness, cramps, and fever, often accompanied by nausea, vomiting, and passage of malodorous flatus. There may be frequent passage of liquid stools containing bloody mucus, but the duration of diarrhea is not usually so prolonged as to cause dehydration. The infection often persists for months or years, causing emaciation and anemia. In severe cases, massive destruction of the colonic mucosa may lead to hemorrhage, perforation, or peritonitis. People with amebic liver abscesses often present with severe right upper quadrant pain, low-grade fever, and weight loss.[82]

Diagnostic methods include microscopic examination of the stool for *E. histolytica*, serum antibody tests, and colonoscopy with specimen collection or biopsy. Treatment includes use of the antimicrobial agents tinidazole and metronidazole, which act against the trophozoites, and diloxanide (not available in the United States), which is effective against the cysts.

Diverticular Disease

Diverticulosis is a condition that commonly occurs on the distal descending and sigmoid colon, in which the mucosal layer of the colon herniates through the muscularis layer.[84] There are often multiple diverticula, most of which occur in the sigmoid colon (Fig. 45.12). Diverticular disease is common in Western society, affecting approximately 40% of the population by age 60 and 60% of the population by age 80.[84] Although the disorder is prevalent in the developed countries of the world, it is almost nonexistent in many African nations and underdeveloped countries. This suggests that factors such as lack of fiber in the diet, a decrease in physical activity, and poor bowel habits (*e.g.,* neglecting the urge to defecate), along with the effects of aging, contribute to the development of the disease.

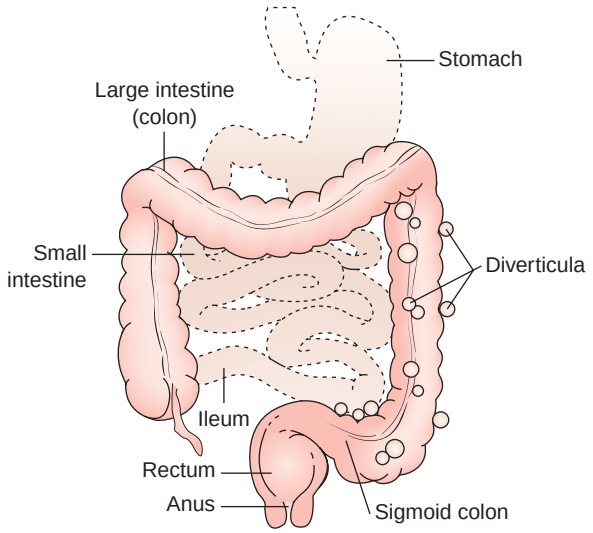


FIGURE 45.12 • Location of diverticula in the sigmoid colon.

In the colon, the longitudinal muscle does not form a continuous layer, as it does in the small bowel. Instead, there are three separate longitudinal bands of muscle called the *teniae coli*. In a manner similar to the small intestine, bands of circular muscle constrict the large intestine. As the circular muscle contracts at each of these points (approximately every 2.5 cm), the lumen of the bowel becomes constricted, so that it is almost occluded. The combined contraction of the circular muscle and the lack of a continuous longitudinal muscle layer cause the intestine to bulge outward into pouches called *haustra*. Diverticula develop between the longitudinal muscle bands of the haustra, in the area where the blood vessels pierce the circular muscle layer to bring blood to the mucosal layer. An increase in intraluminal pressure in the haustra provides the force for creating these herniations. The increase in pressure is thought to be related to the volume of the colonic contents. The scantier the contents, the more vigorous are the contractions and the greater is the pressure in the haustra.

Most people with diverticular disease remain asymptomatic.[86] The disease often is found when x-ray studies are done for other purposes. When symptoms do occur, they often are attributed to IBS or other causes. Ill-defined lower abdominal discomfort, a change in bowel habits (*e.g.,* diarrhea, constipation), bloating, and flatulence are common.

Diverticulitis is a complication of diverticulosis in which there is inflammation and gross or microscopic perforation of the diverticulum. One of the most common complaints of diverticulitis is pain in the lower left quadrant, accompanied by nausea and vomiting, tenderness in the lower left quadrant, a slight fever, and an elevated white blood cell count.[87] These symptoms usually last for several days, unless complications occur, and usually are caused by localized inflammation of the diverticula with perforation and development of a small, localized abscess. Complications include perforation with peritonitis, hemorrhage, and bowel obstruction. Fistulas can form, involving the bladder (*i.e.,* vesicosigmoid fistula) but sometimes involving the skin, perianal area, vagina or small bowel. Pneumaturia (*i.e.,* air in the urine) is a sign of vesicosigmoid fistula.

The diagnosis of diverticular disease is based on history and presenting clinical manifestations. The disease may be confirmed by CT scans or ultrasonographic studies. CT scans are the safest and most cost-effective method.[87] Although barium enema was used in the past, it is no longer recommended because of the risk of extravasation of contrast material if perforation has occurred.[87] Flat abdominal radiographs may be used to detect complications associated with acute diverticulitis.

The usual treatment for diverticular disease is to prevent symptoms and complications. This includes increasing the bulk in the diet and bowel retraining so that the person has at least one bowel movement each day. The increased bulk promotes regular defecation and increases colonic contents and colon diameter, thereby decreasing intraluminal pressure. Acute diverticulitis is treated by withholding solid food and administering a broad-spectrum antibiotic.[87] Hospitalization may be required for people who show significant inflammation, are unable to tolerate oral fluids, are febrile, or have signs and symptoms that suggest systemic involvement.[87] Immunomodulatory agents such as mesalamine and probiotics are two therapies that are becoming more frequently used to manage diverticular disease.[87,88] Surgical treatment is reserved for people experiencing nonresolving symptoms and complications.[84]

Appendicitis

Acute appendicitis is extremely common. In the United States, there is a 12% risk of developing appendicitis for males and a 25% risk for females.[89] The appendix becomes inflamed, swollen, and gangrenous, and it eventually perforates if not treated. Appendicitis is related to intraluminal obstruction with a fecalith (*i.e.,* hard piece of stool), gallstones, tumors, parasites, or lymphatic tissue.

Appendicitis usually has an abrupt onset, with pain referred to the epigastric or periumbilical area. This pain is caused by stretching of the appendix during the early inflammatory process. At approximately the same time that the pain appears, there are one or two episodes of nausea. Initially, the pain is vague, but over a period of 2 to 12 hours, it gradually increases and may become colicky. When the inflammatory process has extended to involve the serosal layer of the appendix and the peritoneum, the pain becomes localized to the lower right quadrant. There may be an elevated white blood cell count but not in all cases.[90] Palpation of the abdomen usually reveals a deep tenderness in the lower right quadrant, which is confined to a small area approximately the size of the fingertip. It usually is located at approximately the site of the inflamed appendix. The person with appendicitis often is able to place his or her finger directly over the tender area. Rebound tenderness, which is pain that occurs when pressure is applied to the area and then released, and spasm of the overlying abdominal muscles are common.

Diagnosis is usually based on history and findings on physical examination. Ultrasonography or CT may be used to confirm the diagnosis in cases where alternative causes of abdominal pain are suspected.[91] Treatment consists of surgical removal of the appendix. Complications include peritonitis, localized periappendiceal abscess formation, and septicemia.

Alterations in Intestinal Motility

The movement of contents through the GI tract is controlled by neurons located in the submucosal and myenteric plexuses of the gut. The axons from the cell bodies in the myenteric plexus innervate the circular and longitudinal smooth muscle layers of the gut. These neurons receive impulses from local receptors located in the mucosal and muscle layers of the gut and extrinsic input from the parasympathetic and sympathetic nervous systems. As a general rule, the parasympathetic nervous system tends to increase the motility of the bowel, whereas sympathetic stimulation tends to slow its activity.

The colon has sphincters at both ends: the ileocecal sphincter, which separates it from the small intestine, and the anal sphincter, which prevents the movement of feces to the outside of the body. The colon acts as a reservoir for fecal material. Normally, approximately 400 mL of water, 55 mEq of sodium, 30 mEq of chloride, and 15 mEq of bicarbonate are absorbed each day in the colon. At the same time, approximately 5 mEq of potassium is secreted into the lumen of the colon. The amount of water and electrolytes that remains in the stool reflects the absorption or secretion that occurs in the colon. The average adult ingesting a typical American diet evacuates approximately 100 to 200 g of stool each day.

KEY POINTS

DISORDERS OF GASTROINTESTINAL MOTILITY

- The luminal contents move down the GI tract as a result of peristaltic movements regulated by a complex interaction of electrical, neural, and hormonal control mechanisms.
- Local irritation and the composition and constituents of GI contents influence motility through the submucosal afferent neurons of the enteric nervous system. GI wall distention, chemical irritants, osmotic gradients, and bacterial toxins exert many of their effects on GI motility through these afferent pathways.

Diarrhea

The usual definition of *diarrhea* is excessively frequent passage of loose or unformed stools. The complaint of diarrhea is a general one and can be related to a number of pathologic and nonpathologic factors. Diarrhea can be acute or chronic and can be caused by infectious organisms, food intolerance, drugs, or intestinal disease. Acute diarrheas that last less than 4 days are predominantly caused by infectious agents and follow a self-limited course.[92]

Acute Diarrhea. Diarrhea that is acute in onset and persists for less than 2 weeks is commonly caused by infectious agents (see previous discussion of infectious enterocolitis). Acute diarrhea is commonly divided into noninflammatory (large-volume) and inflammatory (small-volume) diarrhea, based on the characteristics of the diarrheal stool. Enteric organisms cause diarrhea by several ways. Some are noninvasive and do not cause inflammation, but secrete toxins that stimulate fluid

secretion.[91] Others invade and destroy intestinal epithelial cells, thereby altering fluid transport so that secretory activity continues while absorption activity is halted.[93]

Noninflammatory diarrhea is associated with large-volume watery and nonbloody stools, periumbilical cramps, bloating, and nausea or vomiting. It is commonly caused by toxin-producing bacteria (*e.g., S. aureus*,[94] enterotoxigenic *E. coli*, *Cryptosporidium parvum*, *Vibrio cholerae*) or other agents (*e.g.,* viruses, *Giardia*) that disrupt the normal absorption or secretory process in the small bowel.[95] Prominent vomiting suggests viral enteritis or *S. aureus* food poisoning. Although typically mild, the diarrhea (which originates in the small intestine) can be voluminous and result in dehydration with hypokalemia and metabolic acidosis (*i.e.,* cholera). Because tissue invasion does not occur, leukocytes are not present in the feces.

Inflammatory diarrhea is usually characterized by the presence of fever and bloody diarrhea (dysentery). It is caused by invasion of intestinal cells (*e.g., Shigella*, *Salmonella*, *Yersinia*, and *Campylobacter*) or the toxins associated with the previously described *C. difficile* or *E. coli* O157:H7 infection. Because infections associated with these organisms predominantly affect the colon, the diarrhea is frequent and small in volume[96] and is associated with left lower quadrant cramps, urgency, and tenesmus. Infectious dysentery must be distinguished from acute ulcerative colitis, which may present with bloody diarrhea, fever, and abdominal pain. Diarrhea that persists for 14 days is not attributable to bacterial pathogens (except for *C. difficile*), and the person should be evaluated for chronic diarrhea.

Chronic Diarrhea. Diarrhea is considered to be chronic when the symptoms persist for 4 weeks or greater.[92] Chronic diarrhea is often associated with conditions such as IBD, IBS, malabsorption syndrome, endocrine disorders (hyperthyroidism, diabetic autonomic neuropathy), or radiation colitis. There are four major causes of chronic diarrhea: presence of hyperosmotic luminal contents, increased intestinal secretory processes, inflammatory conditions, and infectious processes[92] (Chart 45.1). Factitious diarrhea is caused by indiscriminate use of laxatives or excessive intake of laxative-type foods.

In *osmotic diarrhea*, water is pulled into the bowel by the hyperosmotic nature of its contents to such a quantity that the colon is unable to reabsorb the excess fluid. It occurs when osmotically active particles are not absorbed. In persons with lactase deficiency, the lactose in milk cannot be broken down and absorbed. Magnesium salts, which are contained in milk of magnesia and many antacids, are poorly absorbed and cause diarrhea when taken in sufficient quantities. Another cause of osmotic diarrhea is decreased transit time, which interferes with absorption. Osmotic diarrhea usually disappears with fasting.

Secretory diarrhea occurs when the secretory processes of the bowel are increased. Secretory diarrhea also occurs when excess bile acids remain in the intestinal contents as they enter the colon. This often happens with disease processes of the ileum because bile salts are absorbed there. It also may occur with bacterial overgrowth in the small bowel, which interferes with bile absorption. Some tumors, such as those of the Zollinger-Ellison syndrome and carcinoid syndrome, produce hormones that cause increased secretory activity of the bowel.[97]

CHART 45.1 CHRONIC DIARRHEA

- Hyperosmotic diarrhea
  - Saline cathartics
  - Lactase deficiency
- Secretory diarrhea
  - Acute infectious diarrhea
  - Failure to absorb bile salts
  - Fat malabsorption
  - Chronic laxative abuse
  - Carcinoid syndrome
  - Zollinger-Ellison syndrome
  - Fecal impaction
- Inflammatory bowel disease
  - Crohn disease
  - Ulcerative colitis
- Infectious disease
  - Shigellosis
  - Salmonellosis
- Irritable colon

Inflammatory diarrhea commonly is associated with acute or chronic inflammation or intrinsic disease of the colon, such as ulcerative colitis or Crohn disease. Inflammatory diarrhea usually is evidenced by frequency and urgency and colicky abdominal pain. It commonly is accompanied by tenesmus (*i.e.,* painful straining at stool), fecal soiling of clothing, and awakening during the night with the urge to defecate.

Chronic parasitic infections may cause chronic diarrhea through a number of mechanisms. Pathogens most commonly associated with chronic diarrhea include the protozoans *Giardia*, *E. histolytica*, and *Cyclospora*. Immunocompromised persons are particularly susceptible to infectious organisms that can cause acute and chronic diarrhea, including *Cryptosporidium*, cytomegalovirus (CMV), and *Mycobacterium avium-intracellulare* complex.

Diagnosis and Treatment. The diagnosis of diarrhea is based on complaints of frequent stools and a history of accompanying factors such as concurrent illnesses, medication use, and exposure to potential intestinal pathogens. Disorders such as IBD and celiac disease should be considered.[92] If the onset of diarrhea is related to travel outside the United States, the possibility of traveler's diarrhea must be considered.

Although most acute forms of diarrhea are self-limited and require no treatment, diarrhea can be particularly serious in infants and small children, persons with other illnesses, elderly persons, and even previously healthy persons if it

continues for any length of time. Thus, the replacement of fluids and electrolytes is considered to be a primary therapeutic goal in the treatment of diarrhea.

Drugs used in the treatment of diarrhea include diphenoxylate (Lomotil) and loperamide (Imodium), which are opium-like drugs. These drugs decrease GI motility and stimulate water and electrolyte absorption. Adsorbents, such as kaolin and pectin, adsorb irritants and toxins from the bowel. These ingredients are included in many over-the-counter antidiarrheal preparations because they adsorb toxins responsible for certain types of diarrhea. Bismuth subsalicylate (Pepto-Bismol) can be used to reduce the frequency of unformed stools and increase stool consistency, particularly in cases of traveler's diarrhea. The drug is thought to inhibit intestinal secretion caused by enterotoxigenic *E. coli* and cholera toxins. Antidiarrheal medications should not be used in persons with bloody diarrhea, high fever, or signs of toxicity because of the risk of worsening the disease. Antibiotics should be reserved for use in persons with identified enteric pathogens.

Acute Diarrheal Disease in Children. Globally, 1.5 million deaths a year are attributed to diarrhea in children under the age of 5 years.[98] Although diarrheal diseases are less prevalent in the United States than in other countries, they place a burden on the health care system. Diarrhea is also the leading cause of malnutrition in children and most frequently affects children under the age of 2 years.[98]

The causes of acute diarrhea in children vary with location, time of year, and population studied. There is increasing recognition of a widening array of enteric pathogens that cause acute diarrhea in children. Viruses are the most common pathogen causing diarrheal illness.[99]

Rotaviruses and noroviruses are the frequently observed pathogens. Other viruses that have been observed in the stools of children include astroviruses and enteric adenoviruses. Many of these pathogens are transmitted easily through food and water or from one person to another. Prevention remains the most vital measure in managing diarrheal disease in children. Important measures to prevent spread of pathogens include proper sanitation methods for food processing and preparation, sanitary water supplies, proper hand hygiene, exclusion of infected people from handling food or providing health care, and exclusion of people with diarrhea from using public recreational water (*i.e.,* swimming pools, ponds, and lakes).

The main objectives in the approach to a child with acute diarrhea are to assess the degree of dehydration, prevent spread of the infection, determine the nature of the etiologic agent, and provide specific therapy as needed. The hydration status of children can be assessed on the basis of easily observed signs and symptoms. Questions about oral intake, frequency and volume of stool output, general appearance and activity of the child, and frequency of urination provide essential information about hydration. Thirst, dry mucous membranes, and decreased skin turgor are common symptoms of dehydration.[99] Data should be obtained about day care attendance, recent travel to a diarrhea-endemic area, use of antimicrobial drugs, and exposure to contaminated water, unwashed fruits or vegetables, or improperly cooked meats because they may indicate the cause of the disorder. Fever is suggestive of an inflammatory process but also occurs with dehydration.[99]

Management of dehydration remains the cornerstone of treatment of children with diarrhea. Infants in particular are more susceptible to dehydration because of their greater surface area, higher metabolic rate, and inability effectively to concentrate their urine. Oral replacement therapy (ORT) is usually the method of choice for infants and children with uncomplicated diarrhea that can be treated at home.

First applied to the treatment of diarrhea in developing countries, ORT can be regarded as a case of reverse technology, in which the protocols originally implemented in these countries have changed health care in industrialized countries as well.[100] Complete ORT solutions contain carbohydrate, sodium, potassium, chloride, and base to replace that lost in the diarrheal stool.[100] Commonly used beverages such as apple juice and cola drinks, which have increased osmolarity because of their high carbohydrate content and low electrolyte content, are not recommended. The effectiveness of ORT is based on the coupled transport of sodium and glucose or other actively transported small organic molecules (see Chapter 44). Bottled ORT solutions are available but can be costly, particularly in cases where large amounts of replacement fluids are needed. The cost can represent a sizable burden for socioeconomically disadvantaged families. Less expensive, premeasured packets and recipes for preparing replacement solutions are available. The use of ORT for treatment of diarrhea in infants and small children is often labor intensive, requiring frequent feeding, sometimes using a spoon or a nasogastric feeding tube.[99] More importantly, the diarrhea does not promptly cease after ORT has been instituted; this can be discouraging for parents and caregivers who desire early results from their efforts. Children who are severely dehydrated with changes in vital signs or mental status require emergency intravenous fluid resuscitation. After initial treatment with intravenous fluids, these children can be given ORT.

Evidence suggests that feeding should be continued during diarrheal illness, particularly in children.[100] It has been shown that unrestricted diets do not worsen the course or symptoms of mild diarrhea and can decrease stool output.[101] Starch and simple proteins are thought to provide cotransport molecules with little osmotic activity, increasing fluid and electrolyte uptake by intestinal cells. The luminal contents associated with early refeeding are also a known growth factor for enterocytes and help facilitate repair after injury. It is recommended that children who require rehydration therapy because of diarrhea be fed an age-appropriate diet. Although there is little agreement on which foods are best, fatty foods and foods high in simple sugars are best avoided. Almost all infants with acute gastroenteritis can tolerate breast-feeding. For formula-fed infants, diluted formula does not provide an advantage over full-strength formula.

Constipation

Constipation can be defined as the infrequent, incomplete, or difficult passage of stools.[92] The difficulty with this definition arises from the many individual variations of function that are normal. What is considered normal for one person (*e.g.,* two or three bowel movements per week) may be considered evidence of constipation by another. Constipation can occur as a primary disorder of intestinal motility, as a side effect of drugs, as a problem associated with another disease condition, or as a symptom of obstructing lesions of the GI tract. Some common causes of constipation are failure to respond to the urge to defecate, inadequate fiber in the diet, inadequate fluid intake, weakness of the abdominal muscles, inactivity and bed rest, pregnancy, and hemorrhoids. The pathophysiology of constipation can be classified into three broad categories: normal-transit constipation, slow-transit constipation, and disorders of defecatory or rectal evacuation. Normal-transit constipation (or functional constipation) is characterized by perceived difficulty in defecation and usually responds to increased fluid and fiber intake.[94] Slow-transit constipation, which is characterized by infrequent bowel movements, is often caused by alterations in the motor function of the colon.[102] Hirschsprung disease is an extreme form of slow-transit constipation in which the ganglion cells in the distal bowel are absent because of a defect that occurred during embryonic development; the bowel narrows at the area that lacks ganglionic cells.[103] Although most persons with this disorder present in infancy or early childhood, some with a relatively short segment of involved colon do not have symptoms until later in life. Defecatory disorders are most commonly due to deficiencies in muscle coordination involving the pelvic floor or anal sphincter.

Diseases associated with chronic constipation include neurologic diseases such as spinal cord injury, Parkinson disease, and multiple sclerosis; endocrine disorders such as hypothyroidism; and obstructive lesions in the GI tract. Drugs such as narcotics, anticholinergic agents, calcium channel blockers, diuretics, calcium (antacids and supplements), iron supplements, and aluminum antacids tend to cause constipation. Older adults with long-standing constipation and straining with defecation may develop dilation of the rectum, colon, or both. This condition allows large amounts of stool to accumulate with little or no sensation. Constipation, in the context of a change in bowel habits, may be a sign of colorectal cancer.

Diagnosis of constipation usually is based on a history of infrequent stools, straining with defecation, the passing of hard and lumpy stools, or the sense of incomplete evacuation with defecation. Rectal examination is used to determine whether fecal impaction, anal stricture, or rectal masses are present. Constipation as a sign of another disease condition should be ruled out. Tests that measure colon transit time and defecatory function are reserved for refractory cases.

The treatment of constipation usually is directed toward relieving the cause. A conscious effort should be made to respond to the defecation urge. A time should be set aside after a meal, when mass movements in the colon are most likely to occur, for a bowel movement. Mimicking a squatting position while sitting on the toilet by elevating the feet may assist in promoting a bowel movement.[102]Adequate fluid intake and bulk in the diet should be encouraged. Moderate exercise is essential, and people on bed rest benefit from passive and active exercises. Laxatives and enemas should be used judiciously. They should not be used on a regular basis to treat simple constipation because they interfere with the defecation reflex and actually may damage the rectal mucosa.

Fecal Impaction

Fecal impaction is the retention of hardened or putty-like stool in the rectum and colon, which interferes with normal passage of feces. If not removed, it can cause partial or complete bowel obstruction. It may occur in any age group but is more common in incapacitated older adults. Fecal impaction may result from painful anorectal disease, tumors, or neurogenic disease; use of constipating antacids or bulk laxatives; a low-residue diet; drug-induced colonic stasis; or prolonged bed rest and debility. In children, a habitual neglect of the urge to defecate in the school setting because of cleanliness of the facilities,[104] modesty, or play interference may promote impaction.

The manifestations may be those of severe constipation, but frequently there is a history of watery diarrhea, fecal soiling, and fecal incontinence.[105] This is caused by increased secretory activity of the bowel, representing the body's attempt to break up the mass so that it can be evacuated. The abdomen may be distended, and there may be blood and mucus in the stool. The fecal mass may compress the urethra, giving rise to urinary incontinence. Fecal impaction should be considered in an elderly or immobilized person who develops watery stools with fecal or urinary incontinence.

Digital examination of the rectum is done to assess for the presence of a fecal mass. The mass may need to be broken up and dislodged manually or with the use of a sigmoidoscope. Oil enemas often are used to soften the mass before removal. The best treatment is prevention.

Intestinal Obstruction

Intestinal obstruction designates an impairment of movement of intestinal contents in a cephalocaudad direction. The causes can be categorized as mechanical or paralytic. Strangulation with necrosis of the bowel may occur and lead to perforation, peritonitis, and sepsis.

Mechanical obstruction can result from a number of conditions, intrinsic or extrinsic, that encroach on the patency of the bowel lumen (Fig. 45.13). Postoperative causes such as external hernia (*i.e.,* inguinal, femoral, or umbilical) and postoperative adhesions are responsible for 75% of intestinal obstruction occurrences.[106] Less common causes are strictures, tumors, foreign bodies, intussusception, and volvulus.

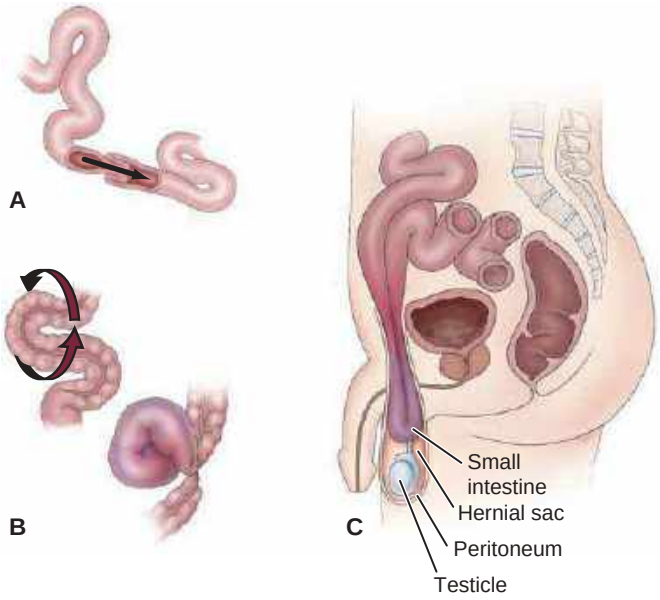


FIGURE 45.13 • Three causes of intestinal obstruction. (**A**) Intussusception with invagination or shortening of the bowel caused by movement of one segment of the bowel into another. (**B**) Volvulus of the sigmoid colon; the twist is counterclockwise in most cases. Note the edematous section of bowel. (**C**) Hernia (inguinal). The sac of the hernia is a continuation of the peritoneum of the abdomen. The hernial contents are intestine, omentum, or other abdominal contents that pass through the hernial opening into the hernial sac. (From Smeltzer S. C., Bare B. G., Hinkle J., et al. (2010) *Brunner and Suddarth's textbook of medical-surgical nursing* (12th ed., p. 1097). Philadelphia, PA: Lippincott Williams & Wilkins.)

FIGURE 45.14 • Intussusception. A cross section through the area of the obstruction shows "telescoped" small intestine surround by dilated small intestine. (From Rubin R., Strayer D. S. (Eds.) (2012). *Rubin's pathophysiology: Clinicopathologic foundations of medicine* (6th ed., p. 645). Philadelphia, PA: Lippincott Williams & Wilkins.)

Remember Ms. Rytel, from the unit opener case study? She is most likely experiencing an obstruction related to the adhesions that occurred due to the necessary repairs required after the motor vehicle collision. At this point, she is monitored closely and receives a nasogastric tube to decompress the upper portion of her GI tract. She is also given nothing by mouth, and her nausea is controlled with intravenous anti-emetics. If her vitals began to deteriorate, she would be prepped to go to the operating room for an open exploration of her abdomen for possible lysis of adhesions.

Intussusception involves the telescoping of bowel into the adjacent segment (Fig. 45.14). It is the most common cause of intestinal obstruction in children younger than 2 years of age.[107] The most common form is intussusception of the terminal ileum into the right colon, but other areas of the bowel may be involved. In most cases, the cause of the disorder is unknown. The condition can also occur in adults when an intraluminal mass or tumor acts as a traction force and pulls the segment along as it telescopes into the distal segment. Volvulus refers to a complete twisting of the bowel on an axis formed by its mesentery (see Fig. 45.13B). It can occur in any portion of the GI tract, but most commonly involves the sigmoid colon (75%), followed by the cecum (22%).[108] Mechanical bowel obstruction may be a simple obstruction, in which there is no alteration in blood flow, or a strangulated obstruction, in which there is impairment of blood flow and necrosis of bowel tissue.

Paralytic, or adynamic, obstruction results from neurogenic or muscular impairment of peristalsis. Paralytic ileus is seen most commonly after abdominal surgery, but it also accompanies inflammatory conditions of the abdomen, intestinal ischemia, pelvic fractures, and back injuries.[109] It occurs early in the course of peritonitis and can result from chemical irritation caused by bile, bacterial toxins, electrolyte imbalances as in hypokalemia, and vascular insufficiency.

The major effects of both types of intestinal obstruction are abdominal distention and loss of fluids and electrolytes (Fig. 45.15). Gases and fluids accumulate in the area; if untreated, the distention resulting from bowel obstruction tends to perpetuate itself by causing atony of the bowel and further distention. Distention is further aggravated by the accumulation of gases. As the process continues, the distention moves proximally (*i.e.,* toward the mouth), involving additional segments of bowel. Either form of obstruction eventually may lead to strangulation (*i.e.,* interruption of blood flow), gangrenous changes, and, ultimately, perforation of the bowel. The increased pressure in the intestine tends to compromise mucosal blood flow, leading to necrosis and movement of blood into the luminal fluids. This promotes rapid growth of bacteria in the obstructed bowel, which has the potential to move into the lymph system and surrounding organs.[108] The movement of the bacteria outside of the digestive tract results in increased inflammation, which can result in further ischemia and organ failure.[108]

The manifestations of intestinal obstruction depend on the degree of obstruction and its duration. With acute obstruction, the onset usually is sudden and dramatic. With chronic conditions, the onset often is more gradual.

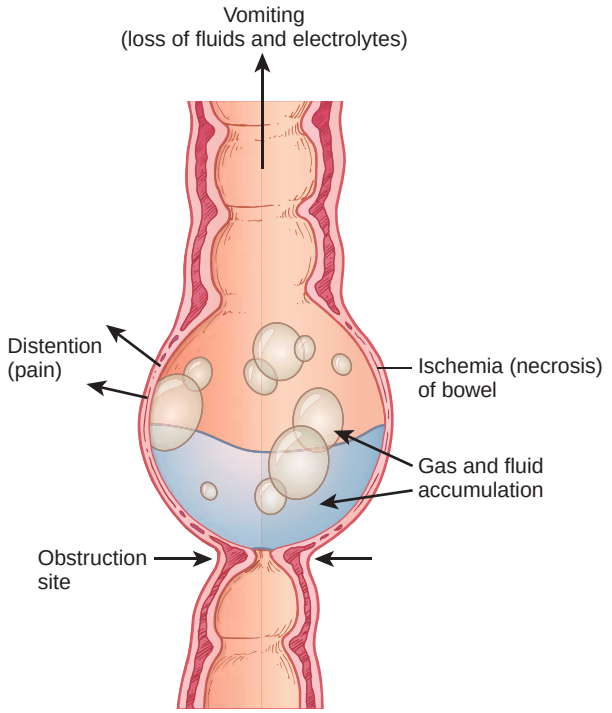


FIGURE 45.15 • Pathophysiology of intestinal obstruction.

The cardinal symptoms of intestinal obstruction are pain, absolute constipation, abdominal distention, sign of fluid volume deficit, and vomiting. With mechanical obstruction, the pain is severe and colicky, in contrast with the continuous pain and silent abdomen of paralytic ileus. There also is borborygmi (*i.e.,* rumbling sounds made by propulsion of gas in the intestine); audible, high-pitched peristalsis; and peristaltic rushes that tend to associate with episodes of abdominal pain.[108] Visible peristalsis may appear along the course of the distended intestine. Extreme restlessness and conscious awareness of intestinal movements are experienced along with weakness, perspiration, and anxiety. Should strangulation of the bowel occur, the symptoms change. The character of the pain shifts from the intermittent colicky pain caused by the hyperperistaltic movements of the intestine to a severe and steady type of pain. Vomiting and fluid and electrolyte disorders occur with both types of obstruction.

Diagnosis of intestinal obstruction usually is based on history and physical findings. Plain film radiography of the abdomen may be used to determine the presence of an obstruction as well as differentiate between partial and complete obstruction by analysis of gas patterns within the bowel.[106] CT scans and ultrasonography may also be used to detect the presence of mechanical obstruction.

Treatment depends on the cause and type of obstruction. Correction of fluid and electrolyte imbalances to baseline levels and measurement of output using a Foley catheter are recommended. Most cases of adynamic obstruction respond to decompression of the bowel through nasogastric suction. Strangulation and complete bowel obstruction require surgical intervention. Intraoperatively, the bowel is observed for return of normal color and peristalsis. If necrotic tissue present, it is resected, and an anastomosis is made.

Peritonitis

Peritonitis is an inflammatory response of the serous membrane that lines the abdominal cavity and covers the visceral organs. It can be caused by bacterial invasion or chemical irritation. Most commonly, enteric bacteria enter the peritoneum because of a defect in the wall of one of the abdominal organs. Causes of peritonitis include perforated peptic ulcer, ruptured appendix, perforated diverticulum, gangrenous bowel, pelvic inflammatory disease, and gangrenous gallbladder. Other environmental causes are abdominal trauma, foreign body ingestion, and infected peritoneal dialysis catheters. Generalized peritonitis, although no longer the overwhelming problem it once was, is still a leading cause of death after abdominal surgery.

The peritoneum has several characteristics that increase its vulnerability to or protect it from the effects of peritonitis. One weakness of the peritoneal cavity is that it is a large, unbroken space that favors the dissemination of contaminants. For the same reason, it has a large surface area that permits rapid absorption of bacterial toxins into the blood. The peritoneum is particularly well adapted for producing an inflammatory response as a means of controlling infection. It tends, for example, to exude a thick, sticky, and fibrinous substance that adheres to other structures, such as the mesentery and omentum, and that seals off the perforated viscus and aids in localizing the process. Localization is enhanced by sympathetic stimulation that limits intestinal motility. Although the diminished or absent peristalsis that occurs tends to give rise to associated problems, it does inhibit the movement of contaminants throughout the peritoneal cavity.

One of the most important manifestations of peritonitis is the translocation of extracellular fluid into the peritoneal cavity (through weeping or serous fluid from the inflamed peritoneum) and into the bowel as a result of bowel obstruction. Nausea and vomiting cause further losses of fluid. The fluid loss may encourage development of hypovolemia and shock. The onset of peritonitis may be acute, as with a ruptured appendix, or it may have a more gradual onset, as occurs in pelvic inflammatory disease. Pain and tenderness are common symptoms. The pain usually is more intense over the inflamed area. The person with peritonitis usually lies still because any movement aggravates the pain. Breathing often is shallow to prevent movement of the abdominal muscles. The abdomen usually is rigid and sometimes described as boardlike because of reflex muscle guarding. Vomiting is common. Fever, an elevated white blood cell count, tachycardia, and hypotension are common. Hiccups may develop because of irritation of the phrenic nerve. Paralytic ileus occurs shortly after the onset of widespread peritonitis and is accompanied by abdominal distention. Peritonitis that progresses and is untreated leads to toxemia and shock.

Treatment. Treatment measures for peritonitis are directed toward preventing the extension of the inflammatory response, intravenously correcting the fluid and electrolyte imbalances that develop, and minimizing the effects of paralytic ileus and abdominal distention. Oral resuscitation is avoided due to the potential need for surgical intervention. Surgical intervention may be needed to address the source of the inflammation, for example, removing an acutely inflamed appendix or closing the opening in a perforated peptic ulcer.

Nasogastric suction, which entails the insertion of a tube placed through the nose into the stomach or intestine, is used to decompress the bowel and relieve the abdominal distention. Fluid and electrolyte replacement is essential. These fluids are prescribed on the basis of frequent blood chemistry determinations. Antibiotics are given to combat infection. Narcotics often are needed for pain relief.

Alterations in Intestinal Absorption

Malabsorption is the failure to transport dietary constituents, such as fats, carbohydrates, proteins, vitamins, and minerals, from the lumen of the intestine into the extracellular fluid compartment for transport to the various parts of the body. It can selectively affect a single component, such as vitamin B_{12} or lactose, or its effects can extend to all the substances absorbed in a specific segment of the intestine. When one segment of the intestine is affected, another may compensate. For example, the ileum may compensate for malabsorption in the proximal small intestine by absorbing substantial amounts of fats, carbohydrates, and amino acids. Similarly, the colon, which normally absorbs water, sodium, chloride, and bicarbonate, can compensate for small intestine malabsorption by absorbing additional end products of bacterial carbohydrate metabolism.

The conditions that impair one or more steps involved in digestion and absorption of nutrients can be divided into three broad categories: intraluminal maldigestion, disorders of transepithelial transport, and lymphatic obstruction. Intraluminal maldigestion involves a defect in processing of nutrients in the intestinal lumen. The most common causes are pancreatic insufficiency, hepatobiliary disease, and intraluminal bacterial growth. Disorders of transepithelial transport are caused by mucosal lesions that impair uptake and transport of available intraluminal nutrients across the mucosal surface of the intestine. They include disorders such as celiac disease and Crohn disease. Lymphatic obstruction interferes with the transport of the products of fat digestion to the systemic circulation after they have been absorbed by the intestinal mucosa. The process can be interrupted by congenital defects, neoplasms, trauma, and selected infectious diseases.

Malabsorption Syndrome

Persons with intestinal malabsorption usually have symptoms directly referable to the GI tract that include diarrhea, flatulence, distention, abdominal pain and/or cramps, and ascites.[110] Weakness, muscle wasting, weight loss, and abdominal distention often are present. Weight loss often occurs despite normal or excessive caloric intake. Steatorrheic stools contain excess fat. The fat content causes bulky, yellow-gray, malodorous stools. In a person consuming a diet containing 80 to 100 g of fat each day, excretion of 7 to 9 g of fat indicates steatorrhea.[110]

Along with loss of fat in the stools, there is failure to absorb the fat-soluble vitamins. This can lead to easy bruising and bleeding (*i.e.,* vitamin K deficiency), bone pain, a predisposition to the development of fractures and tetany (*i.e.,* vitamin D and calcium deficiency), macrocytic anemia,[110] and glossitis (*i.e.,* folic acid deficiency). Neuropathy, atrophy of the skin, and peripheral edema may be present. Table 45.2 describes the signs and symptoms of impaired absorption of dietary constituents.

Celiac Disease

Celiac disease, also known as *celiac sprue* and *gluten-sensitive enteropathy*, is an immune-mediated disorder triggered by ingestion of gluten-containing grains (including wheat, barley, and rye).[112,113] Until recently, celiac disease was considered to be a rare malabsorption syndrome that manifested during early childhood, but today it is known to be one of the most common genetic diseases, with a mean prevalence of 1% to 6% in the general population.[114–116]

The disease results from an inappropriate T-cell–mediated immune response against ingested α-gliadin (a component of gluten protein) in genetically predisposed people. Almost all persons with the disorder share the major histocompatibility complex class II allele HLA-DQ2 or HLA-DQ8.[117] People with the disease have increased levels of antibodies to a variety of antigens, including transglutaminase, endomysium, and gliadin. The resultant immune response produces an intense inflammatory reaction that results in loss of absorptive villi from the small intestine (Fig. 45.16). When the resulting lesions are extensive, they may impair absorption of macronutrients (*i.e.,* proteins, carbohydrates, fats) and micronutrients (*i.e.,* vitamins and minerals). Small bowel involvement is most prominent in the proximal part of the small intestine, where the exposure to gluten is greatest.

There are a number of populations who are at higher risk for celiac disease. These include persons with type 1 diabetes mellitus, other autoimmune endocrinopathies, dermatitis herpetiformis, first- and second-degree relatives of people with celiac disease, and people with Turner syndrome.[118,119] Various malignancies also appear to be a direct result of celiac disease, in that the increased incidence seen in people with celiac disease returns to that of the general population after several years of a gluten-free diet. These malignancies include head and neck squamous cell carcinoma, small intestinal adenocarcinoma, and non-Hodgkin lymphoma.

The classic form of celiac disease presents in infancy and manifests as failure to thrive, diarrhea, muscle wasting, abdominal distention, and, occasionally, severe malnutrition.[120] Beyond infancy, the manifestations tend to be less dramatic. Older children may present with anemia, constitutional short

TABLE 45.2 SITES OF AND REQUIREMENTS FOR ABSORPTION OF DIETARY CONSTITUENTS AND MANIFESTATIONS OF MALABSORPTION

DIETARY CONSTITUENT	SITE OF ABSORPTION	REQUIREMENTS	MANIFESTATIONS
Water and electrolytes	Mainly small bowel	Osmotic gradient	Diarrhea Dehydration Cramps
Fat	Upper jejunum	Pancreatic lipase Bile salts Functioning lymphatic channels	Weight loss Steatorrhea Fat-soluble vitamin deficiency
Carbohydrates			
Starch	Small intestine	Amylase Maltase Isomaltase α-dextrins	Diarrhea Flatulence Abdominal discomfort
Sucrose	Small intestine	Sucrase	
Lactose	Small intestine	Lactase	
Maltose	Small intestine	Maltase	
Fructose	Small intestine		
Protein	Small intestine	Pancreatic enzymes (*e.g.,* trypsin, chymotrypsin, elastin)	Loss of muscle mass Weakness Edema
Vitamins			
A	Upper jejunum	Bile salts	Night blindness Dry eyes Corneal irritation
Folic acid	Duodenum and jejunum	Absorptive; may be impaired by some drugs (*i.e.,* anticonvulsants)	Cheilosis Glossitis Megaloblastic anemia
B_{12}	Ileum	Intrinsic factor	Glossitis Neuropathy Megaloblastic anemia
D	Upper jejunum	Bile salts	Bone pain Fractures Tetany
E	Upper jejunum	Bile salts	Uncertain
K	Upper jejunum	Bile salts	Easy bruising and bleeding
Calcium	Duodenum	Vitamin D and parathyroid hormone	Bone pain Fractures Tetany
Iron	Duodenum and jejunum	Normal pH (hydrochloric acid secretion)	Iron deficiency anemia Glossitis

stature, dental enamel defects, and constipation. In adults, GI symptoms may manifest as diarrhea, constipation, or other symptoms of malabsorption such as bloating, flatus, or belching.

The diagnosis of celiac disease is based on clinical manifestations, supported by serologic tests and confirmed by intestinal biopsy.[112,118] Based on very high sensitivities, the best available tests are the immunoglobulin (Ig) A antihuman tissue transglutaminase (TTG) and IgA endomysial antibody immunofluorescence (EMA) tests.[112] Biopsies of the proximal small bowel are indicated in people with a positive celiac disease antibody test.[118] Usually, additional laboratory tests are done to determine if the disorder has resulted in nutritional disorders such as iron deficiency anemia.

The primary treatment of celiac disease consists of removal of gluten and related proteins from the diet. Gluten is the primary protein in wheat, barley, and rye. Oat products, which are nontoxic, may be contaminated with wheat during processing. Many gluten-free types of bread, cereals, cookies, and other products are available.[113] Meats, vegetables, fruits, and dairy products are free of gluten as long as they are not contaminated during processing. Complete exclusion of dietary gluten generally results in rapid and complete healing of the intestinal mucosa.

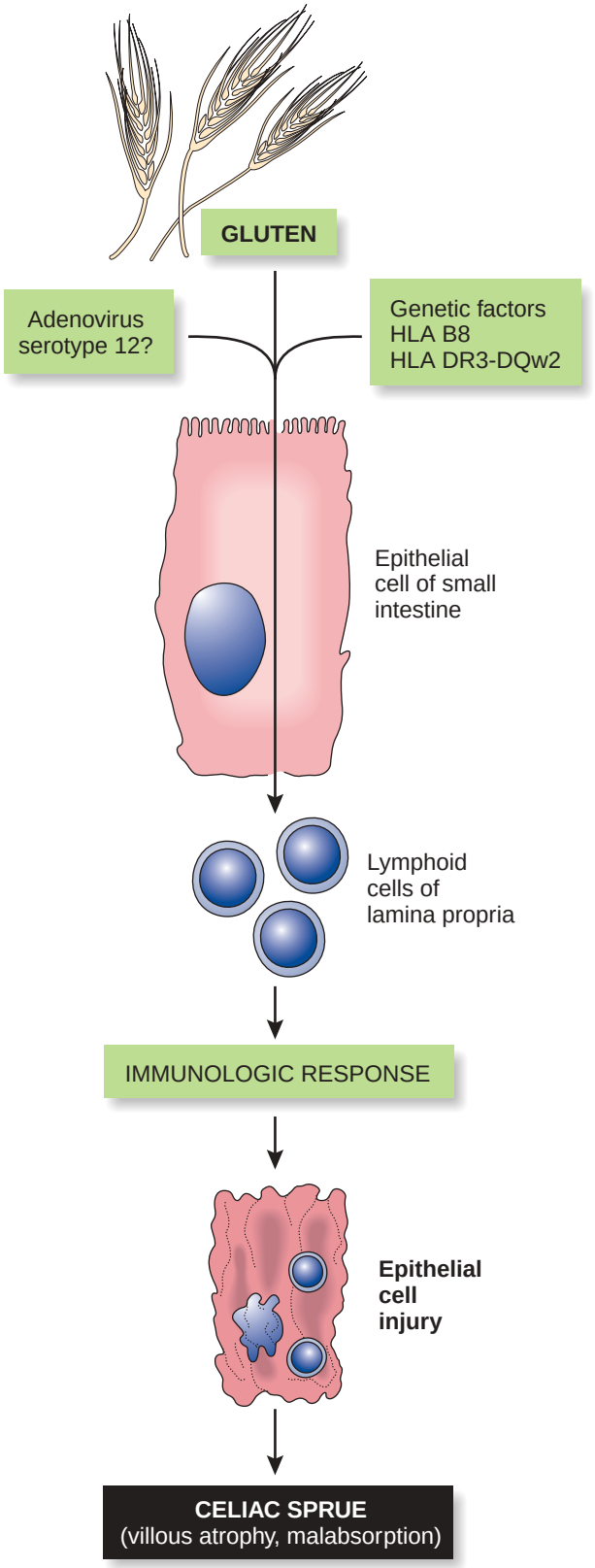


FIGURE 45.16 • Celiac disease (villous atrophy, malabsorption). Hypothetical mechanism in the pathogenesis of celiac disease. HLA, human leukocyte antigen. (From Rubin R., Strayer D. S. (Eds.) (2012). *Rubin's pathophysiology: Clinicopathologic foundations of medicine* (6th ed., p. 642). Philadelphia, PA: Lippincott Williams & Wilkins.)

Neoplasms

Epithelial cell tumors of the intestines are a major cause of morbidity and mortality worldwide. The colon is the site of more primary neoplasms than any other organ in the body.[121] Although the small intestine accounts for approximately 75% of the length of the GI tract, it is an uncommon site of benign or malignant tumors.

Adenomatous Polyps

By far, the most common types of neoplasms of the intestine are adenomatous polyps. A GI polyp can be described as a mass that protrudes into the lumen of the gut.[12] Polyps can be subdivided according to their attachment to the bowel wall (sessile [raised mucosal nodules] or pedunculated [attached by a stalk]), their histopathologic appearance (hyperplastic or adenomatous), and their neoplastic potential (benign or malignant).[12]

Adenomatous polyps (adenomas) are benign neoplasms that arise from the mucosal epithelium of the intestine. They are composed of neoplastic cells that have proliferated in excess of those needed to replace the cells that normally are shed from the mucosal surface (Fig. 45.17). The pathogenesis of adenoma formation involves neoplastic alteration in the replication of the crypt epithelial cells. There may be diminished apoptosis, persistence of cell replication, and failure of cell maturation and differentiation of the cells that migrate to the surface of the crypts.[9] Normally, DNA synthesis ceases as the cells reach the upper two thirds of the crypts, after which they mature, migrate to the surface, and become senescent. They then become apoptotic and are shed from the surface.[12] Adenomas arise from a disruption in this sequence, such that the epithelial cells retain their proliferative ability throughout the entire length of the crypt. Alterations in cell differentiation can lead to dysplasia and progression to the development of invasive carcinoma.

More than half of all adenomatous polyps are located in the rectosigmoid colon and can be detected by rectal examination or sigmoidoscopy.[12] The remainder are evenly distributed throughout the rest of the colon. Adenomas can range in size from a barely visible nodule to a large, sessile mass. They can be classified as tubular, villous, or tubulovillous adenomas.

Tubular adenomas, which constitute approximately 65% of benign large bowel adenomas, typically are smooth-surfaced spheres, usually less than 2 cm in diameter, that are attached to the mucosal surface by a stalk.[12]Although most tubular adenomas display little epithelial dysplasia, approximately 20% show a range of dysplastic changes, from mild nuclear changes to frank invasive carcinoma. *Villous adenomas* constitute 10% of adenomas of the colon.[12] They are found predominantly in the rectosigmoid colon. They typically are broad-based, elevated lesions, with a shaggy, cauliflower-like surface. In contrast to tubular adenomas, villous adenomas are more likely to contain malignant cells. When invasive carcinoma develops, there is no stalk to isolate the tumor and invasion is directly into the wall of the colon. *Tubulovillous adenomas* manifest