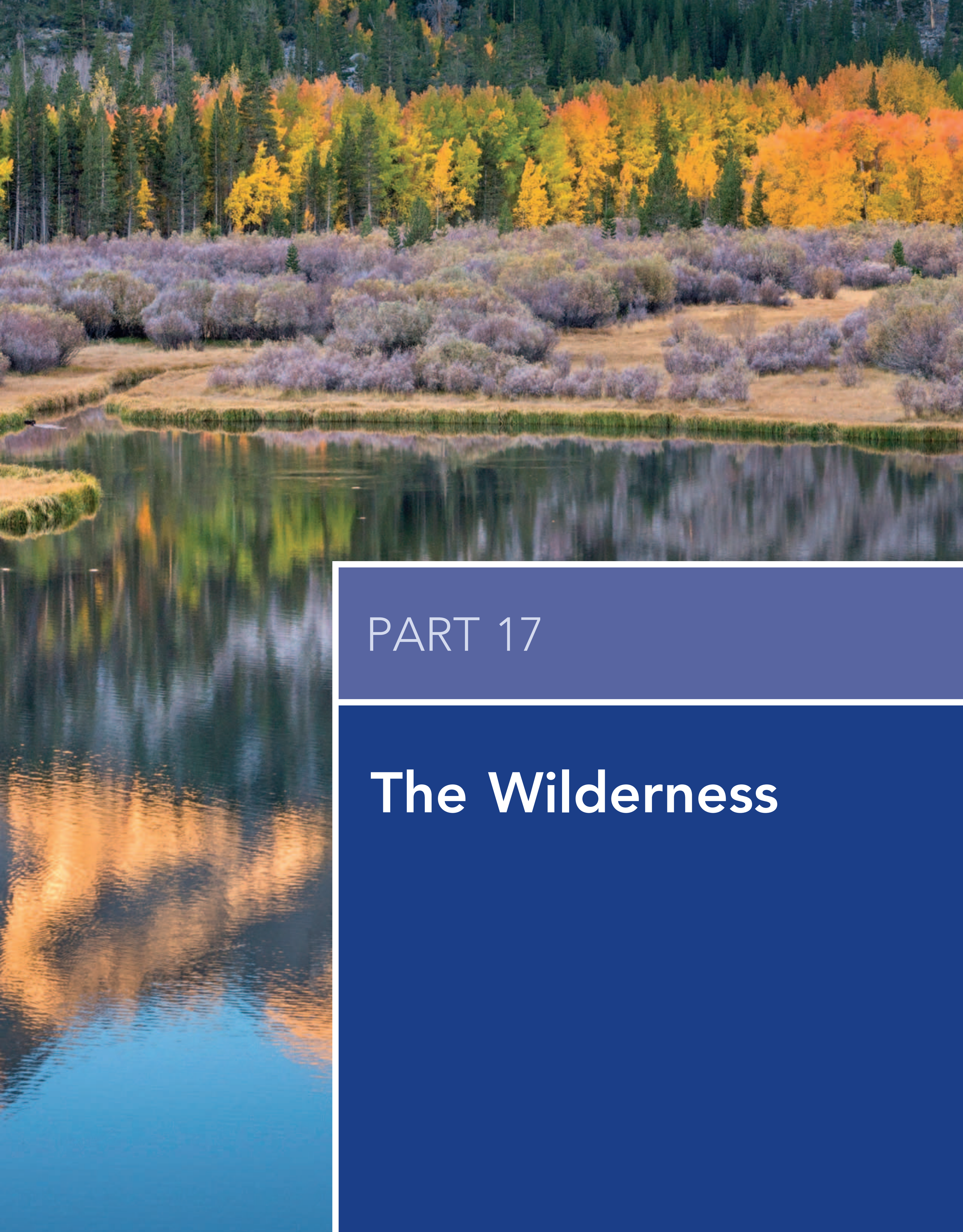


PART 17

The Wilderness

CHAPTER 118

Wilderness Management and Preservation

CHAD P. DAWSON AND JOHN C. HENDEE

The year 2014 marked the 50th anniversary of the Wilderness Act of 1964. Public celebrations of the Wilderness Act were held across the United States to focus on the success of the National Wilderness Preservation System (NWPS). In 2014 the NWPS included more than 796 management units and had 110 million acres of publicly owned lands managed by the four federal land management agencies (Forest Service, National Park Service, Bureau of Land Management, Fish and Wildlife Service).

"Wilderness" is a word that carries various connotations and denotations among different people and cultures. Often, wilderness suggests a special place with human appeal or aversion that invokes an associated emotional, psychological, and mental state because of the natural and undeveloped characteristics of the area, whether real or perceived. Humans have a deep historical and cultural connection with "wild nature" because we have been shaped through human evolution by it, and in turn, we have modified it. Whether a person has a direct experience with wilderness, views wilderness in art and photography, or reads about the adventures of others in wilderness, the human reaction is complex and forms the basis for experiences and both conscious and subconscious memories.

In the United States and several other countries around the world, the term *wilderness* is associated with a legal definition for places that are legislatively protected.[10] The places selected and designated for protection follow certain general guidelines, and this political process is influenced by very strong and diverse groups of stakeholders and the general public. The management objectives for these areas include a variety of values, from preservation of the ecologic conditions and processes to human use and enjoyment (Figure 118-1). Although there is widespread public support for wilderness, there are divergent and polarized viewpoints on how to define wilderness, ranging from extreme protectionists who believe that humans have no place in wilderness to the utilitarian interests that hold that wilderness is a setting for economic development for recreation and tourism activities.

Although there are a variety of definitions, the United States has a legal definition of wilderness, even if somewhat vague, that is the basis for the creation of the NWPS. The purpose of this chapter is to outline the legal designation, management, and preservation of wilderness areas in the United States.[5]

HISTORICAL DEVELOPMENT OF THE WILDERNESS CONCEPT

The term *wilderness* historically was used to describe places that were untamed and not under control of humans, whereas *civilization* was the place of human control.[5,12] Areas of civilization that were cultivated and heavily influenced by human activities often bordered or were surrounded by areas that had minimal human influence. As world population has grown and more land area has come under human influence, wilderness has been lost to the point that it is now scarce in many areas of the world.

There are few places that are not now, or have not been at one time, under human control, habitation, cultivation, or influence. A gradient of human influence and impact exists from urban centers and rural areas to some wild country (e.g., wilderness) that has little or no human influence (Figure 118-2). The so-called human footprint on the world is large and expanding rapidly with population growth, road building, food production, power generation, industrialization, and human habitation. Some identifiable "last of the wild places" exist on each continent and might continue to do so with careful conservation of resources and protection worldwide of some remaining representative or remnant areas of each ecologic community type.[14]

In the United States, older adults have commented on the change in the landscape and extent of development in the country during their lifetime. The early history of the United States during European immigration was one of cultivating and taming the wild places and taking dominion over the land for human habitation. Wilderness was seen as a place for exploration and primitive travel, and most of the population often feared and avoided it. As the amount of land with wild conditions began to diminish, it was more appreciated as a change from cities and civilization. The public's interest in wild places evolved as these areas became scarce. Special places were first set aside as National Parks, such as Yellowstone, Yosemite, and the Grand Tetons. These areas were at first seen as park destinations for development of recreation and tourism, rather than as preserves.

The early interest in wilderness began with employees in the federal land-managing agencies, such as the U.S. Forest Service and U.S. National Park Service. After World War II, a greater public interest began to emerge to save areas for wilderness character. The emerging concern to save certain places by designating them for protection as wilderness was partly the result of interest in recreation experiences, coupled with a growing concern about rapid industrialization and population growth transforming the landscape through human activities. The U.S. population was more than 300 million by 2010.

Some would argue that few places in the world are "wilderness" in the strictest sense of the word. Thus, the more common use of the term *wilderness* is in relation to our perception of areas that are little known or predominantly under the influence of natural processes and forces. Although the term had been commonly applied to any large, remote area with natural characteristics, conditions, and processes, by 1964 it gained a new legal definition that was applied to federally owned land areas designated as wilderness by congressional action in the United States.

WILDERNESS LEGISLATION AND POLICY IN THE UNITED STATES

The U.S. Forest Service and U.S. National Park Service did not begin to set agency policies to protect primitive and roadless areas from development until the 1920s. During the following decades, roadless area inventories and administrative designations of wilderness occurred with increasing public interest. As recreational use and interest in these lands increased, professionals in the agencies and the public raised concerns that the administrative regulation (1) allowed too many development activities, such as mining, grazing, motorized access, and water resource development; (2) shifted boundaries or removed designation to permit resource development; (3) promulgated different regulations and management in different areas; and (4) had neither a distinct policy for wilderness preservation nor a national system with coordinated management.

The concept gradually evolved that legislative protection was needed to create a more permanent and coordinated national

FIGURE 118-1 Day hikers en route to a summit attempt on snow-capped South Sister Mountain (3158 m [10,358 feet]) in the Three Sisters Wilderness, a 286,708-acre area managed by the U.S. Forest Service in west-central Oregon. *(Courtesy Chad P. Dawson.)*

system for wilderness preservation and management. From 1956 to 1964, more than 50 versions of a wilderness bill were introduced in the U.S. Congress, heavily debated, supported, and challenged by different interest groups. Political compromises were deemed necessary to have a wilderness bill finally passed into legislation, and so certain human activities were permitted in some areas, even though they would be nonconforming with the intent of the wilderness legislation.[5,15] These activities included mining, grazing, aircraft landings, and water resources development.

In 1964 the U.S. Congress passed The Wilderness Act (U.S. Public Law 88-577).[18] This legislation created the NWPS and was heralded by the environmental community and the general public as one of the most important pieces of conservation legislation in U.S. history.[15] A historian of wilderness policy and legislation, Scott[16] commented that "before there was a Wilderness Act, wilderness was, at best, an afterthought. Only the U.S. Forest Service had actually delineated wilderness areas, propelled by visionaries within its own ranks."

The Wilderness Act[18] defines a broad statement of policy for designating wilderness under the Act and recognizes the need to set aside significant natural areas for present and future generations because of the rapid loss of such resources, as follows:

> In order to assure that an increasing population, accompanied by expanding settlement and growing mechanization, does not occupy and modify all areas within the United States and its possessions, leaving no lands designated for preservation and protection in their natural condition, it is hereby declared to be the policy of the Congress to secure for the American people of present and future generations the benefits of an enduring resource of wilderness. For this purpose there is hereby established a National Wilderness Preservation System to be composed of federally owned areas designated by Congress as "wilderness areas," and these shall be administered for the use and enjoyment of the American people in such manner as will leave them unimpaired for future use and enjoyment as wilderness, and so as to provide for the protection of these areas, the preservation of their wilderness character, and for the gathering and dissemination of information regarding their use and enjoyment as wilderness. (U.S. Public Law 88-577, section 2a)

FIGURE 118-2 Wildlife is both an attraction and a danger to visitors in the Okefenokee Wilderness, a 353,981-acre wildlife refuge swamp and wilderness area managed by the U.S. Fish and Wildlife Service in southern Georgia. *(Courtesy Chad P. Dawson and Brian Dawson.)*

This paragraph is referred to by land management agencies as the "guiding management intent" because it specifically refers to human "use and enjoyment," provided the areas were "unimpaired" and that management would ensure "preservation of their wilderness character."

Section 2c of The Wilderness Act[18] includes an important and often-quoted definition of wilderness that has led to much controversy and debate because although it is poetic in form, it has left room for interpretation, especially during legal hearings and court cases in the 50 years since its passage:

> A wilderness, in contrast with those areas where man and his own works dominate the landscape, is hereby recognized as an area where the earth and its community of life are untrammeled by man, where man himself is a visitor who does not remain. An area of wilderness is further defined to mean … an area of undeveloped Federal land retaining its primeval character and influence, without permanent improvements or human habitation, which is protected and managed so as to preserve its natural conditions and which (1) generally appears to have been affected primarily by the forces of nature, with the imprint of man's work substantially unnoticeable; (2) has outstanding opportunities for solitude or a primitive and unconfined type of recreation; (3) has at least five thousand acres of land or is of sufficient size as to make practicable its preservation and use in an unimpaired condition; and (4) may also contain ecological, geological, or other features of scientific, educational, scenic, or historical value. (U.S. Public Law 88-577, section 2c)

The definition emphasizes that wilderness is in contrast to civilization, human influences, and habitation. The definition is an ideal that is tempered by four conditions to make it practical and applicable in a world that has historic human activities and impacts that may no longer be noticeable to visitors.

One of the conditions refers to "outstanding opportunities for solitude or a primitive and unconfined type of recreation." This phrase is often referred to as the "guiding principle for recreation and visitor management." A careful reading of the definition of wilderness makes it clear that preservation is the overall principle and reason for designating an area "wilderness." Certain types and amounts of recreation are permitted, provided the area is "protected and managed so as to preserve its natural conditions." This is an important principle to keep in mind when discussing recreation use and management (Figure 118-3), especially when it relates to primitive facilities and trails, backcountry travel, recreational equipment (e.g., removable climbing gear, backpacking stoves), and management interventions.

The use of helicopters and other motorized equipment during search and rescue operations is often regarded as a necessary exception to the natural preservation principle of wilderness, with the humanistic and legal rationale being that human health and safety should take precedence. However, the extent and frequency of search and rescue operations have led to such an intrusion in high-use mountain areas that "outstanding opportunities for solitude" are being diminished. Balancing human interest in risk-taking activities, such as helicopter drop-offs for high-altitude skiing or multiday high-wall rock climbing, with human self-preservation instincts for those injured or needing rescue will remain a controversial subject in wilderness management.

Creation of the NWPS with the Wilderness Act in 1964 was just the beginning of legislative designations. By 2014, more than 170 different laws were passed by the U.S. Congress designating new areas or adding acreage to existing areas.[5] In addition to the Wilderness Act, subsequent legislation over 50 years has clarified congressional intent to protect and manage the wildest remaining U.S. lands as wilderness and has expanded the NWPS. It is difficult to identify any natural resource issue—or any issue—for

FIGURE 118-3 Tourists often test their snow-climbing skills on the easily accessible snowfields above Paradise in the Mt Rainier Wilderness (4393 m [14,410 feet]), a 228,480-acre wilderness area managed by the U.S. National Park Service in western Washington. *(Courtesy Chad P. Dawson.)*

which Congress has so consistently and so often confirmed its intent as it has with wilderness. The initial designation of 9.1 million acres of wilderness was followed by congressional designations in most of the years between 1964 and 2014 to add additional acres and units to the NWPS. The largest single increase was addition of approximately 56 million acres in Alaska under the Alaska National Interest Lands Conservation Act of 1980 (Public Law 96-487).

Many proposals for additional units and acreage are still being brought before the U.S. Congress and its committees. Several authors and organizations have predicted that more acreage will be added to the NWPS in coming decades, but estimates of those additions vary widely.[5,15] Scott[16] observed that, "however much wilderness Americans may choose to designate through their elected representatives, future generations are likely to judge that we preserved too little, rather than too much."

WILDERNESS STEWARDSHIP PHILOSOPHY

Some management of wilderness resources and experiences is necessary as visitor use increases and surrounding land management and use affect the wilderness area. The idea that we need management in an area that was intended to be free of the influences of modern human activities may appear paradoxical. However, wilderness stewardship is the management of human uses of wilderness and internal and external influences on wilderness to protect and preserve an area's solitude and naturalness, including natural processes and conditions. We highlight the following wilderness stewardship philosophy for managers:

> Wilderness management should not mold nature to suit people. Rather, it should manage human use and influences so as not to alter natural processes. Managers should do only what is necessary to meet wilderness objectives and use only the minimum tools, regulations, and enforcement required to meet those objectives. In wilderness, people adapt to nature, to naturalness and solitude, and that is the source of human benefits from wilderness experience as well as the ecological and non-use benefits.

This stewardship philosophy is based on the wilderness legislation and is balanced between two often-debated ends of a continuum. At one end of the continuum is wilderness for its own sake with protection of the naturalness foremost, and at the other end is wilderness primarily for human use and enjoyment. The stewardship philosophy favors natural integrity of the wilderness ecosystems, with some accommodation for primitive styles of recreation and the opportunity for solitude. The long-term results are that the natural forces and processes that shaped and formed the American wilderness will be evident in the wilderness that current stewards leave for future generations.

POTENTIAL THREATS TO WILDERNESS

Designating an area as wilderness is only the first step and must be followed by stewardship to maintain those areas that represent all that is left of many ecosystems, as well as natural landscapes that have not been cultivated, mined, developed, urbanized, or otherwise heavily altered by human activities. Numerous types of internal and external conditions, influences, and changes threaten wilderness resources and values, now and in the future.[9] Three examples of categories of threats are summarized here to highlight the concern about the future sustainability of wilderness conditions and processes.

Wilderness areas in many states are increasingly isolated fragments or remnants of historic ecosystems. As the surrounding landscape becomes more developed and inhabited, wilderness areas become ecologic islands that can continue with various processes, provided they are large enough or are not disconnected from other natural areas. This concern is most pronounced in the eastern United States, with its smaller wilderness areas, but the threat is felt throughout the country as the natural landscape is replaced by human influences, habitation, and manipulation.

Exotic and non-native species of plants and animals are invading wilderness and are direct threats to naturalness and wildness. Efforts to control and manipulate these invasive species can have additional impacts on wilderness conditions that are also undesirable. Invasive plant species, such as knapweed, cheatgrass, and purple loosestrife, can rapidly change an ecosystem and fundamentally alter its historic patterns and conditions for native plant and animal species.

Increasing commercial and public recreation use of wilderness and efforts to control its impacts are serious threats to wilderness resources and values. High visitor use has obvious impacts that can be identified, but what may not be as easily observed are changes to opportunities for solitude and the impacts of regulations and enforcement on social conditions. Regulation takes away some of the freedom of choice and spontaneity to explore that many visitors associate with wilderness experiences.

Nineteen categories of internal and external threats have been identified as change agents that affect wilderness conditions and values[5]:

1. Fragmentation and isolation of wilderness areas as ecologic islands
2. Impacts on threatened and endangered species
3. Increasing commercial and public recreation use
4. Permitted livestock grazing
5. Invasion of exotic and non-native species
6. Administrative access, facilities, and intrusive management
7. Adjacent land management and use
8. Private and public land inholdings within wilderness
9. Established mining claims
10. Wildland fire suppression activities
11. Reduced air quality
12. Reconstruction and maintenance of water projects and reduced water quality
13. Advanced communication and navigation technology that reduces solitude
14. Motorized and mechanical equipment trespass and legal use
15. Aircraft noise and airspace reservations
16. Urbanization and encroaching development
17. Global climate change
18. Legislation designating new wilderness areas with compromised wilderness conditions
19. Lack of political and financial support for wilderness protection and management

The concern is that few of these threats will diminish, and most are projected to increase in the coming decades. Land managers will need to monitor these potential threats to prepare management plans and activities to steward designated wilderness areas and minimize, mitigate, or remove the threats. The responsibilities at the national level for these wilderness

TABLE 118-1 Number of Designated U.S. Wilderness Management Units Under Agency Management and Their Acreage by 2014

U.S. Agency	Wilderness Units	Acres	Percentage of Total
Bureau of Land Management	222	8,736,087	8
Fish and Wildlife Service	71	20,702,488	19
Forest Service	442	36,385,240	33
National Park Service	61	43,932,843	40
TOTAL	796	109,756,658	100

From http://www.wilderness.net.

planning and management activities are primarily under the jurisdiction of four federal agencies.

WILDERNESS MANAGEMENT AGENCIES IN THE UNITED STATES

The NWPS included more than 796 units managed by four federal agencies and totaled 110 million acres of publicly owned lands by 2014 (Table 118-1). The four federal agencies administering the NWPS are the U.S. National Park Service (NPS), U.S. Bureau of Land Management (BLM), U.S. Fish and Wildlife Service (FWS) in the Department of Interior, and U.S. Forest Service (FS) in the Department of Agriculture. The NPS has the greatest total area of wilderness at 43.9 million acres and the fewest units for a federal agency; the largest area in the NWPS, Wrangell–St. Elias Wilderness (>9 million acres) in Alaska, is under NPS management. The FS has the largest number of wilderness units to manage and approximately one-third of the total NWPS acreage. The FWS manages 19% of the NWPS area, including the smallest unit in the system, 5-acre Pelican Island Wilderness in Florida. The BLM has significant acreage in some of the desert ecosystems of the west and manages many smaller units in its 8% of the NWPS.

The four federal land-managing agencies have promulgated regulations based on the wilderness legislation and have developed policy and management documents to steward the lands under their jurisdiction. In addition, the process of evaluating additional lands and managing them for potential inclusion in the NWPS continues for all four agencies. Although the NWPS is a national system operating under the same legislation, each agency has developed its own procedures and organizational approach to accomplishing the task of protecting the "enduring resource of wilderness," based on each agency's administrative mission and structure. Some of the different approaches to visitor and resource management can be confusing to visitors who do not have appreciation or understanding that the overall mission of each agency is different. For example, the FWS has a wildlife management mission that incorporates a national wildlife refuge system.

DISTRIBUTION OF WILDERNESS IN THE UNITED STATES

The 110-million-acre NWPS represents just over 4.5% of the U.S. land area, in contrast to the more than 6% of total acreage in urban and suburban land area or more than 20% of the total in agricultural cropland.[9] The NWPS is an attempt to designate wilderness areas that would represent the different geographic regions and ecosystems of the United States. The representation of ecosystems is not complete (<50% of types are represented), and it is more effective for some arid lands and mountain ecosystems of the west than for coastal lowlands, grasslands, and eastern hardwood forests.[9]

Forty-four of the states have federally designated wilderness, ranging from a 77-acre island wilderness area in Ohio to more than 57 million acres in the state of Alaska. The six states without federally designated wilderness are in the midwestern or northeastern United States. When the number of acres of designated wilderness in each state or region is compared with the total land area and total population, the pattern is an uneven distribution favoring the western United States.[5] Less than 5% of the NWPS is located in the eastern United States, where more than one-half of the population resides on over 40% of the U.S. land area. The Pacific and mountain regions of the western United States have approximately 22% of the population and more than 95% of the NWPS. The greatest disparity is that Alaska has more than one-half of the land area of the NWPS and less than 1% of the U.S. population.

Whereas the NWPS is based on the Wilderness Act of 1964 and federal land ownership, there are also 12 states that have designated state wilderness programs or areas on state-owned lands since the 1970s.[13] Most notable are 22 New York State wilderness areas (1.2 million acres in the Adirondack and Catskill Forest Preserves) and five Alaska state wilderness areas (1.1 million acres). The 12 states with wilderness programs or areas protect more than 3.2 million acres.[13] These areas are managed by the state land-managing agencies and are not part of the NWPS. Many states have legislation and management programs modeled after the federal Wilderness Act.

The NWPS is extensive and complex geographically and is often difficult for visitors to locate because it is generally shown on agency maps only as part of overall public land holdings. For example, the Boundary Waters Canoe Area Wilderness in Minnesota is part of the Superior National Forest. A helpful source to see the geographic distribution and location of the units in the NWPS is available at http://www.wilderness.net and is provided by the Wilderness Institute at the University of Montana's College of Forestry and Conservation, Arthur Carhart National Wilderness Training Center, and Aldo Leopold Wilderness Research Institute. This website also links to the managing federal agency, the wilderness area's legislative history, and visitor information sources.

WILDERNESS VALUES AND PUBLIC PERCEPTIONS

The American public is strongly supportive of wilderness designation and the NWPS.[1,3,4] Scott[16] reported a summary of seven different surveys in the United States from 1999 through 2002 that showed 48% to 81% of respondents supported designating more wilderness land in the United States into the NWPS. Our observations on the relationship between public visitors to wilderness areas and the values they hold for its protection indicate that the striking trend over time is for a broad base of support across American society and intense commitment from a subset to fight for wilderness protection[5]:

> Although wilderness means something different to everyone, four central themes have consistently emerged: experiential, the direct value of the wilderness experience; the value of wilderness as a scientific resource and environmental baseline; the symbolic and spiritual values of wilderness to the nation and the world; and the value of wilderness as a commodity or place that generates direct and indirect economic benefits.

National surveys in 1994 and 2000 reported that more than 50% of the public indicated that 12 of 13 wilderness values were very or extremely important to them[4] (Table 118-2). These surveys showed a trend to increase the percentage for all 13 values, indicating a higher level of value.[3,4] The nonuse values tended to dominate the higher average value scores (scale ranged from "not important" to "extremely important"), with the highest value for protecting water quality and air quality. Strong support for the value of income from tourism related to industry use on wilderness lands tended to be reported by less than 30% of the respondents. Western U.S. residents were somewhat more often aware of the NWPS than were easterners (60% vs. 56%); conversely, eastern more often than western residents (53% vs. 48%) reported there was not enough land in the NWPS. Metropolitan

TABLE 118-2 Percentage of Americans (>16 Years of Age) Indicating a Response of "Very or Extremely Important" for 13 Wilderness Values

Wilderness Value	Percentage
Protecting water quality	93.1
Protecting air quality	92.3
Protection of wildlife habitat	87.8
For future generations	87.0
Protection for endangered species	82.7
Preserving ecosystems	80.0
Future option to visit	75.1
Just knowing it exists	74.6
Scenic beauty	74.0
Recreational opportunities	64.9
For scientific study	57.5
Providing spiritual inspiration	56.5
Income for tourism industry	29.7

From Cordell HK, Tarrant MA, Green GT: Is the public viewpoint of wilderness shifting? *Int J Wilderness* 9:27, 2003, with permission.

and urban more often than rural residents (54% vs. 44%) reported there was not enough land in the NWPS.[4]

WILDERNESS VISITORS

The diversity of wilderness visitors ranges from those who take short walks and view scenery and wildlife in an hour to multiday backpackers, week-long backcountry hunters with pack animals, and mountain climbers on expeditions. The growth in recreation demand and increasing popularity of many forms of recreation in wilderness are the result of U.S. population growth and a general upward trend in participation in many activities, on all types of sites, in the United States[1] on public or private lands. Although backcountry and wilderness use is distributed across the full geographic and sociodemographic spectrum, an identifiable 8.6% of the U.S. population has been labeled as "backcountry actives" by one researcher because of their 2.5 times or greater above-average participation in such activities as backpacking, wilderness visits, cross-country skiing, and day hiking.[1]

Studies at high-use federal agency sites of visitor use of wilderness in the mid-1990s estimated that more than 14 million visitors went to the wilderness per year. Public surveys of the general population in recent years estimated that visitation was closer to 40 million per year.[2] Most estimates of future growth suggest that wilderness use will continue to increase 2% to 4% per year, and that participation will continue across a wide range of activities.[1,2]

Recreation enthusiasts spending 7 or more days in wilderness or primitive areas per year are a growing market segment that includes participation in strenuous physical activities on a regular basis.[1] Improvements in and availability of high-technology gear permit travel in all types of weather conditions and terrain, so that regular use across the entire landscape makes it more difficult to manage use in general, as well as more challenging to conduct search and rescue operations. Risk-taking, exploring, and adventure activities are increasingly prevalent because of media exposure, easier access to sites, more available high-technology gear, more opportunities for initiation into activities, and more training and skill-building opportunities. There is some concern that communication equipment such as satellite cell phones and navigational equipment, such as handheld Global Positioning System (GPS) units, may contribute to the impression that a person could call for help more readily and therefore take greater risks than their skills otherwise would warrant.

One of the ways that people are initiated into wilderness use and gain training in primitive travel and living is through participation in Wilderness Experience Programs (WEPs). The number of WEP organizations and their clientele have grown rapidly in the United States in the past two decades.[7] Wilderness land managers recognize WEPs as a significant user group with a focus on programs in wilderness and primitive areas as one of their defining characteristics.[6]

The most prevalent three types of WEPs are (1) educational programs where the wilderness ecosystem is the focus of instruction, research, and field trips; (2) personal growth and development, where wilderness is the setting and metaphor for everyday life, with insights achieved from challenging activities and reflection; and (3) therapy and healing, where wilderness is the setting to seek restored normal functioning and a healthy balance through primitive living and traveling.[6,7] Although the healing aspects of a natural environment[8] and wilderness areas are well documented, the inherent risks of traveling in remote areas must be recognized through risk management assessment and contingency planning,[17] especially for human health and safety.

DISTRIBUTION OF WILDERNESS VISITOR USE

Visitor use is very unevenly distributed in geographic space and across time. Each wilderness has popular access points because of easy highway access from urban centers, information in guidebooks, word-of-mouth information dissemination, and many other factors. There is often congestion in visitor use at these access sites and crowding along popular trails and at campsites and destinations (e.g., lakes and ponds, scenic overlooks, historic sites). Many wilderness areas have these types of heavy-use access sites; conversely, each area also has places with very low or no use for a variety of reasons. This variation in visitation means that not every acre of wilderness or the associated experience has the potential to be experientially unique. Therefore, a continuum of solitude and naturalness exists across wilderness.

Some of the extreme variations in visitation are caused by seasonality and opportunity for use. Each area has a favored season for a given type of activity. Even though most hiking and camping activities occur during late spring through early fall, some areas have other peak-use times because of weather and opportunities present. For example, spring fishing and whitewater boating may occur in the same area, whereas fall hunting and backpacking during fall foliage may be followed by cross-country skiing and winter camping. Cooler weather in desert areas may bring visitors in nonsummer months (Figure 118-4). Warmer months in alpine areas may bring large numbers of visitors. Weekend and weekday variations are normal fluctuations resulting from work schedules.

Other variation patterns in visitation are caused by geographic location in the United States and proximity to urban population centers. The majority of use in each wilderness area comes from

FIGURE 118-4 The lower Sonoran Desert in the North Maricopa Mountain Wilderness is a scenic and difficult place to access. Visitors may not understand the dangers of heat and dehydration in this 63,020-acre wilderness area managed by the U.S. Bureau of Land Management in southern Arizona, where temperatures commonly exceed 100° F (37.7° C). *(Courtesy Chad P. Dawson.)*

the surrounding region and states, regardless of whether one considers the urban-proximate White Mountain National Forest wilderness in New Hampshire or the Bob Marshall Wilderness in Montana. However, there is considerable long-distance travel and higher visitation to some of the larger and more popular wilderness areas that have unique opportunities and features. Every wilderness area has a great variation in the number of people at any given time and place. Although such variation in visitation is natural, the impacts may be difficult to manage, leading managers to use more direct and heavier-handed management techniques to reduce the negative consequences to resources and other users.

WILDERNESS MANAGEMENT PRINCIPLES

The guiding principles for managing wilderness areas[5] revolve around the idea that wilderness needs to be managed as a pristine extreme in the landscape (>4% of U.S. land area) to maintain the distinctive qualities that define and separate wilderness from other land uses (>95% of U.S. land area). Wilderness is managed from the biologically centered perspective. Environmental integrity and primeval conditions of wilderness are the basis for any human enjoyment, values, and benefits.

Managing wilderness as an *ecosystem* and not as a separate set of resource types (e.g., water, forests, wildlife) focuses managers on a more comprehensive perspective on the protected area. Most wilderness areas represent the remnants of ecosystems or entire ecosystems and, as such, need to be protected for present and future generations if they are to be available for humans to experience and enjoy. In addition, it is imperative that human uses and influences be managed to preserve wilderness conditions and characteristics, because without such stewardship, these remaining areas would lose their unique value in the U.S. landscape.

If wilderness is to be managed to maintain or improve wilderness conditions and not allow degradation on sites or across the area, it is essential to understand the carrying capacity of the area to sustain recreational use. One of the major components in managing wilderness recreational use is to manage use in favor of recreation and human activities that depend on wilderness conditions to achieve their goals, while not degrading wilderness conditions. In other words, there are other places to have various recreational experiences that do not require wilderness conditions. Thus, only those activities that require such conditions should be allowed in wilderness, and only as much as the area can sustain while maintaining its wilderness conditions and processes.

One of the implications of such a management philosophy is that wilderness is not primarily a place for recreational use, although it is permitted as long as it does not impinge on the capacity of the area to maintain its wilderness conditions and processes. All management activities, including search and rescue operations, should have as light an impact on the wilderness and user experience as possible. This is sometimes referred to as the "minimum tool or regulation" that achieves management objectives for the area and maintains the highest levels of naturalness and solitude. Examples are using hand tools instead of gas-powered tools in wilderness maintenance activities, using educational materials in place of direct trip management, and using minimal directional trail signs and not mileage markers.

Section 4c of The Wilderness Act provides for the administrative use of vehicles, such as aircraft and motorized equipment, when it is considered the minimum necessary for certain operations such as search and rescue (SAR). In the case of emergency situations, the use of aircraft, helicopters, and other motorized transport during SAR operations is often thought of as necessary for human health and safety. In some places, however, the extent and frequency of SAR operations have become an intrusion affecting both the environment (e.g., wildlife behavior) and the human experience of solitude. Balancing the need for effective and rapid SAR operations with impacts on the wilderness resources and experiences will be an increasingly controversial subject in wilderness management. Determining the level of SAR operations and activities that is effective with a minimal impact on wilderness resources and experiences is an ideal goal that needs further attention and study. For example, in nonemergency SAR operations, the minimum may be determined to be nonmotorized access using humans on foot to search or to carry out a victim with minor trauma.

WILDERNESS PRESERVATION AS A NATIONAL AND INTERNATIONAL MOVEMENT

The number of and membership in organizations involved in promoting wilderness designations, stewardship, information, and education have grown dramatically over the last 50 years. Organizations can be found at international, national, state, and local levels.

The U.S. legislative model has influenced some forms of international wilderness protection,[11] although variation in level and type of protection is complex and based on the cultural and legislative history in each country. The concept of wilderness is universal, but the national legislative approach used in the United States has been widely adopted by other countries, such as Canada, Australia, Finland, Russia, and South Africa.[11] There are dozens of international organizations that one can join to help protect wilderness. Some examples include the WILD Foundation (http://www.wild.org), Conservation International (http://www.conservation.org), and IUCN—The World Conservation Union (http://www.iucn.org). Many countries have strong public support for wilderness and related organizations that promote wilderness designation and stewardship.

As previously discussed, wilderness protection and management of federal lands in the United States is under jurisdiction of the NPS, BLM, FS, and WFS, each of which has policy and operational information that is important for visitors to understand. This information is available at agency websites but is easy to obtain through websites (e.g., http://www.wilderness.net) that allow access from a geographic map that in turn brings the viewer to useful information, ranging from the designating legislation to the managing agency and local-level contact offices for that wilderness unit. Obtaining local contact information and some general visitor management information makes wilderness visitation more enjoyable and improves the pretrip planning process, including compliance with local regulations.

Wilderness stewardship organizations in the United States include many that have been involved in wilderness issues for decades. Only a few examples are listed here: The Wilderness Society (http://www.wilderness.org), Sierra Club (http://www.sierraclub.org), American Wilderness Coalition (http://www.americanwilderness.org), Campaign for America's Wilderness (http://www.leaveitwild.org), Wilderness Watch (http://www.wildernesswatch.org), National Wildlife Federation (http://www.nwf.org), National Audubon Society (http://www.audubon.org), and the Izaak Walton League (http://www.iwla.org).

Wilderness preservation is a national and international movement comprising grass roots and membership organizations interested in protection and stewardship of dwindling wild areas. Although the concept and values of wilderness are supported by the general population of the United States and many other countries, it is the continued support and work of many people and organizations that stimulate the legislative and administrative branches of government to continue their efforts to maintain parts of the United States "wild" for present and future generations. The reader is encouraged to become part of that wilderness preservation movement. For a comprehensive resource, the reader is referred to Dawson and Hendee.[5]

REFERENCES

> ***Complete references used in this text are available online at expertconsult.inkling.com.***

REFERENCES

1. Cordell HK. Outdoor recreation for 21st century America: A report to the nation: The National Survey on Recreation and the Environment. State College, Pa: Venture Publishing; 2004.
2. Cordell HK. Outdoor recreation in American life: A national assessment of demand and supply trends. Champagne, Ill: Sagamore Publishing; 1999.
3. Cordell HK, Bergstrom JC, Bowker JM. The multiple values of wilderness. State College, Pa: Venture Publishing; 2005.
4. Cordell HK, Tarrant MA, Green GT. Is the public viewpoint of wilderness shifting? Int J Wilderness 2003;9:27.
5. Dawson CP, Hendee JC. Wilderness management: Stewardship and protection of resources and values. 4th ed. Golden, Colo: Fulcrum Publishing; 2009.
6. Dawson CP, Tangen-Foster J, Friese GT, et al. Defining characteristics of U.S.A. wilderness experience programs. Int J Wilderness 1998;4:22.
7. Friese GT, Hendee JC, Kinziger ML. Wilderness experience program industry in the United States: Characteristics and dynamics. J Exper Educ 1998;21:40.
8. Frumkin H. Beyond toxicity: Human health and the natural environment. Am J Prev Med 2001;20:234.
9. Hendee JC, Dawson CP. Wilderness progress after forty years under the U.S. Wilderness Act. Int J Wilderness 2004;10:4.
10. Kormos CF. A handbook on international wilderness law and policy. Golden, Colo: Fulcrum Publishing; 2008.
11. Martin VG, Watson A. International wilderness. In: Dawson CP, Hendee JC, editors. Wilderness management: Stewardship and protection of resources and values. 4th ed. Golden, Colo: Fulcrum Publishing; 2009. p. 50–88.
12. Nash RF. Wilderness and the American mind. 4th ed. New Haven, Conn: Yale University Press; 2001.
13. Propst BM, Dawson CP. State-designated wilderness in the United States: A national review. Int J Wilderness 2008;14:19.
14. Sanderson EW, Jaiteh M, Levy MA, et al. The human footprint and the last of the wild. Bioscience 2002;52:891.
15. Scott DW. The enduring wilderness: Protecting our national heritage through the Wilderness Act. Golden, Colo: Fulcrum Publishing; 2004.
16. Scott DW. The Wilderness Act at forty: Looking back, looking ahead. Int J Wilderness 2004;10:8.
17. Tangen-Foster J, Dawson CP. Risk management programs in wilderness experience programs. Int J Wilderness 1999;5:29.
18. US Public Law 88-577, The Wilderness Act of September 3, 1964, 78 Stat 890.

CHAPTER 119

The Changing Environment

JAMES K. MITCHELL, JAMES M. JEFFERS, AND MICHAEL B. BRADY

Three miles southwest of the Kremlin, not far from Moscow's Olympic Stadium, lies the Novodevichiy Cemetery, one of the most celebrated burying places in Russia. Amid ornate memorials to former leaders such as Gromyko* and Khrushchev,† who were out of favor at the times of their deaths, and to giants of the arts, such as Chekhov‡ and Shostakovich,§ stands a white marble pedestal carrying the sculpted head of Vladimir Illich Vernadsky (1863-1945). Little known in the West, Vernadsky was a prescient observer of the emerging role of humans as makers of the global environment. It was he who first announced that we are living at a time when the power of mankind to change Earth now rivals that of geologic processes.[5] In the past, students of natural history could regard the human life span as a mere blink of a cosmic eye that witnessed little environmental change. At present, we are faced with the prospect that the planet may be fundamentally transformed by humans, perhaps within a few decades, but more probably over the next one or two generations.

This situation has not come about all at once, or equally everywhere. On a global scale, it has gradually built up over centuries, although the local manifestations of increased human agency sometimes have been masked by other processes. For example, conversion of "natural" ecosystems to "managed" ecosystems is a dominant feature on the global scale, but in some parts of the world, especially the United States and Western Europe, managed ecosystems are also being abandoned.

Such is the case in New Jersey, where hilltop 19th-century farmlands have largely reverted to regrowth forests.[21] The complications of environmental change might best be appreciated with the aid of a time machine, such as the one envisioned by H. G. Wells in 1895. Imagine, for a moment, being in a mature pine forest in southern New England. What might be observed as the machine slips into the past at this location? The surrounding landscape comes clearly into view. The pine trees shrink slowly down into youth as the years wind back, because most of today's pines trace their origin to abandonment of farmlands at or near the turn of the century. The pines disappear entirely in the late 19th century and are replaced by shrubs and eventually by grasses. By the mid-1800s, the local vicinity is completely open and appears as a shifting mosaic of agricultural crops and pasture, rotated in time and space. This is the high tide of farming in New England. Thereafter, the sequence goes into reverse. By the 18th century, trees begin to return, connecting the remnant patches of presettlement vegetation. Gradually, the forest closes in, and traces of human presence fade. Little breaks the monotony, apart from occasional fires started by lightning or native Americans clearing seasonal cultivation patches, or major windstorms that topple weaker trees. As the 16th century approaches, the landscape is essentially similar from decade to decade.

The time traveler's dominant impression is one of change. The preceding sequence of changes has been documented by many analysts of the New England landscape.[25] The sequence might be different elsewhere but is no less dynamic. Sometimes the changes are sudden and dramatic, and sometimes they are slow and imperceptible. Sometimes they are "natural" (e.g., storms), and sometimes they are caused by humans (e.g., forest clearance). For the bulk of human history, natural changes have seemed to dominate, although in fact people have been major shapers of the environment for millennia.[13,47,76] A casual observer of the New England landscape might conclude that the well-wooded 21st century scene is more "natural" than the cleared fields of the 1850s. However, today's scene is just as much a product of human choices as that of the 19th century, although different in composition and appearance.

In any event, deciding whether human or natural factors are responsible for a given environmental change is often difficult; these factors operate interdependently (Figure 119-1). It is widely believed that people have reached a critical threshold as environmental modifiers; they are able to equal or surpass the effects of nature. Humans have already significantly modified about half the Earth's land surface; portentous human-forced changes are becoming manifest in the entire biosphere. We can now speak of a human "transformation" of the global environment.[19,31,77]

This chapter addresses environmental change and its human dimensions, with special attention to implications created by the environment for the wilderness and wilderness medicine. What types of changes are likely to occur? How will they affect the natural environment, especially wilderness areas? What will be the consequences for society in general and for medical practitioners in particular? Can anything be done to improve our chances of successfully negotiating this impending time of dislocation and discontinuity?[48]

ISSUES OF ENVIRONMENTAL CHANGE

In recent years, a number of environmental change issues have come to prominence. These include climate change, stratospheric ozone depletion, erosion of biodiversity, population growth, and burgeoning pollution. These issues affect all environments, from urban centers to remote wilderness areas, and are examined on a variety of scales in this discussion. Although each issue is characterized by different expressions of change, all are interconnected. Local changes can aggregate to produce global effects, and global changes have many different disaggregated local effects.[77] In recognition of increasing human prominence, the present era of Earth's history is becoming known as the *Anthropocene*.[15,19] Current investigations of human-forced biogeochemical systems increasingly go beyond the separate issues previously noted, to adopt an integrated (holistic) perspective that spans the natural and social sciences.[25,31]

CLIMATE CHANGE

Weather is the state of the atmosphere at any specific time. *Climate* is the average weather pattern at a particular location. Weather and climate are usually described by such measures as temperature, precipitation, pressure, humidity, and wind speed and direction. In most parts of the world, these measures have been recorded for less than a century, so the actual historical record of direct observations is relatively brief compared with the human tenure of Earth. However, scientists are often able to extend the historical record by constructing synthetic climate data from other evidence, such as tree rings, fossils, concentrations of plankton in ocean sediments, pollen in sedimentary rocks, and isotopes of carbon and oxygen in rocks and glacial ice. For

---

*Andrei Gromyko (1909-1989); long time Foreign Minister of the USSR (1957-1985).

†Nikita Khrushchev (1894-1971); First Secretary Communist Party of the Soviet Union, 1953-1964.

‡Anton Pavlovich Chekhov (1860-1904); Russian novelist, short-story writer, essayist, and memoirist.

§Dmitri Dmitriyevich Shostakovich (1906-1975); Russian composer and one of the most celebrated composers of the 20th century.

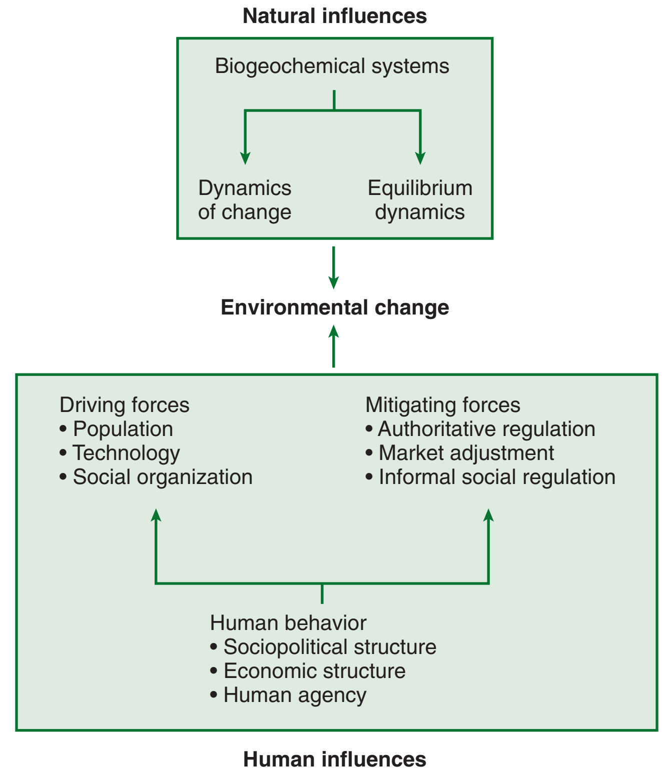


FIGURE 119-1 Human and natural forces for environmental change. *(Modified from Kates RW, Turner BL II, Clark WC: The great transformation. In Turner BL II, editor:* The earth as transformed by human action, *Cambridge, UK, 1990, Cambridge University Press.)*

example, narrow intervals between annual growth rings in trees and thin layers of organic material in lake sediments usually indicate cold, dry conditions. Clues such as these permit investigators to open a window on past climates.

Figure 119-2 illustrates trends in average global temperature during the past 10,000 years. Note that the global temperature has been in flux throughout this period. Not only has weather varied in relation to long-term average conditions, but the averages themselves have changed over time. For example, during the most recent ice age (about 10,000 years ago), average global temperatures were approximately 6°C (42.8°F) cooler than at present.[29] In other words, a massive environmental change (the Wisconsin ice age) was connected with a relatively small climatic change. It is worthwhile remembering that regional changes in climate may or may not parallel global changes. For example, between 2600 and 2700 years ago—around the time of Socrates and Confucius—North America was colder and wetter than was the continental average since the end of the Wisconsin glaciation, whereas conditions in Europe were warmer and drier. Fortunately, the climate has remained within a range that sustains life for most of Earth's history, and the changes have occurred at very slow rates over thousands to millions of years.

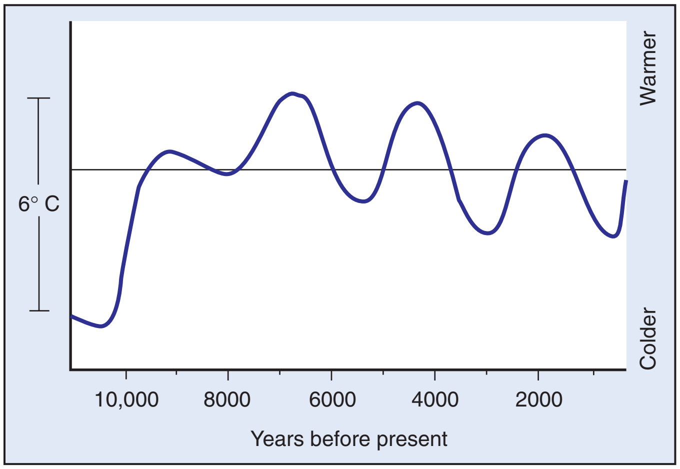


FIGURE 119-2 Variations of mean global temperature during the past 10,000 years. Horizontal line represents present global average temperature. *(Modified from Henderson-Sellers A, Robinson PJ:* Contemporary climatology, *New York, 1991, Wiley.)*

Despite recent controversies regarding the accuracy of some climate science, a strong consensus exists among atmospheric scientists that global temperatures will rise significantly in coming decades.[41,70] One indicator of this trend is the fact that " … during 2014, the average temperature across global land and ocean surfaces was … the highest among all 135 years in the 1880-2014 record, surpassing the previous records of 2005 and 2010."[57] Although the global climate system is enormously complex, two factors point toward warming. First, it is known that certain greenhouse gases warm the atmosphere by trapping short-wave radiation reflected from the earth's surface when it is heated by solar radiation. Second, atmospheric concentrations of these gases, which include carbon dioxide (CO_2), methane, and nitrous oxide, are steadily increasing. Normally, the materials in greenhouse gases pass through long biogeochemical cycles between natural sources and natural sinks. For example, sulfur enters the atmosphere as sulfur dioxide from volcanic eruptions and washes back to the oceans in the form of mildly acidic rainfall, the constituents of which are later incorporated into bottom sediments. Human activities can increase source loads (e.g., emissions) and reduce the absorptive capacities of natural sinks. In the case of CO_2, the greenhouse gas that has raised the most environmental concern, both processes are at work simultaneously. Emissions of CO_2 have been increasing as energy-hungry societies burn petroleum hydrocarbons, coal, and wood. At the same time, forests that usually absorb huge amounts of atmospheric CO_2 continue to be cleared, although at rates less than that predicted in recent decades.[20]

Atmospheric scientists have estimated how climate might change as greenhouse gases accumulate. For this purpose, they rely heavily on *general circulation models* (GCMs) that mathematically simulate the global climate system. The chemistry and physics of climate are complex, and the models, although increasingly sophisticated, are still imperfect. Their accuracy is constrained both by the limits of current knowledge about the dynamics of the atmosphere and by the computational power of the most advanced supercomputers. They are also hedged with other limitations. For example, the present generation of GCMs is too coarse to provide more than a broad-gauge portrayal of atmospheric conditions in a lattice of 150- to 200-km-wide (90- to 120-mile-wide) regions over the earth's surface. The models are able to project some climate variables (e.g., temperature) with a high degree of accuracy, but there is lower confidence in their ability to predict other variables (e.g., precipitation). GCMs do not reveal storm systems that bring most of the weather to middle and high latitudes or smaller-scale features, such as hurricanes, tornadoes, and other extreme local winds. They also do not incorporate the role of clouds as reflectors and absorbers of energy. The models do not satisfactorily account for all the CO_2 believed to have been liberated into the atmosphere through human activities. Nonetheless, many scientists have considerable confidence in the accuracy of GCMs because of their relative success in replicating present and past climates.

The Intergovernmental Panel on Climate Change (IPCC), a large joint United Nations–World Meteorological Organization committee of leading Earth scientists, won the Nobel Peace Prize in 2007 for its work synthesizing existing research on climate change. It has completed five comprehensive assessments of the state of scientific knowledge and reached sobering conclusions. The most recent assessment (2014) concluded that there is now "unequivocal" evidence that global average temperatures are increasing and projected to rise by 1.5° to 2.0°C (by 2100, under most scenarios). Similarly, there is likely to be increasing disparity between wet and dry seasons and between wet and dry regions, as well as a continuing rise in global sea level at an accelerating rate.[36,75] Although the IPCC estimates embody a consensus about global warming, the level of agreement declines as researchers attempt to forecast the resulting impacts, especially at regional and local levels. An enhanced greenhouse effect would have a greater effect on global climate than would temperature alone. Solar radiation provides the energy that drives the climate system. The effects of a warmer atmosphere could produce a cascade of changes in many climate variables.[14] For example, precipitation and evaporation would also likely increase, especially over high latitudes, but with strong regional variations elsewhere. Increasingly accurate climate projections for regional patterns of climate change have been developed.[12,23] We will soon know a great deal more about regional patterns of climate change, because a large number of studies that combine GCM data and other indicators of climate are being conducted.[2]

The GCMs generally indicate that lower latitudes and lower elevations will be less affected by anticipated climate changes than will upper latitudes and higher elevations.[12] However, climate change impact simulations paint no simple pictures. For example, temperature changes in the tropics may be relatively small, but their effects on insects could be more deleterious than in higher latitudes, where insects are typically more heat tolerant.[18] Most likely, a mosaic of regional and local changes with varying impacts will occur along a spectrum from strongly positive to strongly negative, depending on how, when, and where they occur[25] (Table 119-1). For example, some tropical islands, particularly in the Indian Ocean, are likely to experience heavy precipitation combined with more frequent severe storms and rising sea levels. Other islands, such as those in the Caribbean, are also likely to see increasing sea levels, but may experience a decline in summer rainfall. The net impacts of such changes are difficult to assess, but the experiences of Indian Ocean islands and coastlines during the massive tsunami of December 26, 2004, show just how vulnerable these regions already are to natural disasters. For the Andamans, Maldives, Seychelles, and other heavily populated low-lying islands of the Indo-Pacific Ocean, the results of sea level rise could be disastrous, whereas other places, such as high-standing islands of the Caribbean, could see offsetting agricultural benefits.[60]

TABLE 119-1 Effects of Global Climate Change on Different Regions

Region	Climate Change Effects	Level
Africa	Water stress	High
	Food shortages	Very high
	Mosquito-borne and waterborne diseases	Very high
Europe	River and coast flooding	Medium
	Water stress	High
	Heat waves and air pollution	High
Asia	Flood damage	High
	Extreme heat deaths	Very high
	Drought-related malnutrition	Medium
Australasia	Coral reef and species loss	High
	Coastal flooding	Medium
North America	Wildfire destruction	Very high
	Heat wave deaths	High
	Extreme rainstorm damage	High
Latin America	Water stress in semiarid areas	Very high
	Urban flooding caused by extreme rainfall	Very high
	Decreased food production	Very high
Polar regions	Risk from permafrost, snow, and ice changes	High
	Food insecurity; unsafe drinking water	Very high
Small islands	Property losses caused by rising seas	High
	Loss of coastal land	Very high
Oceans	Decline in low-latitude fish catches	Medium
	Biodiversity loss from damaged corals	Very high
	Erosion and sedimentation of coasts	High

Modified from Climate change 2014: Impacts, adaptation, and vulnerability. In Field CB, Barros VR, Dokken DJ, et al (eds). Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge: Cambridge University Press, United Kingdom.

More than any other factor, the rate of climate change is of concern to humans. GCMs indicate that absolute changes in temperature will be smaller than those that have occurred at other times during Earth's history. However, anticipated climate changes would still occur at a rate and magnitude that are unprecedented in human experience. Whereas past changes usually occurred slowly enough for plants and animals to adapt or migrate, examples exist of mass extinctions following rapid change. Many scientists fear that today and in the future, insufficient time and undeveloped areas will be available for plants and animals to make similar adjustments.

Although changes in average climate would have important long-term consequences, variations in extreme weather might produce the most immediate and significant impacts.[36] Droughts, floods, and tropical cyclones are unusual events in the current climate. If mean climates change, changes in frequency and severity of these extremes would probably also occur and become manifest well before permanent shifts could be confirmed.[28] Geographic distribution of such events would also be affected. According to the IPCC, in the future, heat waves likely will be "more intense, more frequent and longer lasting," and declines in frost days and cold waves are projected over all land areas. The frequency of extreme precipitation events would also likely increase everywhere.[35] There would be increased incidence of drought and water shortages. This may also have profound impacts for the local environment, particularly in locations where the local flora and fauna may not be adapted to prolonged dry periods. Sea level rise is likely to increase on all coasts, except in a few locations where land uplift negates its impacts. This is likely to lead to increased erosion and coastal flooding, increasing pressures on sensitive coastal environments, such as dunes and mangroves. Changes in both minimum and maximum temperatures experiences, particularly in the high latitudes and upload areas, are likely to have important impacts on the ability of some plant species to survive. In other areas, these changes may have a positive impact on some species because of increased length of the growing season.

Overall, the impacts of global climate change for wilderness areas will be complex and varied, influenced by a variety of local and regional factors as well as by global trends. As a result, natural hazards would likely pose increased risks to society. Moreover, exposure and vulnerability to extreme events would probably be exacerbated, because populations at risk might respond to the new conditions on the basis of outdated information and assumptions.[51] We might find that our previous experience prepared us to "fight the last war" rather than the current one.

STRATOSPHERIC OZONE DEPLETION

The stratosphere is a distinct layer of the upper atmosphere that occurs between 14.5 and 56 km (9 and 35 miles) above the ground. It contains significant concentrations of ozone (O_3), a gas that is formed when solar radiation splits oxygen atoms.* The stratospheric ozone layer absorbs most of the ultraviolet (UV) radiation from space that would otherwise damage plant and animal species.

During the 1970s and 1980s, researchers discovered that stratospheric ozone was being depleted, and that the ozone layer was thinning to the point of disappearance, particularly in polar

---

*Ozone also accumulates near ground level as a byproduct of the photochemical modification of exhaust gases from automobiles and other sources of pollution. Concentrations of this type of ozone are sometimes reported in local news media, but the ground-level "ozone problem" should not be confused with the stratospheric one.

regions.[27,49] Chlorofluorocarbons (CFCs) were held to be at fault. For decades, these synthetic compounds had been manufactured in large quantities, mainly for use as aerosol propellants and refrigerants. Once CFCs escape into the atmosphere, they remain stable until reaching the stratosphere, where they decompose under the action of UV radiation. Chlorine atoms are released and bond with ozone atoms, breaking them down into oxygen and other products. As a result, the ozone shield is weakened or removed.

If the ozone layer is sufficiently depleted, intensity of UV radiation that reaches the earth's surface could be significantly increased. This could have deleterious consequences for human populations and plant and animal species. For humans, increased incidence of skin cancer, cataracts, and immune system suppression are three recognized effects of high UV exposure. Although humans might take precautions to protect themselves against UV radiation, such as reducing time spent outdoors or adding sunblock, sunglasses, and clothes, nonhuman species may not be able to make the necessary adaptations. Serious disruptions of human and agricultural systems are possible.

During their winter seasons, the Antarctic[61] and to a lesser extent the Arctic[46,73] have experienced elevated levels of UV radiation. "Ozone holes" have been clearly traced to CFCs. An international agreement, the Montreal Protocol, was reached to phase out CFC use by 1996. Much progress has been made toward that goal, but these compounds are still being produced in some developing countries, and the substitute chemicals that were introduced elsewhere may also contribute to global warming.[42] Chlorine atoms are extremely long-lived in the stratosphere, persisting perhaps for 100 to 200 years. There will continue to be some potential for additional ozone depletion in the decades to come.

EROSION OF BIODIVERSITY (See Chapter 120)

Loss of species or the habitats that support them is a controversial and potentially serious global problem that comes under the heading "erosion of biodiversity." Biodiversity is not an agent of change as are greenhouse gas buildup and ozone depletion. Rather, it is an index against which environmental changes can be assessed.[72] As with climate change and ozone depletion, biodiversity has a global to local range of dimensions.[34] Two aspects of biodiversity of great importance are numbers and interconnections of species.

Estimates of the number of existing species range widely because the state of knowledge about the planet's biologic resources is both uneven and incomplete.[74] It is estimated that the earth hosts between 5 and 15 million species. About 1.75 million of these have been named.[44] Higher-order mammals and birds in temperate ecosystems are well documented, but insects, worms, and microscopic life-forms in tropical regions are much less known. In the United States, approximately 100,000 species are recognized, but only about one-fifth of these have been surveyed to date.[74]

Paleobiologic research indicates that the number and type of species have varied greatly over time. New species evolve through adaptive genetic mutations, while others perish because of competitive pressures of natural selection. Emergence and disappearance rates depend on the speed and direction of environmental change and the ability of species to adjust. What is most troubling about the recent record is the disappearance of so many species. "Between 1600 and 1994, at least 484 species of animals and 654 species of plants (mostly vertebrates and flowering plants) became extinct. During this period, the rate of extinction in groups such as birds and mammals also increased dramatically. Nearly three times as many species of birds and mammals became extinct between 1810 and 1994 (112 species) as were lost between 1600 and 1810 (38 species)."[64] Plant losses are presumed to have been much greater. On some oceanic islands, such as Hawaii, disappearance of native animal species is almost total. Of 269 extinct Hawaiian species, most were either invertebrates (135 species) or plants (105 species). A majority of the rest are birds (15 species) and land snails (11 species).[74] Commercial forestry and fishing have proved particularly injurious to biodiversity because they simultaneously harvest desirable species and destroy undesirable species.[33] Agriculture and animal husbandry also contribute to species extinctions, especially by modifying habitats that support biota. Particular concerns have been expressed about threats to tropical forests and near extinction of certain marine species such as the northern cod, blue whale, and leatherback turtle. However, the problem is general in scope and may be most important for the "noncelebrity" species that do not elicit much human compassion.

It is estimated that approximately 32% of all U.S. species are now under serious pressure, sometimes to the point of threatened extinction (Tables 119-2 and 119-3).[74] However, the picture varies widely among particular groups of species (Table 119-4). Inhabitants of freshwater ecosystems, such as shellfish, crustaceans, amphibians, and fish, are much more likely to be in danger than are flowering plants, conifers, mammals, and birds.

TABLE 119-2 Percentage of U.S. Species at Risk

Status	%
Extinct	1.0
Critically imperiled	6.5
Imperiled	8.8
Vulnerable	15.4
Secure	69.3

Modified from Stein BA, Flack SR: Conservation priorities: The state of U.S. plants and animals, *Environment* 39:6, 1997.

TABLE 119-3 U.S. Species Groups at Risk

	Extinct	Imperiled	Vulnerable	Secure
Freshwater mussels	16.4	39.7	11.8	32.1
Crayfish	17.3	32.8	0.9	49.0
Amphibians	14.0	23.9	2.5	59.6
Freshwater fish	14.1	22.0	2.6	61.3
Flowering plants	16.6	15.8	0.9	66.7
Conifers	12.2	14.0	0.0	73.8
Ferns	11.9	8.9	0.7	78.5
Tiger beetles	13.6	6.3	0.0	80.1
Dragonflies	10.4	7.6	0.4	81.6
Reptiles	11.9	6.1	0.0	82.0
Butterflies	12.3	4.0	0.5	83.2
Mammals	9.1	7.2	0.2	83.5
Birds	5.4	5.8	3.3	85.5

Modified from Stein BA, Flack SR: Conservation priorities: The state of U.S. plants and animals, *Environment* 39:6, 1997.

TABLE 119-4 Species Threatened With Extinction: a Global Picture

Status	Numbers
Extinct	801
Extinct in the wild	63
Critically endangered	3947
Endangered	5766
Vulnerable	10,104
Near threatened	4467
Data deficient	10,497
Least concern	27,837
Other	255
Total species assessed	63,837

From 2012 IUCN *Red List of Threatened Species.* http://www.conservation.org/NewsRoom/pressreleases/Pages/Securing-the-web-of-life-31-Percent-of-Species-Threatened-with-Extinction.aspx. Accessed February 9, 2015.

Likewise, on the U.S. mainland, loss of biodiversity is more acute in Sunbelt states and east of the Mississippi River than in states of the northern Great Plains and the northern Rocky Mountains.

For many people, protection of threatened species is a moral imperative. For others, it is a luxury. Quite apart from moral issues, the rising rate of species extinction has practical implications. For example, loss of the planet's genetic stock hampers the search for wild strains of domestic crops that are resistant to pests and diseases that plague high-yield domestic varieties. The so-called Green Revolution that has helped to alleviate world hunger in recent decades owes much of its success to introduction of resistant wild genetic strains into commercial agriculture.

Biodiversity is also important for stability of global ecosystems. For example, the extent to which entire species can be eliminated from an ecosystem before it collapses is unknown. Likewise, the extent to which some nominally "wild" species may thrive under human management, while others succumb, is hotly debated.[24] Most ecologists believe that ecosystems containing a wide diversity of organisms are more resilient to change than are those with few species. Regardless of the degree of resilience, biodiversity and environmental change may be connected by negative feedback relationships. Environmental change may lead to loss of biodiversity that in turn produces lowered resistance to pressures for further change.

Despite intuitive, theoretical, and case study arguments in favor of preserving biodiversity, it has been difficult to agree on standardized measures of biodiversity or its loss. Deforestation of South American rain forests is a case in point. The Amazon Basin is one of the world's premier wilderness regions and is regarded as Earth's most important source of biodiversity. Perhaps impelled by dramatic and widely publicized reports of forest clearances by ranchers, homesteading farmers, and mineral firms in Brazil during the late 1980s, levels of international concern about loss of biodiversity in Amazonia have been high. As in most developing countries, however, comprehensive and reliable data on Brazilian deforestation are difficult to secure and interpret.[70,88]

Given the foregoing uncertainties, it is difficult to predict future rates of loss of biodiversity. The best available estimates suggest that about 3.5% of current bird species will likely become extinct by the year 2050, together with most large marine predators and much of species richness of freshwater ecosystems.[37]

POPULATION GROWTH

Human population is frequently cited as one of the primary driving forces behind contemporary environmental change. Beginning with the Reverend Thomas Malthus (1766-1834), many have argued that rising populations must eventually deplete resources and degrade environments, with potentially catastrophic consequences for the biosphere and the human populations who depend on it, because the earth is, for practical purposes, a closed system.[40] In the absence of interplanetary space travel on a scale impossible at present to destinations that are now unknown and perhaps nonexistent, Earth is our only home. This does not mean that it will be impossible for our planet to hold additional human populations. The record of the past four centuries demonstrates that global carrying capacity is highly elastic up to some, as yet unreached, limit.[10] Leaving aside the argument that human ingenuity can make possible the support for larger populations indefinitely, clearly from the perspective of burdens on the physical environment, how people live is more important than the number of people. All other factors being equal, more affluent societies place heavier burdens on the physical environment than do poorer ones. For example, per-capita consumption of energy in the United States is more than 15 times higher than that in India.[65]

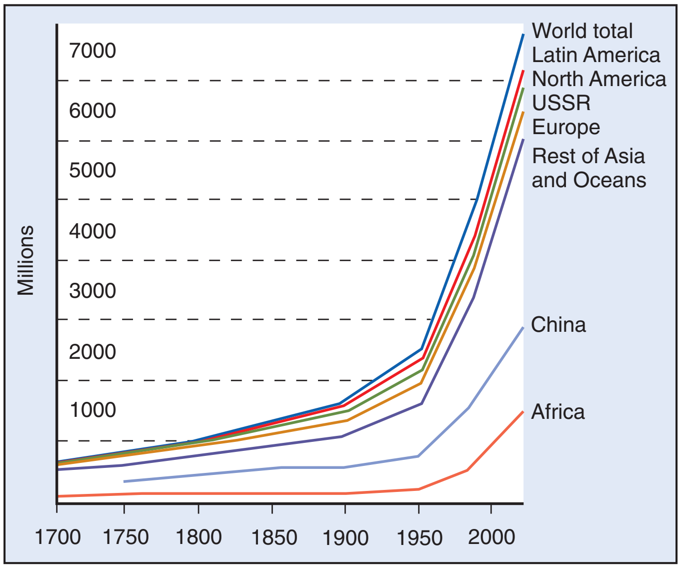


FIGURE 119-3 Global population (1700-2020). *(Modified from Demeny P: Population. In Turner BL II, editor:* The earth as transformed by human action, *Cambridge, UK, 1990, Cambridge University Press.)*

The global population has undergone unprecedented growth in the past several centuries (Figure 119-3). By 1800, Earth's population was approximately 1 billion. By 1920 it was approaching 2 billion. Three billion was reached by 1960, and the present number is more than 7 billion. The United Nations estimates that 9 to 11 billion people will be on Earth by 2050. Most of the new growth is likely to occur in developing countries of Asia, Africa, and Latin America (Table 119-5). The recent experience of China suggests that the process of development can itself perpetuate or increase historic rates of population growth, at least until economic conditions improve significantly.

Composition of future populations is an increasing concern of governments and individuals. In places such as Japan and Eastern Europe, natural increase is now well below the rate necessary to replace existing populations, whereas in the United States, increasing numbers are maintained largely by immigration. As a result, the fraction of national human populations older than

TABLE 119-5 Average Annual Percentage Rates of Population Increase

	1700-1750	1750-1800	1800-1850	1850-1900	1900-1950	1950-1985	1985-2020	2014
Africa	0.0	0.0	0.1	0.4	1.0	2.6	2.7	2.5
Asia	0.3	0.5	0.5	0.3	0.8	2.1	1.4	1.1
Europe	0.3	0.6	0.7	0.7	0.6	0.6	0.1	0.0
Russia	0.3	0.7	1.0	1.0	0.7	1.2	0.6	–0.03
North America	0.8	1.0	3.2	2.6	1.2	1.3	0.6	0.4
Latin America	0.8	0.5	1.2	1.6	1.6	2.6	1.6	1.2
Oceania	—	—	—	—	1.6	1.9	1.2	1.1
World total	0.25	0.44	0.55	0.54	0.84	1.88	1.45	1.14

From Demeny P: Population. In Turner BL II, editor: *The earth as transformed by human action*, Cambridge, UK, 1990, Cambridge University Press. With additional information supplied by the author.

65 years is growing rapidly in the developed world. Meanwhile, increasing expertise in genetic manipulation holds out the prospect of significantly longer life span for populations who can afford the scientific research and medical care that will make this possible. Some segments of the world's population will become increasingly healthier and older, while others may remain caught in a cycle of brief, sickness-prone lives followed by early deaths.

Apart from the staggering societal impacts that such a change would produce, the implications for wilderness are considerable. For those who can be assured of longer lives, the quality of life experience, including the quality of their environments, may become of great importance. Wilderness areas would be among the most cherished places, and decisions about their future all the more portentous. Among developing countries, a different future might emerge, perhaps dominated by intense pressure to convert all available resources into support for survival. Although such scenarios are not difficult to envision, they carry a danger of indulging in stereotypic dichotomizations that ignore possibilities for a range of more nuanced outcomes.

The greatest uncertainties about future populations pertain to rates of migration and composition of families.[11] Rates of migration are shaped by many factors, including the extent to which the negative consequences of environmental change may lead to population displacement. If current patterns of migration continue, the majority of migrants will settle in large countries such as the United States, India, Pakistan, France, and Germany, where depending on their internal destinations, they may add to the burden of users on existing wilderness areas.

POLLUTION

Unwanted byproducts of production and consumption that exceed the absorptive capacity of the environment are known as *pollution*. Pollution comes in many forms, including solid physical materials, liquid chemical compounds, and energy (e.g., thermal pollution). Some pollutants (e.g., certain isotopes of plutonium) are highly toxic even in small amounts. Many materials that are beneficial in small amounts can be deleterious in large quantities. For example, phosphorus is a nutrient that limits biologic productivity in coastal and marine ecosystems. Small amounts of phosphorus can increase algal growth at the bottom of marine food chains. However, when large amounts of phosphorus-rich runoff from fertilizers or septic systems enter these environments, the entire population of algae can begin a period of explosive growth ("bloom"). Extensive blooms can produce "red tides" or "brown tides."[9] This occurs when algae prevent light from penetrating coastal waters and decomposition of dead algae consumes dissolved oxygen. Large fish kills are a frequent result.

The preferences of people for quick and convenient disposal of pollutants into available environmental sinks (e.g., soil, streams, groundwater, oceans, atmosphere) have sometimes been validated by incomplete science. For decades, in the United States and elsewhere, scientists advised policy makers that "the solution to pollution is dilution." As a result, physical and chemical wastes have been released into environments that had finite capacities for absorbing them. Once the absorptive capacities were reached, a variety of serious problems occurred. These included biologically "dead" rivers (e.g., Cleveland's Cuyahoga River), lakes (e.g., Lake Erie), and seas (e.g., sewage sludge dumping ground in the New York Bight off the coasts of New Jersey and Long Island). Although some of these conditions can be reversed, the remediation processes are slow, costly, contentious, and often incomplete. A growing body of evidence suggests that the aggregate effect of pollution may be jeopardizing functions of fundamental Earth systems. Buildup of atmospheric CO_2 is an excellent example of this concept.

Despite a large volume of evidence, effects of pollutants on receiving environments are not fully known.[13] This is partly because of lack of scientific knowledge about the normal (unpolluted) functioning of some environments, such as deep oceans and tropical forests. The volume and variety of materials released into the environment and their interactions complicate the study of effects of any single pollutant. Sometimes the effects of pollutants are subtle, long delayed, and far removed from the point of origin, making it difficult to connect causes and consequences. Occasionally, experts disagree about evidence of pollution impacts collected in the field and acquired from laboratory experiments. Even the impacts of well-studied events, such as the *Exxon Valdez* tanker grounding, are in dispute.[6,56] Nonetheless, there is broad consensus that the absorptive capacity of receiving media is not inexhaustible, and that pollution is a growing world problem pushing society against the limits of environmental resilience.

IMPACTS OF ENVIRONMENTAL CHANGE ON WILDERNESS AREAS

The task of assessing environmental change impacts in wilderness areas poses particularly difficult challenges to researchers. To begin with, the term *wilderness* is rarely used by scientists. For example, readers of the voluminous reports of the IPCC[36] might search in vain for references to wilderness impacts. Instead, there are comments about the effects of climate fluctuations on specific types of ecosystems, land covers, or land uses, such as forests or nature reserves, some of which may be defined as "wildernesses" by different interest groups.[53] Lack of references to wilderness in the scientific literature can be explained by emergence of a widely shared conviction among scientists that no part of the world is now truly "natural." As one prominent ecologist put it, "Overall, any clear dichotomy between pristine ecosystems and human-altered areas that may have existed in the past has vanished."[80] In other words, our environments are arranged on a continuum from intensely human-constructed places, such as cities, to places where the human presence is small, intermittent, or nonexistent, such as wilderness areas. Most places show evidence of both human and natural influences.

If scientists shy away from the term *wilderness*, the same cannot be said of political leaders and the general public. Unfortunately, in this larger arena, there is little agreement about meanings of wilderness, because some of them are rooted in religious views about the perfectibility of the earth.

In addition, people in different countries often interpret *wilderness* in different ways. "Naturalness" and "remoteness" are two frequently mentioned wilderness attributes, but measures of both can be highly subjective. Depending on the definition adopted, a wilderness area might include some of the world's most biologically productive ecosystems (e.g., tropical rain forests of Brazilian Amazon Basin) together with some of its least productive (e.g., Sahara desert), as well as some that have been thoroughly transformed by human activities (e.g., so-called urban wilderness areas like Portland, Oregon's Forest Park). Some analysts have attempted to cut through this Gordian knot by equating wilderness with "uninhabited areas" or "roadless places," but neither of these indicators is comprehensive in scope when it comes to accounting for all the places that humans perceive as "wild." Using the criterion of human intrusion, all of Antarctica might qualify as wilderness, for by most measures it is one of the least human-impacted places on Earth. However, very few of the specially protected areas that have been established in Antarctica were designated because of their wilderness values; although aesthetics and wilderness characteristics are among the main criteria for designating these protected areas, historic sites of human activity are much more protected.[32] In other words, in Antarctica, the rarity value of human impacts outranks the importance of more or less pristine wilderness.

In view of the potential for confusion that exists in discussions of wilderness, it is worthwhile to address some of the philosophic and evolutionary backgrounds of the term.[53]

The concept of wilderness can be found in several of the world's earliest cultures,[63] but formulation of a powerful philosophic and political movement that espouses the value of unaltered natural areas did not occur in the United States until the 19th and early 20th centuries.[67] Although the wilderness movement subsequently diffused elsewhere, widespread public concern for preservation of wild lands remains a characteristically American preoccupation,[69] perhaps because the contrast between

TABLE 119-6 Global Crop Land Changes (1700-1990) in 1000 ha

Region	1700	1750	1800	1850	1900	1950	1990
United States and Canada	3077	6606	15,009	170,444	199,413	235,327	232,771
Central and South America	15,348	15,333	16,602	18,980	31,950	79,024	149,456
Europe	67,292	73,315	79,878	87,028	106,734	143,983	139,129
Asia and Russia	135,514	179,054	239,826	324,859	424,820	603,186	698,279
Oceania	2147	3500	5788	9659	16,291	27,750	53,063
Japan	1425	1458	1492	1527	2119	4720	4596
World	265,631	321,302	401,611	537,060	813,425	1,229,985	1,477,600

From History Base of the Global Environment, HYDE. http://arch.rivm.nl/env/int/hyde/.
ha, Hectares (1000 hectares = 10,000,000 m^2).

juxtaposed human-dominated landscapes and ostensibly "natural" ones is so apparent in the United States (e.g., California's Central Valley and Sierra Nevada mountains; Florida's Everglades and Gold Coast). Despite the volume of public debate about wilderness, the concept itself remains poorly defined, even in the United States.[*]

Most analysts recognize that *wilderness* refers to places that have one or all of three characteristics: (1) few or no permanent resident human populations, (2) unmanaged biogeochemical systems, and (3) no significant modification by modern technology. Places that meet these criteria might include deep oceans, high mountains, deserts, circumpolar lands, certain oceanic islands, coastal fringes, most areas of active vulcanicity, and some of the world's great forests (e.g., taiga, tropical rain forests).[†]

These three criteria are best regarded as necessary but not sufficient to identify an area as wilderness. Spatial dimensions must also be taken into account. An acre of wetland surrounded by shopping malls would not be considered wilderness, even if it is in biologically pristine condition. As a rule of thumb we have adopted, a wilderness usually encompasses at least several square miles.

If applied to the United States, the previous criteria would identify a great diversity of environments. Most would be marginal lands and waters beyond the boundaries of areas that are permanently settled, and perhaps without prospects for human occupancy or use in the long term. Some protected areas within the ecumene (inhabited lands) might also qualify as wilderness because they are administered as such. Most protected areas, however, such as national parks, national forests, and national recreational areas, would probably not meet all three major wilderness criteria because they are often subject to intensive management of residual plant and animal populations, as well as frequent human visits.[28]

CONVERSION OF WILDERNESS

In many parts of the world, the frontiers of wilderness areas are being pushed back as land is converted to managed uses. Economic growth and population increases are the ultimate driving forces of this conversion at the global scale. At local and regional levels, a variety of conversion processes are apparent. These include resource extraction industries (e.g., mining, forestry), agriculture, animal husbandry, tourism, and commercial and residential uses. These processes are not confined to land. They also affect freshwater and shallow-water marine environments. Tourism exerts pressure on coral reefs in Belize, Kenya, and many other countries.

Most land conversion is driven by demands for additional cropland. The highest levels of land conversion are found in developing countries with rapidly growing populations. By comparison, land formerly devoted to crops is reverting to an uncultivated state in much of Europe, North America, and Japan (Table 119-6). Most of the world's prime agricultural lands have been brought under cultivation, so attention has turned to other terrains that are spatially and agriculturally marginal.[3] These are often wilderness areas. For example, during the past six decades, many formerly unpopulated parts of Sumatra and Borneo have been settled by government-sponsored "transmigrants" from the heavily populated Indonesian island of Java.

Land conversion may fragment existing wilderness areas by dividing them into smaller blocks. This process is well advanced in the Amazon rain forest of Brazil, where new, long-distance, government-built roads bring settlers.[50] As a result, ecologic "islands" are created that may not be sustainable. Forest-edge environments replace deep-forest ones. The islands may be too small to retain the previous diversity of species. Governments often attempt to protect such islands by designating them as parks or wilderness areas, but this may be insufficient to prevent further changes. In any case, to be effectively protected, such places often require intensive management of ecosystems and visitors, which defeats the objective of designating them as wilderness. Moreover, management actions may ripple through the ecosystems in unforeseen ways, perhaps contributing to the long-term conversion process.

HUMAN PENETRATION OF WILDERNESS AREAS

The number of people visiting wilderness areas is on the rise. In the United States, approximately 70 million persons per year visited them in 2008. Increase in visitation rates is not confined to the United States. The deadly South Asian tsunami of December 26, 2004, was notable in part for the number of foreign tourists included among the estimated 300,000 dead. Citizens of 44 different countries from outside the region became victims because they were vacationing in resorts located on exposed islands and remote coasts. Many such places were in or near wilderness areas. Among other issues, this "global" disaster highlights the increasing spread and penetration of humans into formerly remote places.

People have not figured prominently in conventional definitions of wilderness, but wilderness areas often contain significant human populations. Many of the remaining habitats of endangered tropical species survive because they are located in places far removed from the pressures of modern society. However, this does not mean that these areas are devoid of people. For example, the Bonobo chimpanzees of the central Congo and the rhinoceri and tigers of Assam are sheltered by thick forests that also are home to humans numbering in the hundreds of thousands to millions. Sometimes, the remoteness of these places has

---

[*] The definition included in the Wilderness Act (1964) is typically vague: " … an area where the Earth and its community of life are untrammeled by man, where man himself is a visitor who does not remain." For practical purposes "roadless areas" are often used as an indicator of wilderness in North America. Even in a well-researched region such as North America, exhaustive inventories of the species present in wilderness areas are generally lacking; at the global level, only the approximate distribution of wilderness areas has been mapped.

[†] How should formerly developed areas that have reverted to unmanaged states be classified? Many such areas can be found in Western Europe and North America (e.g., Adirondack Mountains of New York). Often, radical differences exist between predevelopment conditions and reverted conditions. For the purposes of this discussion, such areas are considered wilderness.

made them havens for dissident political movements and sites of civil conflicts that have had devastating effects on indigenous plants and animals. As advocates of international programs to mitigate the effects of climate change have observed, anti-deforestation programs will not be successful if they fail to gain the support of wilderness-resident human populations.[30,66]

Wilderness areas may be degraded without being converted to other uses. This usually occurs in one of three ways: direct impacts from increasing human presence, indirect effects of conventional industrial technologies in adjacent areas, and global effects of innovative, powerful, and often high-risk technologies.

DIRECT IMPACTS

Few parts of the planet have remained unexplored by humans at ground level. Formerly remote areas are penetrated for a variety of reasons. Winter sports entrepreneurs are shifting attention to Europe's ecologically fragile High Alps, because snowfields at lower altitudes shrink under the forcing action of rising temperatures.[18] In Canada, James Bay has been altered by a huge hydropower scheme, and extraction of bitumen from the Athabasca Oil Sands is having profound impacts on local environments and ecosystems.[43] Gold prospecting has intruded into the innermost recesses of Amazonia and Angola. Philippine coral reefs are subject to cyanide poisoning in pursuit of aquarium fish.[17]

Penetration of wilderness is facilitated by modern industrial technologies, especially transportation technologies. For example, road building encourages invasion of wilderness areas for recreation, resource extraction, and other purposes. The roads themselves have environmental impacts ranging from vegetation clearance to drainage impedance, but their roles as conduits of change are even more significant. They bring new people, exotic materials, and different lifestyles to remote places. Similar inroads are made by boats and aircraft and their support facilities.

Because economic gain is an important incentive for wilderness penetration, recreational and esthetic needs also increase visitation. Hunting and fishing have long attracted visitors to wilderness areas, such as the Boundary Waters Canoe Area of northern Minnesota. Such pursuits are reinforced by "ecotourism." Increasing numbers of people want to visit remote areas to appreciate pristine beauty. For many people who formerly might have sought out Yellowstone National Park and the Grand Canyon, the destinations of choice include such places as Antarctica, the high Himalayas, Amazonia, and even Siberia. For many persons, the more remote the destination, the more attractive it is. Since most ecotourists want to visit the wilderness for only brief periods, they are whisked in and out by the most modern transportation technologies.

Visits from ecotourists can change wilderness environments. Seemingly insignificant impacts that are repeated can eventually become major problems. In the Masai Mara Reserve of Kenya's Serengeti Plains, the savanna ecosystem has been altered by photographic safaris. Safari camps require open campfires; fuel wood is scavenged from fallen trees that would otherwise provide important ecologic niches for local plants and animals. Climbing expeditions on Mt Everest have reported large volumes of garbage left by earlier expeditions. Decomposition is slow in the dry mountain air. Scarring of scientific sites in Antarctica by discarded refuse and vehicle tracks is well known. The Galápagos Islands, the one-time archetypical wilderness of Charles Darwin, are succumbing to the effects of their popularity with ecotourists. Geographers from the United States have assisted the government of Ecuador in carrying-capacity studies that form the basis for land-use regulations and other development controls to limit further degradation of these internationally valued sites.

INDIRECT IMPACTS

One of the most potent indirect impacts on wilderness areas follows introduction (inadvertent or intentional) of non-native species. Negative impacts have been demonstrated in the United States countless times, such as after the introduction of English sparrows, Asian gypsy moths, and Africanized "killer" bees. Everglades National Park in Florida is now one of the best places to find imported Burmese pythons living in the wild.[22] In Glacier National Park, pack trips within the park were curtailed because horses were introducing exotic species of grasses picked up from stable feed and passed through the digestive tract within their feces.

The problems of small islands and introduced species are legendary. Guam's experience with the brown tree snake is a good example. These snakes are native to New Guinea, but several managed to travel to Guam on airplanes in 1962. They thrived in the absence of native snakes or predators. Now Guam has as many as 30,000 brown tree snakes per square mile, and they have devastated native bird species. These snakes are beginning to show up in the Hawaiian Islands, where conditions are also favorable for colonization. Although efforts to intercept the snakes are being increased, the potential outcome is discouraging.

HIGH-RISK TECHNOLOGIES

Technologic risks are increasingly familiar threats to modern industrial society. Such risks are usually perceived as limited to accidents in urban industrial zones such as Bhopal, India, where more than 3000 people died following accidental release of methylisocyanate gas in 1984. However, some technologies have the potential to affect very large areas at great distances from their point of origin, up to and including the entire global environment. For example, the *Deepwater Horizon* explosion and subsequent oil spill in the Gulf of Mexico produced widespread impacts across a range of marine ecosystems, including deepwater corals.[82] Increasing oil and gas exploration presents a variety of risks to wilderness areas.

Biotechnology exemplifies some of these powerful, high-risk technologies. Through genetic engineering, new organisms are being created, primarily for agricultural purposes. Nuclear technologies also carry environmental change risks. For decades after World War II, a massive nuclear war between the United States and Soviet Union was a serious possibility. This would have brought catastrophic changes to the earth as a whole.[62] Many military nuclear facilities were located in remote areas. With the end of the Cold War, this threat has diminished, but regional nuclear conflicts among lesser powers are still possible. The risks of accidents involving nuclear weapons remain a threat to some wilderness areas. Nuclear bombs have been lost at sea; improperly managed nuclear wastes have exploded in the Ural Mountains and elsewhere; and military nuclear wastes are buried on small Pacific islands, often within reach of rising sea levels. Environmental contamination around nuclear weapon–manufacturing plants in the United States has been reported, and nuclear submarine propulsion systems have been discarded into the Arctic Ocean north of Russia.

Civilian uses of nuclear technologies pose risks to wilderness areas. Accidents such as the explosion and fire at the Chernobyl nuclear power station and the meltdown at Fukushima, Japan, triggered by the combined effects of an earthquake and tsunami, had global repercussions. Deposition of highly radioactive fallout in Arctic areas of Scandinavia demonstrates that no wilderness is immune from the effects of major nuclear accidents. Proposed placement of a repository for high-level nuclear waste in Yucca Mountain in the middle of semiarid Nevada provides another example of the connection between high-risk technologies and wilderness areas.

CONSEQUENCES OF ENVIRONMENTAL CHANGE

Research on global environmental change continues to reveal an ever-greater number of connections between human and natural systems. Linkages among species in a given ecosystem, among different ecosystems, and among global biogeochemical systems have been described. Providing details about all vulnerable systems is not possible, but the range of interconnections can be illustrated by two examples: Canadian wilderness use and coral reefs.

Scientists have recently explored the likely impacts of environmental changes on users of wilderness areas in northern Canada. In one case, rising temperatures and increased precipitation were judged likely to pose few problems for rafters and canoeists on the Mackenzie River, but accompanying forest fires were seen as much greater threats. Farther north on Bathurst Island, the likelihood of increased winter snowfall, combined with larger summer insect populations, seemed likely to stress the existing large caribou herds to a point where hunting might have to be curtailed. Throughout the region, a shift from consumptive uses (e.g., hunting) of wilderness lands to nonconsumptive uses (e.g., scenic tourism) is a potential outcome.[8] Elsewhere in the Arctic, there might be serious effects on resident and visitor populations. For example[2]:

> Indigenous people, dependent on climate conditions that support specific vegetation like forage for cattle and tundra climate commercial crops will need to change their lifestyle and adapt to suit the new environment. These changes in lifestyle would have long-term implications in all aspects including health. It is projected that although health conditions from frostbites and hypothermia would decrease with the reduction of cold stress in the region, heat-related diseases would become more common. Indirectly changes due to adjustments of dietary practices and weather conditions would change bacterial and viral proliferations that would result in specific health effects.

Coral reefs provide a second illustration of environmental change effects. Such reefs are among the most prized of wilderness ecosystems. Major reefs such as the Great Barrier Reef of Australia and the reefs off Belize are national and international treasures. Coral reefs cover only 0.17% of the ocean floor, an area approximately the size of Texas.[68,81] However, the importance of such reefs far exceeds their physical extent. Their biologic diversity is second only to that of tropical forests, and their productivity is among the highest in the world. They protect adjacent lands from wave action, nourish valuable fish populations, and generate millions of dollars in tourist revenues.

When subject to physical or chemical stress, coral "bleaches," losing color because of biochemical changes. Such stresses may be caused by fluctuations in sea level, temperature, or salinity and by pollution. Although reefs sometimes recover, bleaching often leads to death of the coral organisms and decomposition or disintegration of the reefs. In 1987, marine scientists began to notice high levels of coral bleaching and mortality off Puerto Rico. A worldwide pattern of severe coral bleaching began to emerge. Some scientists interpreted the problem as a harbinger of global warming, but it is unclear that this is the case. Nonetheless, coral reefs are vulnerable to temperature changes and sea level increases, so the threat of future damage is considerable. The best estimate is that sea level may rise an average of 1 m (3.3 feet) by 2100. Healthy reefs can grow upward by as much as 10 cm (4 inches) per decade, which may allow some reefs to adjust to rising sea level. However, if reefs are unhealthy, as the evidence of bleaching suggests, the rate of inundation may well exceed coral's ability to keep pace.[71]

Among the stresses that afflict coral reefs are coral mining for cement, dredging for navigation, coral collection for aquariums, and disruption by divers and commercial fishing. Many places also experience significant biochemical effects from coastal pollution and sediment or pollution runoff from land.

Loss of coral reefs is already significant. Estimates suggest that 5% to 10% of the world's living reefs have been destroyed by human activities. An additional 60% are thought to be at risk over the next 20 to 40 years.[85] The consequences for society are potentially enormous. Physical protection of coastlines could be drastically reduced. Locally, rich fisheries of coral islands could be diminished to the impoverished levels that typify deep oceans. Prized tourist attractions would disappear along with the revenues they generate. Opportunities for recovery of medicinal products (e.g., kainic acid) from reef organisms could be lost. Finally, the genetic resources of the planet could be further eroded. These are just some of the consequences of environmental change for one type of wilderness area. Similar, perhaps larger, effects may occur elsewhere.

ENVIRONMENTAL CHANGE AND MEDICAL EMERGENCIES

The causes and characteristics of many medical emergencies, and perhaps also the appropriate responses, are directly and indirectly connected with the environment in which they occur. This text contains many examples of medical challenges that are posed by environments in general and wilderness environments in particular. In some cases, an environmental agent (e.g., reptile bite, altitude sickness, wild animal attack) causes a medical emergency. In others, the environment affects the treatment of problems that are not environmentally created (e.g., wilderness trauma and surgical emergencies, hunting injuries, wilderness medical liability). In many cases, the environment serves as both agent and context. Inasmuch as the process of environmental change is global in scope, it probably will also affect wilderness medicine. A number of examples follow.

Increasing human penetration of wilderness areas by hikers, hunters, skiers, climbers, white-water boaters, and others is steadily driving up the number and cost of wilderness emergencies. For example, the U.S. National Park Service (NPS) spent $5.2 million on 2876 rescues in the United States in 2012.[79] NPS personnel, the U.S. military, and volunteers may be exposed to high risk when called on to retrieve inexperienced and underequipped parties. Given the rising cost of such operations, it has been proposed that individuals who participate in risky adventures should post rescue bonds before departing into the wilderness. California enacted a law that permits local authorities to charge persons who were aided up to $12,000 for each rescue performed by public agencies.[45] Similar laws exist in Hawaii, Idaho, Oregon, and New Hampshire. The combination of increasing populations and projected changes in environmental conditions can only add to future costs and difficulties of search and rescue in wilderness areas.

As settlement advances into wilderness areas, new patterns of disease are likely to emerge. For example, African land conversion from unmanaged wetlands to irrigated agriculture may spread the range of schistosomiasis and other waterborne diseases that are associated with drainage canals. Likewise, more people may be exposed to virulent diseases that are characteristic of wilderness ecosystems. Conversion of tropical forest in Africa may increase exposure to malaria carried by mosquitoes, onchocerciasis (river blindness) carried by *Simulium* flies, and trypanosomiasis (sleeping sickness) carried by tsetse flies.[4]

Pollutants often migrate into wilderness areas ahead of people. Air pollution is particularly mobile. Higher smokestacks are a common means of diluting airborne pollutants, but they also allow these materials to disperse more widely. Trees and lakes in New York State's Adirondack Mountains have been affected by acid rains transported from the Ohio Valley, and once clear vistas in the Grand Canyon have been obscured by smoke from a distant coal-fired power plant. The growing severity of winter haze in the Arctic is a problem.[73] Although the Arctic is a remote area, increasing haze has been observed there for almost a century. This smog consists of many different industrial pollutants that originate far to the south in industrial areas, especially the heavy manufacturing industries of Russia. Intense cold is the most obvious environmental health hazard in the Arctic, but buildup of industrial air pollutants may also have significant health effects, both directly on the body and indirectly through uptake by food sources from the Arctic environment.

The effect of weather on human mortality has long been a focus of biometeorologic research.[41] Box 119-1 lists a range of medical conditions that are weather related. For example, well-established linkages exist between high summer temperatures and human mortality, especially among elderly people. Although "global warming" need not mean that all parts of the earth will experience significantly elevated temperatures, some researchers are convinced that summer heat waves are likely to become more extreme, leading to increased mortality from this cause.[39]

The medical effects of UV radiation are known. Further erosion of the stratospheric ozone layer will undoubtedly increase the incidence of cataracts, skin cancers, and immune system

BOX 119-1 Causes of Death Considered to Be Weather Related

- Active rheumatic fever
- Adverse effects of medicinal agents
- Cerebrovascular disease
- Complications of medical care
- Complications of pregnancy and childbirth
- Contusion and crushing of intact skin surface
- Diseases of the arteries, arterioles, and capillaries
- Diseases of the blood and blood-forming organs
- Diseases of the digestive system
- Disease of the musculoskeletal system and connective tissue
- Diseases of the nervous system and sense organs
- Diseases of the skin and subcutaneous tissue
- Diseases of the veins and lymphatics
- Effects of foreign body entering through orifice
- Endocrine, nutritional, and metabolic diseases
- Fractures of the skull, spine, trunk, and limbs
- Hypertensive disease
- Influenza
- Injury to nerves and spinal cord
- Intracranial injury
- Ischemic heart disease
- Neoplasms: benign and malignant
- Superficial injury
- Toxic effects of substances of chiefly nonmedical sources

Modified from Kalkstein LS, Davis RE: Weather and human mortality: An evaluation of demographic and interregional responses in the United States, *Ann Assoc Am Geogr* 79:44, 1989.

diseases. Reduction of biodiversity threatens to reduce availability of natural materials that have medicinal value. Ethnobotanists are currently working with traditional shamans in Amazonia to catalog medicinal properties of plants in tropical forests. Marine species are also an important source of new medicines. Their decline will impair new drug discovery.

The "ozone hole" is a dramatic example of the expanding capacity of humans to modify the biosphere. Usually the process is inadvertent, and wilderness areas are not singled out for attention. Sometimes, however, the very remoteness and isolation of wilderness areas encourage dramatic environmental changes. Such was the case in northwest Alaska in 1962 when the U.S. government buried 15,000 pounds of radioactive soil at Point Hope.[55] The project was conducted by the U.S. Geological Survey acting in conjunction with the Atomic Energy Commission. The intent was to study effects of Arctic environments on radioactive isotopes. However, the burial was illegal; no public hearings were held, no markers were erected, and high-level wastes instead of low-level wastes were included. When the land was returned to the Inupiat (Eskimos) in 1971, they were not informed about the buried soils. They now attribute current elevated cancer rates to living and hunting for many years in a contaminated area. Government officials reject this view. The Point Hope case is not an isolated example. There is significant evidence that metropolitan governments have often tended to regard wilderness peripheries and their populations as dispensable when issues of national security and the welfare of metropolitan residents are at stake.[16]

COMPLEXITY AND UNCERTAINTY

Although we have ample reason to be concerned about environmental changes that lie ahead for wilderness areas, the subject is hedged with complexity and uncertainty. The potential for change exists, but it is difficult to be certain how fast and how far such changes will proceed. The following two cases illustrate some of the dimensions of complexity and uncertainty.

The north (Na Pali) coast of the Hawaiian island of Kauai is representative of wilderness areas that are particularly vulnerable to climate change. It is one of the most remote and beautiful places in Hawaii, accessible only on foot, from the ocean, or by air, weather permitting. The potential for increased rainfall, storminess, and sea level rise could radically alter this wilderness. For example, increased rainfall on Kauai's massive central peak, Mt Waialeale ("the wettest place on Earth"), would make for difficult hiking on steep Na Pali access trails that are already subject to erosion and landslides. The few available campsites near beaches may be eliminated by rising sea level. Sea caves that can be entered only by small, inflatable powerboats during calm conditions may become inaccessible. Flash floods in Na Pali streams may erode archeological sites, and increased moisture in the air would add to the mistiness that is now only an occasional feature of the area. Offshore waters host migrating whales that can be seen from the coast, but increased soil erosion might add to sediment loads and discourage the presence of these majestic mammals.

As the Na Pali coast becomes increasingly hazardous to visitors on foot, larger numbers may try to enter by helicopter, with more high technology–dependent visitors and fewer low technology–dependent ones. Health and safety emergencies may increase, or the mix of emergencies may change. The skies over Na Pali are already crowded with noisy aircraft. Crashes and deaths would likely increase. Leptospirosis from Na Pali streams may become more frequent. The bacteria were introduced from Southeast Asia in imported rats and pigs. In 1989, the Hawaiian Islands reported 66 cases of leptospirosis, with two resulting deaths.[54] Despite the potential for problems, no one can yet say with certainty which, if any, of these changes will occur. Still, we see strong indications that the Na Pali coast will not remain in its present state.

A second case that illustrates the complex interplay of environmental linkages and the potential for problems is provided by the highlands of Papua New Guinea.[1] Since the sweet potato was introduced to this area in the 1500s, it has become a staple crop for residents of remote mountain valleys. Sweet potatoes are susceptible to frost damage and tend to deplete mountain soils. In response to these constraints, villagers have developed specialized social and agricultural adjustments, including the practice of "mounding" and a complex system of resource exchanges between residents of higher elevations and lower elevations. Global warming might reduce the frost hazard, but increased precipitation or increased UV radiation could also threaten crop survival. At present, we have no way of confirming the extent and severity of possible changes. Clearly, however, a delicately balanced system of human ecology such as this would not remain unaffected by climate changes of the type anticipated in the next decades.

WHAT MIGHT BE DONE ABOUT LIMITING ENVIRONMENTAL CHANGE?

We have suggested that change is a dominant, perhaps "normal," feature of the world's landscapes and environments. What is different about the present era of environmental change is the extent to which it is directly attributable to human decisions and actions. It seems unlikely that people will do nothing if the anticipated changes are perceived as threatening, especially if they are also perceived as caused by humans. However, it is unlikely that responses to environmental change will be motivated solely by concern about environmental hazards, including medical emergencies in wilderness areas. Recognition is growing worldwide that improved environmental quality is an appropriate goal for all countries, not just developed ones. Therefore, public policies toward the environment will seek both to mitigate risks such as those connected with environmental emergencies and to secure rewards by safeguarding and enhancing valued resources, such as wilderness areas.

CHANGES IN ENVIRONMENTAL SCIENCE AND POLICY MAKING

This chapter initially appeared in the third edition of *Wilderness Medicine* (1995). It was written after the 1992 United Nations Conference on Environment and Development (UNCED) prompted creation of several landmark institutional instruments intended to guide governments and peoples toward economically and environmentally sustainable futures. Since then, the pace of international cooperation in support of these goals has greatly accelerated. By 2012, with the exception of a half-dozen developing countries and (notably) the United States, all of the world's nations had signed at least 10 of the 14 most prominent international governance agreements aimed at managing threats to environmental sustainability.*[78] Wilderness does not feature specifically in any of these agreements as a legal designation, but the roles of wild fauna and wild flora as indicators and maintainers of healthy environments loom large in several.

The picture is somewhat different at the level of national governments. Many countries have set aside wilderness areas and possess wilderness management systems, but these have typically evolved in piecemeal fashion and exhibit complex governance arrangements that imperfectly straddle jurisdictions of different agencies and organizations with differing agendas. Institutional complexity is the norm even in such places as New Zealand and Iceland, where wilderness has high cultural significance and enormous economic salience.[59,87] In the United States, the federal government has an explicit commitment to maintenance and protection of wilderness areas that is expressed in different policies among four major federal agencies, almost 200 separate wilderness-related laws, and thousands of guidance documents that govern wilderness management arrangements in different locations.[83]

The U.S. National Wilderness Preservation System (NWPS) includes 796 areas totaling 110 million acres in 44 states and Puerto Rico;[84] the separate Wild and Scenic Rivers System includes 12,598 miles of 203 rivers in 38 states and Puerto Rico.[52] The purpose of these designations is to protect relevant areas against significant modification by humans.† Most of the designated wilderness acreage is in Alaska (60%), and the bulk of the remainder in the 11 westernmost coterminous states (see Chapter 118). Typically, a wilderness area is embedded within and surrounded by other types of public land, such as national forests, national parks, or fish and wildlife reserves. Many types of human uses and activities permitted in the surrounding lands have spillover effects on the wilderness areas. Wilderness management is usually in the hands of the same agencies that administer the surrounding public lands and has frequently been a neglected stepchild of those agencies. Moreover, recent research has disclosed that the boundaries of wilderness areas are often poorly suited to permit survival of many species they contain. Especially in the case of migratory animals or those with large territorial ranges, what happens to them outside the wilderness is just as important as what happens within. Therefore, it makes little sense to restrict efforts for limiting environmental change solely to wilderness areas. Such efforts usually need to be applied to the private lands and waters that interdigitate with federally managed wilderness and nonwilderness areas.

Holistic management principles and tools that can be applied across governmental boundaries between and within countries are increasingly in demand. Some of these are available, but more are needed. For example, Environmental Impact Statements are decision support tools intended to provide holistic integrative assessments of proposed development actions with potentially undesirable effects on human environments, including wilderness areas.[7] Newer forms of holistic planning and management instruments (e.g., Health Impact Assessments) might take into account contributions of wilderness areas to human well-being that have previously been overlooked or undervalued.[58]

Although national and international government initiatives may have grabbed the headlines, adoption of international norms for private business practices has grown apace during the past 20 years. During the past two decades, the International Standards Organization has issued more than 230,000 compliance certificates to firms that met tighter environmental performance specifications.[78] Most of these focus on broad classes of environmental impact; they have positive indirect effects on wilderness areas and natural systems that sustain them.

Nongovernmental organizations (NGOs) have played key roles in fostering increased awareness of deleterious environmental changes and the growing human contribution to them. The IPCC's leadership of international efforts to apply scientific knowledge to the redress of unwanted climate changes is only the tip of a much larger iceberg founded on enhanced involvement of laypersons and local communities in decision making about environmental governance. This is made possible by modern electronic information and communication technologies (e.g., handheld computers, smartphones, Internet, social media, cloud sourcing, participatory mapping using geographic information science). These have ushered in a new era of volunteered geographic information that provides opportunities for increased use of lay knowledge in expert decision making and co-production of wilderness management systems previously a domain of experts.[26] Widespread availability of Global Positioning System (GPS) devices has reduced some of the uncertainties of wilderness navigation and, especially when coupled with satellite phones and other communications devices, affected the sense of remoteness formerly a signal characteristic of human experiences in the wilderness.

New environmental interest groups are joining the fray, including several connected with medicine, health, and human well-being. For example, the American Medical Association formed an Environmental Health Task Force charged with studying harmful environmental issues such as waste disposal and ozone depletion. In addition, the National Association of Physicians for the Environment was established in April 1992 to educate physicians about environmental hazards to human health and to develop recommendations for policy makers. In 2009, the World Health Professionals Alliance, a body that brings together representatives of national nursing, dentistry, pharmacy, and medicine associations across the globe, issued a statement on combating the effects of climate change that contained recommendations for changes in professional and public policies.[86]

Traditional conceptions of wilderness have emphasized the value of nature in a pristine condition—a condition that must be protected and preserved against human modifications. As argued here, such a view does not square with the vast bulk of scientific knowledge that recognizes the pervasiveness of past human impacts on natural systems and looks toward future environments that are even more completely human dominated. Such recognition implies that science and society might both consider restoration of degraded environments as well as preservation of minimally disturbed ones. *Restoration ecology* has become a new growth area in the environmental sciences and a new tool for environmental managers.[38] The implications for wilderness areas of such a shift are considerable. On the one hand, a restored watershed is not the same as one that has never been allowed to deteriorate, and a hand-reared endangered species is not the same as one that survives without direct human help. On the other hand, it is possible that environments might be restored to states that are functionally equivalent to wilderness. This raises fascinating but as yet unanswered questions. Would the availability of human-constructed alternatives to natural areas mean a slackening in political pressures to preserve "authentic" wilderness? Would "synthetic wildernesses" contain suites of medical risks similar to other types of wilderness? Current debates about the wisdom of changing existing wilderness protection statutes to accommodate climate change adaptation measures, which might involve more active human intervention in those places, are representative of the issues in play.

Looking across the range of environmental change issues and responses, new institutions and philosophies of human-nature relations are emerging and being linked to a broad range of

---

*Basel Convention, Cartagena Convention, Convention on Biological Diversity, Convention on International Trade in Endangered Species of Wild Fauna and Flora, Convention on Migratory Species, World Heritage Convention, Kyoto Protocol, Secretariat for the Vienna Convention and for the Montreal Protocol, Ramsar Convention, Rotterdam Convention, Stockholm Convention, Convention to Combat Desertification, Convention on the Law of the Sea, Framework Convention on Climate Change.

†http://www.fs.fed.us/recreation/programs/cda/wilderness.shtml.

public concerns. Issues of environmental change are seen as intertwined with issues of economics and security. The principle of diversity in natural systems, which imparts resilience in the face of stress, is being replicated in social systems. This is a hopeful sign at a time when environmental changes are unprecedented in rate and magnitude.

REFERENCES

Complete references used in this text are available online at expertconsult.inkling.com.

CHAPTER 120

Biodiversity and Human Health

RICHARD S. SALKOWE

Biodiversity is defined as the variety of all life-forms that inhabit Earth. From the earliest prokaryotic microorganisms that resided on this planet approximately 3.5 billion years ago to the megafauna that presently roam the vast plains of the Serengeti (Figure 120-1), this diversity of life is a result of competitive and cooperative relationships among species that have resulted in a delicate balance of natural processes that are essential to maintenance of human health. More than a century ago, the famous naturalist and preservationist John Muir stated, "Whenever we try to pick out anything by itself, we find it hitched to everything else in the universe." He was referring to inherent interrelationships that exist among the physical and biologic components of our environment. These relationships result in diverse ecosystems that serve to filter air, purify water, protect us from hazards, and provide essential food resources. What has become alarmingly evident is that the present rate of ecosystem destruction, species extinction, and loss of genetic variety on planet Earth is associated with a concurrent increase in prevalence of invasive species, severity of damage associated with natural disasters, and spread of infectious disease. Biodiversity is in a state of crisis, and the balance of nature that is critical to our sustainable existence is at risk.

Loss of species diversity is occurring at a rate that is 1000 to 10,000 times greater than the natural background rate.[36] This has an insidious effect on planetary and individual well-being. The effects are seen in the compromise of coastal estuaries that serve as natural waste filters and barriers to storm surges. The consequences are evident in bleached coral reefs that provide habitats for fish species and in clearing of tropical rain forests that serve as carbon sinks and provide oxygen for the environment.

In 1992 at the United Nations Earth Summit in Rio de Janeiro, 150 government leaders agreed to sustainable conservation of biologic diversity for preservation of planetary health. This agreement, adopted as the "Convention on Biodiversity," defined biologic diversity as "the variability among living organisms from all sources including, inter alia, terrestrial, marine and other aquatic ecosystems and the ecological complexes of which they are part; this includes diversity within species, between species and of ecosystems."[7] The United Nations Educational, Scientific, and Cultural Organization, in recognizing that biodiversity is the "basis for human existence," declared 2010 to be the International Year of Biodiversity in an attempt to increase awareness of the importance of biodiversity to human well-being.

UNDERSTANDING THE ETIOLOGY OF THE BIODIVERSITY CRISIS

> *What is man without the beasts? If the beasts were gone, men would die from great loneliness of spirit, for whatever happens to the beasts also happens to man. All things are connected. Whatever befalls the earth befalls the children of the earth.*
> *—Chief Seattle of the Suquamish, 1854*

Chief Seattle's words are emblematic of the historic sentiments of people who were directly involved with the land for sustenance and shelter. Early civilizations revolved around small communal hunter-gatherer societies. These societies had integral dependence on interaction with nature. They were in constant contact with natural resources and depended on basic respect for the rhythms of nature to maintain societal sustainability. Native American Indians of the Eastern Cherokee Nation developed these ideals into a harmony ethic of noncompetitive and reciprocal symbiotic relations with nature and their fellow man[11] (see Chapter 112).

Successful hunter-gatherer societies gradually increased in population. Additional demand for food resources coincided with discovery of plant cultivation technology, leading to agrarian civilizations and additional needs pertaining to land ownership and permanent settlements. Agrarian culture and resultant success of permanent settlements eventually led to development of cities, states, and empires. The industrial revolution took hold as a result of advances in scientific discovery and need for greater productivity to meet the demands of a burgeoning populace. Civilization's advances gradually moved individuals further and further away from the necessity of physical contact with the natural world. Intermediaries with nature, such as farmers, fishers, and merchants, satisfied the sustenance needs of city dwellers. It is not coincidental that environmental degradation and biodiversity loss secondary to the byproducts of industrialization occurred without apparent knowledge in a society so seemingly independent of nature. It became easy to ignore an unknown and intangible threat, namely, biodiversity loss. Robert Ornstein and Paul Ehrlich[23] hypothesize that humans are affected by a lack of natural selection for response to slowly developing threats such as biodiversity loss. They explain this factor as follows[23]:

> Hundreds of thousands or millions of years ago, our ancestors' survival depended in large part on the ability to respond quickly to threats that were immediate, personal, palpable: threats like the sudden crack of a branch as it is about to give way or the roar of a flash flood racing down a narrow valley. Threats like the darkening of the entrance of a cavern as a giant cave bear enters. Threats like lightning, threats like a thrown spear. Those are not threats generated by complex technological devices accumulated over decades by unknown people half a world away. Those are not threats like the slow atmospheric buildup of carbon dioxide from auto exhausts, power plants and deforestation; not threats like the gradual depletion of the ozone layer. Thus, the human mind evolved to register short-term changes from moment to moment, day to day, and season to season, and to overlook the backdrop against which those take place.

REFERENCES

1. Allen BJ. Adaptation to frost and recent political change in highland Papua New Guinea. In: Allan NJR, Knapp GW, Stadel C, editors. Human impacts on mountains. Totowa, NJ: Rowman & Littlefield; 1993.
2. Arctic Climate Impact Assessment. Impacts of a warming Arctic: Arctic Climate Impacts Assessment. Cambridge: Cambridge University Press.; 2004 <http://amap.no/acia/>.
3. Arnell NW. Climate change and global water resources. Glob Environ Change 1999;9:S31.
4. Aron JL, Schiff CJ, Beck A. Malaria and global ecosystem change. In: Aron JL, Patz JA, editors. Ecosystem change and public health: A global perspective. Baltimore: Johns Hopkins University Press; 2001. p. 353–78.
5. Bailes KE. Science and Russian culture in an age of revolution, vol. I. Vernadsky and his scientific school (1863–1945). Bloomington: Indiana University Press; 1990.
6. Birkland TA. In the wake of the Exxon Valdez: How environmental disasters influence policy. Environment 1998;40:4, 27.
7. Blumm MC, Wisehart LM. The underappreciated role of the national environmental policy act in wilderness designation and management. Law Rev 2014;44(2):323–72.
8. Brotton J, Staple T, Wall G. Climate change and tourism in the Mackenzie Basin. In: Cohen SJ, editor. Mackenzie Basin Impact Study: Final report. ON [Canada], Climate, Atmos Res Dir, Downsview: Environment Canada; 1997. p. 253–64.
9. Cherfas J. The ocean fringe—Under siege from the land. Science 1990;248:163.
10. Cohen JE. How many people can the earth support? New York: WW Norton & Co; 1995.
11. Cohen JE. Human population: The next half-century. Science 2003; 302:1172.
12. Cohen J, Screen JA, Furtado JC, et al. Recent Arctic amplification and extreme mid-latitude weather. Nat Geosci 2014;7:627–34.
13. Conzen MP, editor. The making of the American landscape. New York: Routledge; 1990.
14. Cooney CM. Downscaling climate models: Sharpening the focus on local-level changes. Environ Health Perspect 2012;120(1):22–8.
15. Crutzen PJ, Steffen W. How long have we been in the Anthropocene Era? Clim Change 2003;61:251.
16. Dalby S. Ecopolitical discourse: "Environmental security" and political geography. Prog Hum Geogr 1992;16:503.
17. Derr M. Raiders of the reef. Audubon 1992;92:48.
18. Deutsch CA, Tewksbury JJ, Huey RB, et al. Impacts of climate warming on terrestrial ecotherms across latitude. Proc Natl Acad Sci U S A 2008;105(18):6668–72.
19. Ellis EC, Fuller DQ, Kaplan JO, Lutters WG. Dating the Anthropocene: Towards an empirical global history of human transformation of the terrestrial biosphere. Elem Sci Anth 2013;1:000018.
20. Food and Agriculture Organization. Global forest resources assessment 2010, Rome, p 3. <http://foris.fao.org/static/data/fra2010/KeyFindings-en.pdf>.
21. Forest Guild. Restoration and resilience in New Jersey's forests, Santa Fe, NM, 2014. <http://www.forestguild.org/publications/research/2014/New_Jersey.pdf>.
22. Forget the gators: Exotic pets run wild in Florida, New York Times, February 29, 2004.
23. Frauke F, Burkhardt R, von Storch H, et al. Regional climate models add value to global model data: A review and selected examples. Bull Am Meteorol Soc 2011;92:1181–92.
24. Freese CH, editor. Harvesting wild species: Implications for biodiversity conservation. Baltimore: Johns Hopkins University Press; 1997.
25. Gillson L, Willis KJ. As Earth's testimonies tell: Wilderness conservation in a changing world. Ecol Lett 2004;7:993.
26. Goodchild MF. Citizens as sensors: The world of geography. GeoJournal 2007;69(4):221.
27. Haas PM. Policy responses to stratospheric ozone depletion. Glob Environ Change 1991;1:224.
28. Hassol SJ, Katzenberger J. Elements of change: Natural hazards and global change. Aspen, Colo: Aspen Global Change Institute; 1997.
29. Henderson-Sellers A, Robinson PJ. Contemporary climatology. New York: Wiley; 1991.
30. Homji V, Meher M. Wilderness woes—Tribals versus tigers: Pros and cons. Curr Sci 2005;89:737.
31. Hooke RL, Martin-Duque JF, Pedraza J. Land transformation by humans: A review. GSA Today 2012;22(12):4–10.
32. Hughes KA, Convey P. The protection of Antarctic protected ecosystems from inter- and intra-continental transfer of non-indigenous species by human activities: A review of current systems and practices. Glob Environ Change 2010;20:96.
33. Hunter ML, editor. Maintaining biodiversity in forest ecosystems. Cambridge: Cambridge University Press; 1999.
34. Huston MA. Biological diversity: The coexistence of species on changing landscapes. Cambridge: Cambridge University Press; 1994.
35. Intergovernmental Panel on Climate Change. Climate change 2001: The scientific basis. Working Group I of the Third Assessment. New York: Cambridge University Press; 2001.
36. Intergovernmental Panel on Climate Change. Climate change 2014: Impacts, adaptation, and vulnerability—Synthesis report. IPCC; Special report on renewal energy sources and climate change mitigation, 2012; Managing the risks of extreme events and disasters to advance climate change adaptation, Cambridge: Cambridge University Press; 2012.
37. Jenkins M. Prospects for biodiversity. Science 2003;302:1175.
38. Jordan WR, Gilpin ME, Aber JB, editors. Restoration ecology: A synthetic approach to ecological research. Cambridge: Cambridge University Press; 1987.
39. Kalkstein LS, Davis RE. Weather and human mortality: An evaluation of demographic and interregional responses in the United States. Ann Assoc Am Geogr 1989;79:44.
40. Kates RW. The human environment: The road not taken, the road still beckoning. Ann Assoc Am Geogr 1987;77:525.
41. Kates RW. What kind of science is sustainability science? Proc Natl Acad Sci U S A 2011;108(49):19449–50.
42. Krueger JP, Rowlands IH. Protecting the earth's ozone layer. Glob Environ Change 1996;6:245.
43. Kurek J, Kirk JL, Muir DCG, et al. Legacy of a half century of Athabasca oil sands development recorded by lake ecosystems. Proc Natl Acad Sci U S A 2013;110(5):1761–6.
44. Lawton JH, May R. Extinction rates. Oxford: Oxford University Press; 1995.
45. Los Angeles Times. Thrill geekers pay the price, January 18, 2005.
46. Mahlman JD. Global change: A looming Arctic ozone hole. Nature 1992;360:209.
47. Marsh GP. Man and nature; or the earth as modified by human action [orig 1864]. Cambridge: Belknap Press of Harvard University Press; 1965.
48. Mitchell JK, Cutter SL. Hazard continuities and discontinuities in an era of environmental and societal transformation. Environ Hazards 1999;1:1.
49. Morrisette PM, et al. Prospects for a global greenhouse accord: Lessons from other agreements. Glob Environ Change 1991;1:209.
50. Myers N. The shrinking ark. Oxford, England: Pergamon Press; 1980.
51. National Research Council. Adapting to the impacts of climate change. Washington, DC: National Academies Press; 2010. p. 221.
52. National Wild and Scenic Rivers System. About the WSR Act—River classification, 2015. <http://www.rivers.gov/wsr-act.php>.
53. Nelson MP, Callicott JB. The wilderness debate rages on: Continuing the great new wilderness debate. Athens: University of Georgia Press; 2008.
54. New York Times. Alien species threaten Hawaii's environment, December 29, 1992.
55. New York Times. Eskimos learn they've been living amid secret plots of radioactive soil, December 12, 1992.
56. New York Times. Oil-spill damages disputed by Exxon, April 30, 1993.
57. NOAA National Climatic Data Center. State of the climate: Global analysis for annual 2014. <http://www.ncdc.noaa.gov/sotc/global/2014/13>.
58. North Slope Borough. Climate change and health impact assessment reports, Barrow, Alaska, 2014. <http://www.north-slope.org/departments/health-social-services/climate-change-health-impact-assessment-reports>.
59. Olafsdottir R, Runnstrom MC. How wild is Iceland? Wilderness quality with respect to nature-based tourism. Tourism Geographies 2011; 13(2):280–98.
60. Overseas Development Institute and Climate and Development Knowledge Network. The IPCC's Fifth Assessment Report: What's in it for small island developing states? London, 2014.
61. Pitari G, Viscunti G, Verdecchia M. Global ozone depletion and the Antarctic ozone hole. J Geophys Res 1992;97:8075.
62. Pittock AB, et al. Environmental consequences of nuclear war, SCOPE 28. New York: Wiley; 1986.
63. Redman CL. Human impact on ancient environments. Tucson: University of Arizona Press; 1999.
64. Reid WV. Strategies for conserving biodiversity. Environment 1997; 39:16, 19.
65. Robbins P. Political ecology: A critical introduction. Oxford: Wiley-Blackwell; 2012.
66. Robbins P, McSweeney K, Chhangani AK, Rice JL. Conservation as it is: Illicit resource use in a wildlife reserve in India. Hum Ecol 2009; 37:559.
67. Rudzitis G. Wilderness and the changing American West. New York: Wiley & Sons; 1996.
68. Salvat B. Coral reefs—A challenging ecosystem for human societies. Global Environ Change 1992;2:12.

69. Schmitt PJ. Back to nature: The Arcadian myth in urban America. New York: Oxford University Press; 1989.
70. Smith ZA. The environmental policy paradox, Upper Saddle River. NY: Prentice Hall; 2000. p. 239–40.
71. Smith SV, Buddemeier RW. Global change and coral reef ecosystems. Annu Rev Ecol Systemetics 1992;23:89.
72. Solbrig OT. The origin and function of biodiversity. Environment 1991;33:16.
73. Soroos MS. The odyssey of Arctic haze: Toward a global atmospheric regime. Environment 1992;34:6.
74. Stein BA, Flack SR. Conservation priorities: The state of U.S. plants and animals. Environment 1997;39:6–34.
75. Stocker TF, Qin G, Plattner K, et al. Climate change 2013: The physical science basis. Cambridge: Cambridge University Press; 2013.
76. Thomas WL Jr. Man's role in changing the face of the earth. Chicago: University of Chicago Press; 1956.
77. Turner BL II, editor. The earth as transformed by human action: Global and regional changes in the biosphere over the past 300 years. Cambridge, UK: Cambridge University Press; 1990.
78. United Nations Environment Programme. Keeping track of our changing environment: From Rio to Rio +20 (1992–2012), Nairobi, 2011. <http://www.unep.org/geo/pdfs/keeping_track.pdf>.
79. Villella A, Keen TK. Into the wild: Can regulation of wilderness recreational activities improve safety and reduce search and rescue incidents? Jeffrey S Moorad Sports Law J 2014;21(2).
80. Vitousek PM. Beyond global warming: Ecology and global change. Ecology 1861;75:1994.
81. Weber R. Reviving coral reefs: State of the world 1993. Worldwatch Institute, New York: WW Norton; 1993. p. 42.
82. White HK, Hsing PY, Cho W, et al. Impact of the Deepwater Horizon oil spill on a deepwater coral community in the Gulf of Mexico. Proc Natl Acad Sci U S A 2012;109(50):20303–8.
83. Wilderness.net. Document map, University of Montana, 2015. <http://www.wilderness.net/NWPS/docmap>.
84. Wilderness.net. The beginnings of the National Wilderness Preservation system, US National Park Service, 2015. <http://www.wilderness.net/nwps/fastfacts>.
85. Wilkinson CR. Coral reefs are facing widespread extinctions: Can we prevent these through sustainable management practices? Address at International Symposium on Coral Reefs, Guam, 1992.
86. World Health Professionals Alliance. Statement on health and climate change, Geneva, 2009. <http://www.whpa.org/jsClimate_change10.htm>.
87. Wray KA. The culture of the wild: An exploration of the meanings and values associated with wilderness recreation in New Zealand, PhD Dissertation, Lincoln University. 2009. <http://researcharchive.lincoln.ac.nz/bitstream/10182/2372/3/Wray_PhD.pdf>.
88. Zuidema PA, Sayer JA, Dijkman W. Forest fragmentation and biodiversity: The case for intermediate-sized conservation areas. Environ Conserv 1996;23:290.

FIGURE 120-1 Zebra in the Serengeti during wildebeest migration. *(Courtesy David Dennis, cc-by-sa-2.0.)*

The insidious processes of planetary degradation have reached a crisis phase. The miracle of the combustion engine and invention of plastics have become the potential bane of our existence, as shown by the consequences of injudicious burning of fossil fuels and discovery of toxic byproducts, such as dioxin and bisphenol A. Loss of biodiversity represents a unique challenge that demands an agenda for action. It is difficult enough to effectively communicate the risks associated with known technologic hazards, such as lead, mercury, and greenhouse gases. Convincing the public of the adverse consequences of the extinction of the dusky seaside sparrow (Figure 120-2) or the Monteverde golden toad (Figure 120-3) represents an even greater challenge. Loss of these flagship species is a story that must be told, because the sparrow's demise is a tale of the health-related dangers of dichlorodiphenyltrichloroethane (DDT) and mismanagement of marshland in the United States. In this regard, extinction of a seemingly inconsequential avian species serves as an indicator of the ecologic dangers associated with pesticide exposure and loss of the water filtration and hazard protection services that marshlands provide to protect human health in coastal areas.

The Monteverde golden toad succumbed to a multitude of pressures associated with invasive species introduced by tourists and aquatic chytrid fungal infections that are theorized to be correlated with El Niño–induced climate change in the toad's former home range of Costa Rica. Loss of this particular species is indicative of the threat that amphibians face worldwide. A recent study revealed that 32% of the 6000 species of amphibians under analysis were threatened and 43% were in decline.[25] The potential medicinal value of chemical compounds that have been extracted from amphibians is evident in the 200 psychoactive alkaloids that have been extracted from the skin of frogs and toads.[1] Some of these compounds have been used in medical research pertaining to nerve and muscle disorders. The alkaloid known as *epibatidine,* which is synthesized from skin of the phantasmal poison frog, is being tested as a nonaddictive and nonsedating analgesic that exhibits 200 times the potency of morphine.[1] Bufogenin and bufotoxin, substances that have been extracted from parotid glands of toads from the same *Bufo* genus as the extinct Monteverde golden toad, exhibit adrenal and cardiovascular effects in humans.[1]

FIGURE 120-2 Dusky seaside sparrow. *(Courtesy U.S. Fish and Wildlife Service. Public domain.)*

FIGURE 120-3 Monteverde golden toad. *(Courtesy U.S. Fish and Wildlife Service. Public domain.)*

Further investigation of the adverse sequelae of biodiversity loss, as evidenced by ecosystem degradation, species decline, and loss of genetic diversity, provides additional validation of the corollary threat to human health.

THREATENED ECOSYSTEMS

Ecosystems represent abiotic and biotic components of an environment. Variation in ecosystems, from arid deserts of the South American Atacama to Siberian subarctic taiga forests, is determined by climatic and geologic characteristics of the respective regions. Mineral components of soil and nitrogen-fixing capacities of soil microbes; salinity, turbidity, and other hydrologic aspects of the respective environment; and the variety of flora and fauna that inhabit a region create a syncretic balance that is sustainable in a healthy ecosystem. Speciation within these respective biomes and ecosystems has developed over millions of years, as chronicled by successes and failures of living organisms that did (or did not) establish sustainable niches in a vast array of seemingly hospitable and inhospitable locales.

Ecosystems provide a multitude of features that are essential to human well-being. Food resources, fresh water, sediment retention, nutrient cycling, disease regulation, erosion control, air quality, and climate change depend on healthy ecosystems.[18] Human activities, ranging from land-use patterns associated with increased urbanization to clear-cutting of rain forest for agricultural purpose, have degraded the quality of ecosystems worldwide. The British ecologist, Norman Myers, developed the concept of "biodiversity hotspots" to identify areas of the planet with a high number of endemic species under extreme threats

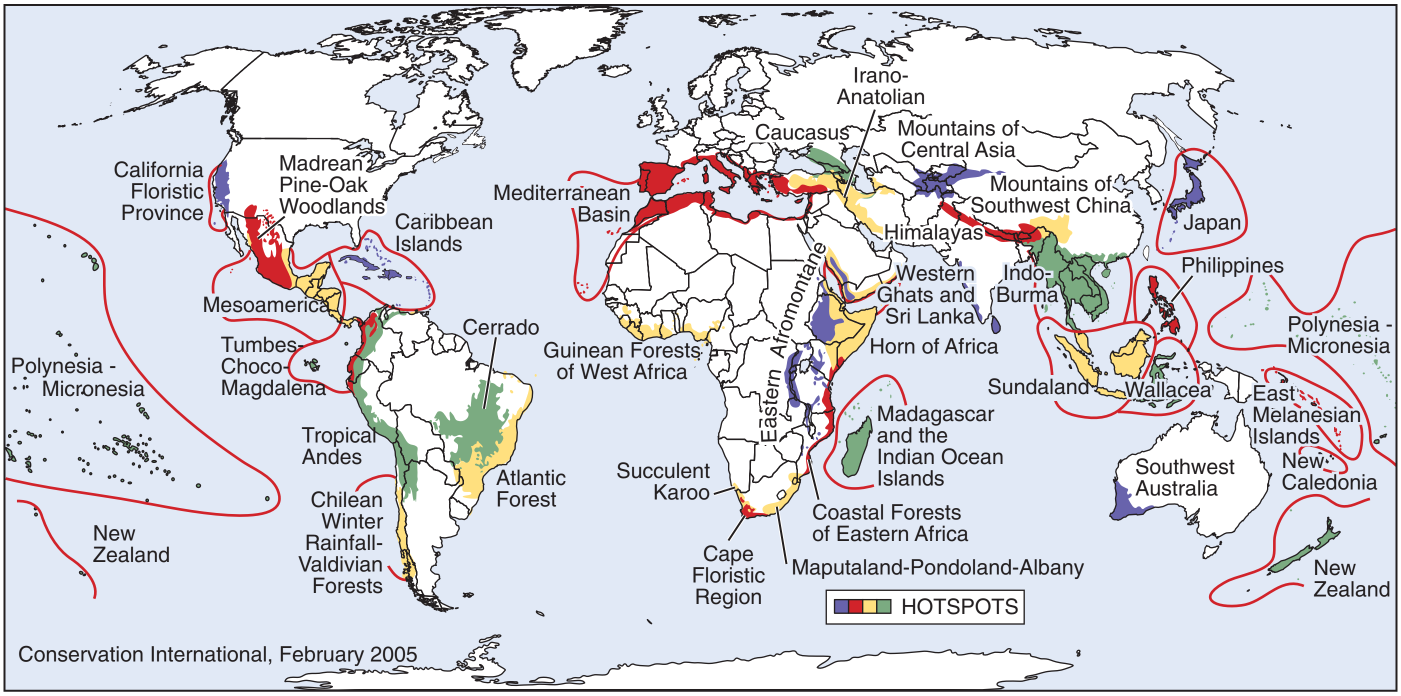


FIGURE 120-4 "Biodiversity hotspot" map. *(With permission from Conservation International.)*

to their ecosystems, as evidenced by loss of at least 70% of the natural vegetation.[6] Conservation International[6] states that "over 50 percent of the world's plant species and 42 percent of all terrestrial vertebrate species are endemic to the 34 biodiversity hotspots" (Figure 120-4).

Tropical rain forests exemplify the human health risks associated with ecosystem compromise. These biomes provide habitats for more than one-half of all plant and animal species in the world. Prior studies have counted 100 to 300 species in a 1-hectare area in South America. Approximately 25% of all pharmaceutical agents are estimated to contain compounds that are found in tropical rain forest plants.[2] The United Nations estimates that 13 million hectares of tropical forest are destroyed each year for logging and agricultural land clearing.[10] In South America, Mexico, and Guyana, rain forests have been cleared for cattle grazing, soybean farming, logging, and gold mining (Figure 120-5). Less than 1% of all tropical rain forest species have been evaluated for pharmacologic benefit. Degradation of this essential ecosystem has human health consequences that range from loss of potential lifesaving medicines to alterations in climate associated with carbon sequestration by forest trees and burning of forest land.

Mangrove estuaries and wetland sloughs represent another class of threatened ecosystems. Land development, agriculture, and aquaculture pursuits have led to loss of at least 35% of this ecosystem worldwide.[32] Considered to be one of the planet's most productive ecosystems, wetlands are referred to as "nature's kidneys," because they filter sediments and pollutants and regulate water flow. Mangroves act as a natural buffer to prevent coastal erosion and provide essential habitat for crustaceans, fish, and several other species. Runoff from excessive amounts of nitrogen- and phosphorus-based fertilizers used in monoculture farming leads to eutrophication of coastal forests and increased mortality of mangrove species as a result of root damage (Figure 120-6).

Coral reefs support more than 4000 species of fish and are home to approximately 25% of all marine species. In addition, reefs provide important breakwater protection for coastal areas during tropical storms. Scientists estimate more than 50% of the world's coral reefs face potential destruction by the year 2030.

FIGURE 120-5 Tropical rain forest destruction in Mexico. *(Courtesy Azari Nicks. Public domain.)*

FIGURE 120-6 Everglades National Park mangrove tidal estuary. *(Courtesy Moni2. Public domain.)*

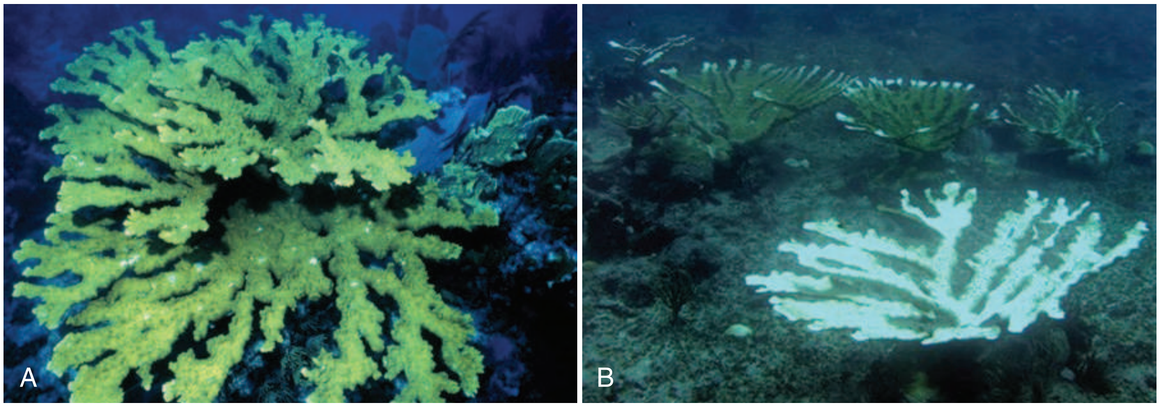


FIGURE 120-7 Florida Keys critically endangered elkhorn coral. *(Courtesy the National Oceanic and Atmospheric Administration. Public domain.)*

Global warming is a major threat to this important ecosystem; a sea temperature rise of 1° to 2° C (1.8° to 3.6° F) has been associated with physiologic stress and immune compromise to corals, which leave them with subsequent increased susceptibility to bacterial and fungal pathogens[3] (Figure 120-7).

SPECIES DECLINE

Species decline is another form of biodiversity loss. Variety of species in an ecosystem is critical to sustainability of the respective habitat. Excluding bacteria and viruses, approximately 1.5 million species have been taxonomically identified, and approximately 10 million species are believed to exist on Earth.[5] The 2008 *Living Planet Report* indicates that, between 1970 and 2005, the earth's wildlife populations declined by a third.[37] The International Union for Conservation of Nature has estimated that in 2010, 22% of all vertebrates, 34% of all invertebrates, 70% of all plants, and 50% of all fungi and protists were listed as critically endangered, endangered, or vulnerable species. Species that are considered threatened worldwide include 30% of all amphibians, 21% of mammals, and 86% of mosses.[16] Species are disappearing at the alarming rate of 1000 to 10,000 times the natural background rate of 1 to 10 species per year. E. O. Wilson, "the father of biodiversity," estimates the current extinction rate is 137 species per day in tropical rain forests alone[36] (Figure 120-8).

Invertebrate species, which represent approximately 76% of all life-forms, are experiencing a significant rate of extinction. Dam construction, water pollution, and deforestation have challenged the capacity of several invertebrate species to retain a foothold in ecosystems worldwide. A keystone species in the Antarctic ecosystem, the Antarctic krill, is indicative of the threats that face invertebrate species (Figure 120-9). Krill are an important food source for whales, seals, squid, penguins, and fish. In addition, these small crustaceans act as an essential component of the ocean's capacity to sequester carbon. Recession of the Antarctic ice pack and acidification of ocean waters associated with carbon dioxide emissions and global warming are challenges to vitality of the Antarctic krill.

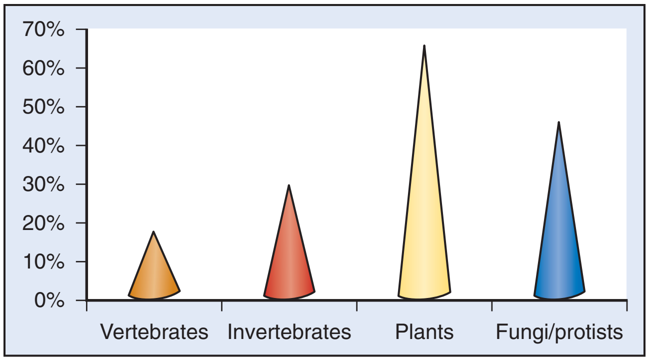


FIGURE 120-8 Extinction rate per millennium.

LOSS OF GENETIC DIVERSITY

Some analysts believe that the greatest threat to human welfare comes from losses of genetic diversity within species.[35] Farmers and pastoralists have used selective planting and breeding techniques for centuries to increase crop yield and product output. With the advent of techniques to genetically engineer crops for resistance to variations in climate and susceptibility to disease and pests, biodiversity of the gene pool has drastically altered. In 1970, the United States lost 15% of its Midwest corn crop as a result of a fungus that the genetically modified crop was unable to resist. By 2007, an estimated 73% of the U.S. corn crop was genetically modified or engineered. Uniformity of plant crops has led to pesticide-tolerant species. There is continuous need for further engineering of crops to resist pests and climatic influences, which were previously tolerable because of the capacity for a diverse gene pool to provide protection from adverse influences through the processes of natural selection. Although humans have dramatically increased crop yields, this has occurred at the risk of increased susceptibility to unanticipated pathogens

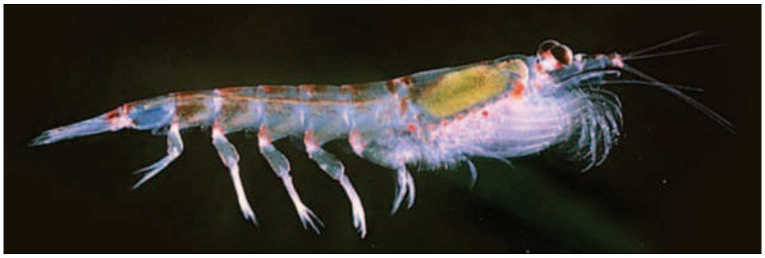


FIGURE 120-9 Antarctic krill. *(Courtesy Uwe Kils, Creative Commons, Share Alike 3.0.)*

and environmental extremes as a result of loss of indigenous strains that were well adapted to local ecosystems.[35]

Fragmentation of habitats caused by urbanization and deforestation has prevented species from interbreeding because of the loss of historic migration corridors, leading to genetic bottlenecks. Species ranging from the grizzly bear in the Central Canadian Rockies to the Florida panther in the Everglades face dangers associated with increased susceptibility to disease and genetic mutation caused by inbreeding of small populations. Loss of vitality in these species affects predator-prey relationships and ultimately contributes to spread of human pathogens (e.g., *Borrelia burgdorferi,* the causative agent of Lyme disease).

Genetic uniformity among honeybees has resulted in problems with reproduction and disease. A condition known as *colony collapse disorder* has led to an alarming die-off among honeybee colonies. This is a significant concern to beekeepers, who have witnessed decline of 30% to 90% of their hives in some parts of the United States. The disorder has been associated with lack of genetic diversity in managed beehives. The potential effect on apple, peach, soybean, and other honeybee-dependent crops could be devastating if these keystone pollinators continue to decline in number.

INVASIVE SPECIES

Introduction of species that are not indigenous to a particular ecosystem has become a global threat to biodiversity. This has posed a historic challenge to human health. The spread of disease to susceptible native populations led, in part, to the downfall of Mesoamerican civilizations as far back as the 16th century. Some alien species were intentionally introduced with intent to improve the local environment without awareness of possible negative repercussions. Africanized bees were imported to improve honeybee productivity in tropical regions of South America. Nile perch were introduced in an attempt to control aquatic weeds, and Brazilian pepper was planted as an attractive ornamental. In each case, lack of species competition in the introduced community allowed the invasive species to overwhelm the native species, with deleterious consequences for the entire ecosystem.

Zebra and quagga mussels are native to the Black and Caspian seas and are believed to have been inadvertently introduced to the Great Lakes region of the United States by emptying of ballast water from transatlantic commercial ships. Prolific female zebra and quagga mussels can produce up to 1 million eggs per year. These mollusks are clogging intake pipes that supply municipal water in Great Lakes cities. Both native mussel populations and freshwater ecosystems are threatened by these invasive species (Figure 120-10). In Florida, 10 of the 16 native bromeliad plant species are listed as threatened or endangered. The Mexican bromeliad weevil, inadvertently introduced by means of infested imported plants, represents a serious challenge to these vulnerable plants. Extracts from bromeliads of the *Tillandsia* genus exhibit analgesic and antiviral properties (Figure 120-11). The danger of invasive species predation is exemplified by the risk of extinction of potentially beneficial bromeliad species. Exotic species, such as Burmese and African Rock pythons, invaded the Florida Everglades after pet owners released their animals into the wild. It is estimated that the Everglades are infested with a population of more than 10,000 of these snakes. Pythons compete with native species for food and habitat resources and represent a threat to several endangered species in the region (Figure 120-12). Approximately 50,000 species of introduced plants, animals, and microbes cause more than $120 billion in annual damages and control costs in the United States.

FIGURE 120-11 Florida bromeliad. *(Courtesy Richard Salkowe.)*

PUBLIC HEALTH CONCERNS

“More than any other biodiversity-related issue, effects on human health may make saving biodiversity an important societal goal.”[35]

FIGURE 120-10 Zebra mussels. *(Courtesy GerardM, Creative Commons, Share Alike 3.0.)*

FIGURE 120-12 Burmese python and American alligator. *(Courtesy Lori Oberhofer, National Park Service.)*

The Wilderness Medical Society has defined five areas that relate biodiversity loss to public health concerns:

1. Altered epidemiology of diseases
2. Loss of biologic raw materials
3. Loss of models for medical research
4. Threatened food production
5. Threatened water resources

ALTERED EPIDEMIOLOGY OF DISEASES

Degraded ecosystems lose capacity to resist the challenges of disease and pestilence. The roles of invasive species, climate change, genetic uniformity, and altered predator-prey relationships have been previously highlighted with respect to relationships between biodiversity loss and disease epidemiology. Lyme disease exemplifies the ability for pathogens to spread in environments where an altered habitat disrupts normal relationships between predators and prey. Spread of *B. burgdorferi* has been enhanced by increased prevalence of the disease-carrying tick vector *Ixodes scapularis*. Altered habitats associated with urban sprawl and altered predator-prey relationships have allowed a burgeoning population of competent disease reservoirs in white-footed mice and deer that the female ticks use for blood meals before egg laying.

Nutrient runoff from fertilization and animal waste has become a significant concern with respect to biodiversity loss caused by eutrophication of nutrient-laden waters, with subsequent formation of harmful algal blooms (Figure 120-13). The Mississippi Dead Zone is an area where algal blooms have formed as a result of nitrogen and phosphorus runoff from fertilized cornfields in the U.S. Midwest. Decaying algal blooms lead to proliferation of oxygen-dependent bacteria. Resultant hypoxic waters, with 90% depletion in oxygen level, threaten vitality of the shrimp industry in the Gulf of Mexico. In addition, areas of high algal growth and warm water serve as ideal reservoirs for pathogens such as *Vibrio cholerae*. For areas in which hypoxic conditions are insufficient to affect productivity of copepod species, algal blooms are associated with a greater number of cholera-carrying copepods. Naturally occurring hydrocarbon-consuming bacteria have been active in clearing the remaining oil in the Gulf of Mexico after the 2010 *Deepwater Horizon* spill. These oil-related bacterial blooms are associated with 35% depletion in oxygen levels. The adverse effect on oxygen-dependent ocean species in the region remains under investigation.

LOSS OF BIOLOGIC RAW MATERIALS

Natural environments provide a vast pharmacologic potential: 57% of the 150 top prescription drugs sold in the United States in 1993 are in some way linked to natural products.[35] The World Health Organization (WHO) estimates that plant medicines account for more than $60 billion in worldwide sales. Revenue from chemotherapeutic drugs derived from the plant species *Taxus baccata* was $2.3 billion in 2000.[35] Many prescription medications are derived from plant products. The foxglove plant *Digitalis purpurea* is the source of digitalis. Deadly nightshade *Atropa belladonna* is the original plant source for atropine (Figure 120-14). Taxol is derived from bark of the threatened Pacific yew tree *Taxus brevifolia*. The Madagascar rosy periwinkle *Catharanthus roseus* is the source of vincristine and vinblastine[35] (Figure 120-15).

FIGURE 120-14 Foxglove flowers. *(Courtesy Jensflorian, Creative Commons Share Alike 3.0; and BerndH, Creative Commons Share Alike 2.5.)*

LOSS OF MODELS FOR MEDICAL RESEARCH

Species decline is associated with loss of biologic models that may help to understand human physiology and disease. Ancrod, a defibrinogenating agent derived from venom of the Malaysian pit viper, is being investigated as a treatment for acute ischemic stroke. Understanding the renin-angiotensin system evolved

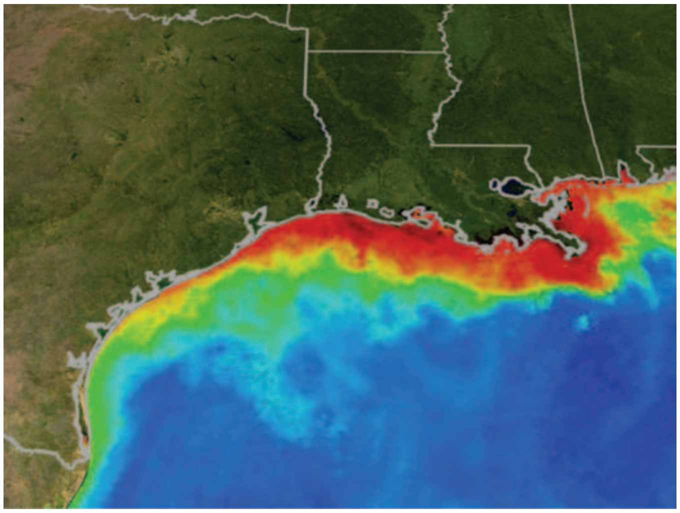


FIGURE 120-13 Mississippi Dead Zone. *(Courtesy National Aeronautics and Space Administration. Public domain.)*

FIGURE 120-15 Madagascar rosy periwinkle. *(Courtesy titanium22, cc-by-sa-2.0.)*

FIGURE 120-16 Malaysian pit viper. *(Courtesy Al Coritz, Creative Commons, Share Alike 3.0.)*

from study of South American pit viper venom that contains angiotensin-converting enzyme (ACE) factors. This led to development of synthetic ACE inhibitors[35] (Figure 120-16). The Gila monster of the southwestern United States is a venomous lizard that produces a glucagon-like peptide in its saliva (Figure 120-17). The synthetic version of this protein is the incretin mimetic, exenatide.

THREATENED FOOD PRODUCTION

At present, 15 crop plants provide 90% of the world's food-energy intake. High levels of monoculture farming and fertilizer use, in conjunction with genetic engineering of crops, have increased yield and resistance to pests and disease-causing organisms. However, loss of naturally acquired immunity to the full range of deleterious environmental factors that have affected our food crops is a concern. High-yielding varieties of genetically modified crops rely on heavy inputs of water, fertilizer, and pesticides. More than 900 pests now resist more than one pesticide. Pesticides also kill the natural enemies of pests, permitting an especially large explosion of those pests.[35]

FIGURE 120-17 Gila monster. *(Courtesy Jeff Servoss, U.S. Fish and Wildlife Service.)*

Use of chemical dispersants as part of the containment response to the 2010 *Deepwater Horizon* oil spill disaster in the Gulf of Mexico has potential long-term deleterious effects on the region's ocean, coastal, and estuarine ecosystems. The Gulf of Mexico produces 25% to 30% of the annual seafood harvested in the United States, including 59% of the oyster production and 75% of the wild shrimp catch.[30] Scientists at the Dauphin Island Sea Lab, the state of Alabama's marine science education and research laboratory, are investigating the role of the coastal region's biodiversity in mitigating the adverse consequences of the event. The biopsychosocial impact on human health associated with any significant loss of biodiversity in this region extends from economic hardship to potential illness caused by environmental contaminants in the food chain.

THREATENED WATER RESOURCES

Water pollution from pesticides and industrial byproducts containing polychlorinated biphenyls (PCBs), bisphenol A, and phthalates has been associated with compounds that interfere with the function of endocrine hormones, resulting in developmental abnormalities and altered reproductive capacity. Alligator populations in Lake Apopka, Florida, have been in decline; nitrate-related endocrine disrupters may be the cause of observed abnormalities in sexual development.[14] The potential for endocrine disrupters to affect human health has been recognized, although there is no confirmatory evidence that these pollutants have definitively altered human reproductive function.[26]

CASE STUDY

The combined effects of ecosystem degradation and species decline as a risk to human health have raised concerns about one of the great river systems of the world. The Colorado River basin is considered the lifeline of the southwest. More than 20 million people depend on the Lake Powell portion of the river system for water and electricity. Lake Powell was formed between 1963 and 1980, when Glen Canyon slowly filled with backflow from the Colorado River after completion of the Glen Canyon Dam. This massive project was a component of the U.S. Bureau of Reclamation's plan to provide water and electricity to the growing population of the southwestern United States. The Glen Canyon and Colorado River ecosystem had previously developed over the course of 5 million years. However, during the span of 17 short years, it was drastically altered by the dam project. Examining this situation provides a window into the consequences of such an endeavor with respect to biodiversity and human health. Extinct and threatened native species in this region serve as indicators of ecosystem disruption and raise concerns pertaining to sustainability of a drastically altered natural environment.

"The native fish of the Colorado River system make up one of the most unusual assemblages of fish specially adapted to their environment found anywhere in the world."[29] The unique combination of geologic, hydrologic, and climatic factors that were found in the predam Colorado River basin led to natural-selection processes resulting in peculiar morphologic and behavioral characteristics in the native fish. Adaptations of the fish to extreme and severe river conditions include large streamlined bodies, large fins, and thick skin.[20] Larger fish of the Colorado River live an exceptionally long time and have depressed skulls with large predorsal humps and small eyes.[19] Niche partitioning for available resources led to development of structural variations within the fish native to this region, such as razorback sucker fish with "protrusible mouths and special gill rakers for sieving plankton or detritus."[19] Some of the native fish fauna have existed for 20 million years.[20] Speciation in this region was associated with demands placed on the aquatic biota. The native warm-water fish adapted to the "challenges of living in a highly variable environment subject to seasonal extremes of flow and water temperature, short term flow changes from local storm events, and highly turbid conditions"[29] associated with high-volume sediment transport. The previously mentioned unique morphologic features were most likely adaptations to the high-flow,

FIGURE 120-18 Humpback chub. *(Courtesy Melissa Trammell, National Park Service. Public domain.)*

sediment-laden waters of the Colorado River ecosystem. There are eight species of native fish in Glen and Grand Canyons, and six are endemic to the area: humpback chub (endemic); razorback sucker (endemic); Colorado squawfish (endemic); bonytail chub (endemic); roundtail chub (nonendemic); flannelmouth sucker (endemic); bluehead sucker (endemic); and speckled dace (nonendemic).[19] Four of the native fish are federally listed endangered species: humpback chub, razorback sucker, Colorado squawfish, and bonytail chub (Figure 120-18). Three of the native fish species are believed to be extirpated from the Glen and Grand Canyons: Colorado squawfish, bonytail chub, and roundtail chub.[29] Three native fish species retain relatively stable populations in the region: flannelmouth sucker, bluehead sucker, and speckled dace. The flannelmouth sucker is listed as a federal endangered species candidate and is protected in Arizona as a result of concerns regarding species decline.

Minckley[19] and Smith[28] have classified the native fish species of the Grand Canyon as dietary generalists. However, Minckley,[19] in reference to his earlier work, "characterized trophic relations in the native fish species based on qualitative and quantitative differences in selected foods and spatial segregation in feeding. Adult squawfish were piscivorous (fish eating). Flannelmouth suckers fed on insects and other benthic (bottom dwelling) animals. Bluehead suckers were adapted for scraping algae. Razorback suckers fed on detritus and plankton. Humpback chubs, speckled dace, and bonytail chub tend to be insectivores; although speckled dace have exhibited facultative omnivorous behavior." Classification of the native fish of this region as dietary generalists is tempered by empirical evidence of specialized structures and feeding behaviors that developed in response to specific demands of the river environment in this region.

"Before completion of the Glen Canyon Dam, the Colorado River was thought to be largely heterotrophic with little primary production in the sediment laden water."[4] Schmidt and colleagues[24] made note of the "allochthonous pre dam aquatic system," which refers to limited primary production of algae and existence of nutrient resources transported downstream by the river flow. This is confirmed by dietary habits of the native species, because the only fish (i.e., bluehead sucker) that used primary-producing algae as a food source was also able to sustain itself on other food sources. The predominant primary producer in the Glen Canyon and Grand Canyon aquatic ecosystem is the filamentous green algae *Cladophora*. The turbid sediment-laden waters of the Colorado River prevented sufficient penetration of sunlight to create an ideal environment for primary producers. Webb and colleagues[34] note that "the green alga *Cladophora glomerata* was present but not abundant in the river" before completion of the Glen Canyon Dam. As a result of the limited algal resources, the native fish were dependent on carbon-based nutrients that were transported by river current. This process was contingent on scour of upstream flood-level algal remnants, diatoms, invertebrates, insects, fish, and plant debris from native riparian vegetation (e.g., willow, cottonwood).

A survey of Glen Canyon in 1958, before building the Glen Canyon Dam, revealed 17 species of fish, including 10 non-native types.[19] Intrusion of non-native fish is a significant factor in the demise of the native fishery. "Remarkable numbers of non-native fishes have been intentionally and inadvertently stocked into the Colorado River."[19] At least 20 species of non-native fish were planted in Utah before 1900.[27] Additional non-native fish were added for sport fishing, vegetation control, and as bait for the sport fishery. Minckley[19] states that "non-native fish have invaded essentially every habitat." Accidental introduction of gizzard shad during a 1998 largemouth bass stocking in Morgan Lake in New Mexico led to presence of this species in Lake Powell. This non-native shad species, also known as *stink shad,* may provide a food resource for the Lake Powell non-native striped bass fishery. However, there is significant concern that potential downstream spread of gizzard shad below the dam will compromise the food supply for the remaining native fish because of resource competition. Resultant competitive and predatory exclusion by non-native species, including channel catfish, trout, and voracious sunfish, has had a deleterious effect on native fish species. It is evident that even before development of the Glen Canyon Dam, there was significant pressure on native fish resources. The dam compounded the challenge to extant native fish.

A seasonal warm-water body with intermittent periods of inundation by floodwaters was immediately transformed into a constant cold-water river (i.e., 10° C [50° F])[21] with diurnal variations in flow and limited controlled flooding caused by opening the dam locks. The backwater and eddy habitats of native fish were compromised with respect to flood-induced habitat management, and the previously mentioned carbon-based debris-dependent ecosystem had been transformed. Clear discharge from the dam was established as a result of retention of sediment in the newly established Lake Powell. This provides for increased penetration of sunlight and significant increase in the primary-producing algal life-form *Cladophora*. Detritus and carbon-based debris flow were affected by the comparatively limited alteration in flow rate after the dam and the limited scour of higher-altitude vegetation that occurred from predam floods ranging from 2330 to 6230 m^3/sec.[33]

The Glen Canyon Dam has created a compromise in the historic food supply, shelter, and thermal gradient conducive to spawning for the native fishery. Combined with predation and competition from non-native species, these adverse factors have led to the previously mentioned extirpation of three native species and "endangered" listing of two of the five remaining species. Valdez[31] summarized that as a result of the Glen Canyon Dam, several native fish species in this area are susceptible to "major threats from flow depletion, altered water chemistry, flooded habitat from reservoirs, introduced parasites and diseases, competition and predation from introduced non-native fish."

Introduced species have fared better in the postdam environment. Cold, clear outflow from the Glen Canyon Dam provides excellent conditions for growth of *Cladophora,* which is a food source for the freshwater scud *Gammarus lacustris*. Scud, annelids, and midge species are the primary dietary choices for rainbow trout in the area below Glen Canyon Dam. "The Lee's Ferry water below Glen Canyon Dam holds an estimated 50,000 trout over six inches long (17,000 over 12 inches) per mile in over 15 miles of water, according to the Arizona Game and Fish Department."[9] Daily change in water-flow release associated with electrical demand creates ideal feeding conditions for trout. When water levels drop significantly, millions of scuds become stranded on exposed gravel bars lining the banks. After 2 days of drying in the hot Arizona sun, the desiccated scuds are flushed into the current when the water rises again. As the dried shells float downstream, live fish go on a binge.[9] Although this carbon-based detritus would seem ideal for native fish accustomed to allochthonous food resources, cold-water temperature in the area between the dam and Lee's Ferry is not a conducive habitat or spawning ground for the native fishery. This leaves the abundant scour resource for the prime benefit of the cold-tolerant trout species.

The dam resulted in artificial creation of a prime non-native fish habitat in Lake Powell and the Grand Canyon. However, there is evidence that this is subject to a limited time period. "Lake Powell, an artificial reservoir on the Colorado Plateau, is beginning to suffer from high selenium levels in its sediments."[13] Two independent studies of largemouth bass from Lake Powell report that selenium concentrations in the fish greatly exceed national averages. This abnormally high selenium concentration in fish reflects the high concentration in the reservoir[8] and is a result of the dam blockage of normal downstream sediment transport. A growing body of literature continues to document extensive contamination of aquatic environments with selenium and the adverse effects in aquatic organisms.[15] Regardless of sediment toxicity, accumulation of a sediment load equal to 100 million tons per year is projected to potentially block the river outlet valves within 100 years.[12] The downstream trout fishery depends on a continued clear river-outlet discharge.

Creation of the dam has resulted in a variety of ecologic impacts regarding fishery resources. The native fish are the most vulnerable as a result of the effects of environmental pollutants, species competition, habitat modification, and gradual accumulation of sediment load in Lake Powell. There are continued efforts to ameliorate the effects of the dam on the native fishery while preserving the economically productive introduced sport fishery. The intractable nature of this environmental balancing act is firmly entrenched in the policy disputes associated with river management in this region. The inevitable fact is that choices and options will become progressively limited as the sediment load in Lake Powell eventually renders the dam ineffective for any economic or ecologic purpose consistent with present or historic use patterns. In addition, evaporative water losses in Lake Powell and potential diminished levels of stream flow associated with climate change could place the 20 million people who depend on this area for water and electricity at extreme risk.

The challenge to biodiversity in Lake Powell and the Colorado River basin is a challenge to human health. Biophysical alteration of a natural environment after construction of the Glen Canyon Dam has resulted in an unsustainable artificial ecosystem doomed to extinction under the pressures of sediment load and accumulation of arsenic, lead, selenium, boron, and mercury from upstream runoff sources. Resultant exposure to environmental contaminants and loss of water, electric, and economic resources will potentially create psychological and physiologic hardships for inhabitants of this region.

CONCLUSION

Lisa Newton[22] defines *sustainability* as activity that "can be maintained profitably and indefinitely without degrading the systems on which it depends." She suggests that no practice will be regarded as sustainable unless it can be continued without degrading the environment that nurtures it though the seventh generation from its initiation. It is evident that our historic land-use practices have led to ecosystem degradation, species decline, and loss of genetic diversity. The consequences of these actions are foreboding for human health and planetary well-being as we are exposed to emergent and resurgent infectious diseases and our food and water resources are placed at potential risk.

When referring to an ecologic conscience, Aldo Leopold[17] stated:

> Obligations have no meaning without conscience, and the problem we face is the extension of the social conscience from people to land. Land ethic, then, reflects the existence of an ecological conscience, and this in turn reflects a conviction of individual responsibility for the health of the land. Health is the capacity of the land for self-renewal. Conservation is our effort to understand and preserve this capacity.

Those of us with an interest in wilderness medicine have a unique opportunity to serve as stewards of the environment and of our patients' well-being by increasing our understanding of the importance of biodiversity with respect to human health.

REFERENCES

Complete references used in this text are available online at expertconsult.inkling.com.

CHAPTER 121

Health Implications of Environmental Change

CAROLYN SIERRA MEYER AND JAY LEMERY

For 11,700 years, Earth's natural processes and related systems have remained relatively stable. For example, weather patterns, nutrient cycles, freshwater repositories, biodiverse forests, and prairie systems have nurtured growth of human societies. It is increasingly recognized that human actions modify these processes. Since the middle of the 19th century, human impact has accelerated, and it now threatens resilience of the planet. Humans burn fossil fuels, fertilize depleted soil, and irrigate deserts. The cumulative impact of environmental stressors directly afflicts humans by causing changing patterns of disease. Climate change and loss of biodiversity, among other impacts, threaten the quality and quantity of human life. Earth's planetary processes have global thresholds, or "tipping points"; crossing these boundaries may lead to changes that are not hospitable for human social structure or even survival. For example, there are limits to food production. Agricultural crops and livestock have physiologic limitations in certain thermal and water stress situations. Staple crops, such as maize, rice, wheat, and soybeans, grow only within the range of 40° C (104° F) to 45° C (113° F).[58] There are thresholds of global warming beyond which current agricultural practices will no longer be able to support large human civilizations. The global risk to food security will become alarmingly great if an increase occurs in global mean temperature of 4° C (7.2° F) to 6° C (10.8° F) or more.[56] Therefore, impacts of

REFERENCES

1. Adler K. Amphibians and humans. In: Hutchins M, Duellman WE, Schlager N, editors. Grzimek's animal life encyclopedia, vol. 6. 2nd ed. Amphibians. Farmington Hills, Mich: Gale Group Division, Thomson Learning; 2003. p. 51–5. <http://www.answers.com/topic/amphibians-and-humans>; [accessed 10.02.16].
2. Balick MJ, Elisabetsky E, Laird SA. Medicinal resources of the tropical forest. New York: Columbia University Press; 1995.
3. Bruno JF, Selig ER, Casey KS, et al. Thermal stress and coral cover as drivers of coral disease outbreaks. PLoS Biol 2007;5:1220.
4. Carothers SW, Brown BT. The Colorado River through Grand Canyon: Natural history and human change. Tuscon: University of Arizona Press; 1991.
5. Chivian E, Bernstein B. Sustaining life: How human health depends on biodiversity. Oxford, UK: Oxford University Press; 2008.
6. Conservation International. The biodiversity hotspots. <http://www.conservation.org/explore/priority_areas/hotspots/Pages/hotspots_main.aspx>.
7. Convention on Biological Diversity. Convention text. Article 2. Use of terms. <http://www.cbd.int/convention/articles.shtml?a=cbd-02>.
8. Dansie T. Uranium mill tailings in the Colorado River basin. <http://www.glencanyon.org/publications/hiddenpassage/hp5uraniumtailings.php>.
9. Deeter K. Lee's Ferry rainbows. <http://flyfisherman.com/southwest/kdleesferry/>.
10. Food and Agriculture Organization of the United Nations. Global forest resources assessment 2005. Progress towards sustainable forest management. <http://www.fao.org/DOCREP/008/a0400e/a0400e00.htm>.
11. Garrett JT, Garrett M. Medicine of the Cherokee. Rochester, Vt: Bear & Co; 1996.
12. Glen Canyon Institute. Sediment and Lake Powell. <http://www.glencanyon.org/library/sediment.php>.
13. Graf W. Selenium dynamics of the Colorado River basin: SCERP Project Number W-13, 1991.
14. Guillette LJ, Edwards TM. Is nitrate an ecologically relevant endocrine disruptor in vertebrates? Integr Comp Biol 2005;45:19.
15. Hamilton S, Lemly D. Water-sediment controversy in setting environmental standards for selenium. Ecotoxicol Environ Saf 1999;44:227.
16. International Union for Conservation of Nature. IUCN Red List version 2010.1: Table 1. Numbers of threatened species by major groups of organisms (1996-2010). <http://www.iucnredlist.org/documents/summarystatistics/2010_1RL_Stats_Table_1.pdf>.
17. Leopold A. A Sand County almanac. New York: Ballantine Books; 1986.
18. Millennium Ecosystem Assessment: Living beyond our means: Natural assets and human well-being. <http://www.millenniumassessment.org/en/BoardStatement.aspx>.
19. Minckley WL. Background Paper 7. Native fishes of the Grand Canyon region: An obituary? From Colorado River ecology and dam management: Proceedings of a symposium, May 24-25, 1990, Santa Fe, NM, Washington, DC, 1991, National Academy Press.
20. Molholland R. Endangered and threatened species in the lower Colorado River basin, delta and Sea of Cortez, Salt Lake City, Utah, 1999, Glen Canyon Institute.
21. National Parks Conservation Association-Center for Park Research. National parks of the Colorado River Basin: Water management, resource threats, and economics. <http://www.npca.org/cpr/colorado_river_basin/Colorado-River-Report.pdf>.
22. Newton L. Ethics and sustainability: Sustainable development and the moral life. Upper Saddle River, NJ: Prentice Hall; 2003.
23. Ornstein R, Ehrlich P. New world new mind: Moving toward conscious evolution. Los Altos, Calif: Malor Books; 2000.
24. Schmidt JC, Andrews ED, Wegner DL, et al. Origins of the 1996 controlled flood in the Grand Canyon—Geophysical Monograph 910. In: Webb RH, Schmidt JC, Marzolf GR, et al., editors. The controlled flood in the Grand Canyon. Washington, DC: American Geophysical Union; 1999.
25. Science Daily: Pollution drives frog deformities by ramping up infections. <http://www.sciencedaily.com/releases/2007/09/070925130649.htm>.
26. Sharpe RM, Irvine DS. How strong is the evidence of a link between environmental chemicals and adverse effects on human reproductive health? BMJ 2004;328:447.
27. Sigler WF, Sigler JW. Fishes of the Great Basin: A natural history. Reno: University of Nevada Press; 1987.
28. Smith ML. Late Cenozoic fishes in the warm deserts of North America: A reinterpretation of desert adaptations. In: Naiman RJ, Soltz DL, editors. Fishes in North American deserts. New York: Wiley & Sons; 1981.
29. US Department of the Interior, Bureau of Reclamation. Operation of Glen Canyon Dam, Colorado River Storage Project, Arizona: Final environmental impact statement. <http://www.usbr.gov/uc/envdocs/eis/gc/pdfs/Cov-con/cov-con.pdf>.
30. US Environmental Protection Agency. Gulf of Mexico program: General facts about the Gulf of Mexico. <http://www.epa.gov/gmpo/about/facts.html#resources>.
31. Valdez RA. Native fishes of the Grand Canyon. <http://gcrg.org/bqr/6-1/nativefish.htm>.
32. Valiela I, Bowen JL, York LK. Mangrove forests: One of the world's threatened major tropical environments. Bioscience 2001;51:807.
33. Webb RH, Melis TS, Griffiths PG, et al. Reworking of aggraded debris fans—Geophysical Monograph 110. In: Webb RH, Schmidt JC, Marzolf GR, et al., editors. The controlled flood in the Grand Canyon. Washington, DC: American Geophysical Union; 1999.
34. Webb RH, Wegner DL, Andrews ED, et al. Downstream effects of Glen Canyon Dam on the Colorado River in Grand Canyon: A review—Geophysical Monograph 110. In: Webb RH, Schmidt JC, Marzolf GR, et al., editors. The controlled flood in the Grand Canyon. Washington, DC: American Geophysical Union; 1999.
35. Wilderness Medical Society. Biodiversity loss impact on human health, Educational lecture series, 2010.
36. Wilson EO. Biophilia and the conservation ethic. In: Penn D, Mysterud I, editors. Evolutionary perspectives on environmental problems. Piscataway, NJ: Transaction Publishers; 2007.
37. World Wildlife Federation. Living planet report. <http://assets.panda.org/downloads/living_planet_report_2008.pdf>.

suboptimal nutrition and malnourishment may become severe as planetary temperature rises.

The wilderness medicine practitioner has a dual mandate to promote both patient wellness and a healthy environment. This chapter serves as a primer for understanding growing threats to human health caused by a changing environment. It addresses human-imposed impact on Earth's natural environmental processes and the effects on human health.

CLIMATE CHANGE

There is overwhelming scientific consensus that anthropogenic climate change is accelerating[23] (see Chapter 119). The Intergovernmental Panel on Climate Change (IPCC) is the leading international body for assessment of climate change. In 1988, the United Nations Environment Programme and the World Meteorological Organization established the IPCC to provide the world with a clear scientific view on the current state of knowledge regarding climate change and the potential environmental and socioeconomic impacts. Thousands of scientists worldwide contribute on a voluntary basis to these quadrennial reports. The thresholds for data to be included in the IPCC surpass that of contemporary peer-reviewed journals, and the science encompassed is considered to be of exceedingly high quality. The IPCC 2014 Synthesis Report states that anthropogenic "greenhouse" gases—atmospheric concentrations of carbon dioxide, methane, and nitrous oxide—have exponentially increased beyond historical cyclic variations measured over the past 800,000 years. Their effects are believed to be the dominant cause of the observed warming since the middle of the 20th century.[44] High levels of these greenhouse gases are linked to increases in intensity, frequency, and duration of heat waves,[48] melting of the Greenland and Antarctic ice sheets,[54] and increases in precipitation extremes, including heavy rainfall in some regions and drought in others.

VULNERABLE POPULATIONS

Climate variability and change do not affect all people equally. The vulnerable populations, whether measured in socioeconomic or demographic (e.g., extremes of age) terms, will bear a disproportionate burden of climate-related health effects.[27] Working Group II of the IPCC finds that regions of Africa with poor governance, tenuous public health and health care systems, and food and water insecurity will suffer the most from global climate change.[50] Factors that increase vulnerability include inadequate or no mosquito protection and limited to no access to health care facilities. Intense heat waves will increase mortality and morbidity in elderly people and persons with preexisting medical comorbidities. Increases in heavy rainfall and temperature will increase the risk of diarrheal diseases, dengue, and malaria, with the effects of these compounded by poor public health infrastructure. Increases in floods and droughts will exacerbate rural poverty in parts of Asia through negative impacts on certain crops, such as rice, resulting in increases in food prices with associated effects on nutrition.

Vulnerability is multifactorial, and data have clarified the etiology of climate exposure. Geographic location influences deterioration of health caused by climate change. Warming most affects persons who work outdoors in hot temperatures at the limits of thermal tolerance.[25] Populations in proximity to the present limits of the range of transmission of vector-borne diseases are most vulnerable to changes attributed to rising temperatures.[31] Communities situated on low-lying coral atolls experience the health impacts of soil salination, flooding, and freshwater reservoir contamination because of sea level rise.[41]

Age and gender are factors in the loss of health caused by climate change. Children and elderly persons are at increased risk for climate-related injuries and illnesses. Children are more physiologically susceptible to the destructive effects of malaria, diarrhea, and poor nutrition, all of which show increases from climate change.[35] Elderly people have a limited ability to respond to physiologic stressors, such as heat and air pollution, because they often have preexisting health conditions.[17] They also find it difficult to avoid the hazards and destruction of floods, heat waves, and other extreme natural events because they tend to be less agile and less aware than are younger adults.

According to the World Health Organization (WHO), worldwide mortality from natural disasters, including droughts, floods, and storms, is higher among women than men.[60] Pregnant women in particular are at increased risk because they are more vulnerable to extreme heat, malaria, food-borne infections, and influenza.

Financially poor populations are at increased risk of loss of health caused by climate change. Mortality risk associated with tropical cyclones from 1970 to 2009 showed dependence on three major factors: storm intensity, quality of governance, and poverty level.[47] A study on the impacts of flooding in Bangladesh found that as average income and number of income sources increased, household risk was reduced. Poorer households took preventive action less often, were more severely affected by flooding, and received less assistance after the flooding than wealthier households.[7]

Indigenous peoples, populations of small island developing states, and resource-poor urban communities in developed countries, including the United States, are also particularly vulnerable to negative health impacts from climate change.[50]

DIRECT IMPACTS OF CLIMATE CHANGE ON HUMAN HEALTH

HEAT-RELATED HEALTH IMPACTS

Global warming is a relative misnomer. Although the past few years have been the warmest on record, some of the coldest and stormiest winter weather has occurred in the northeastern United States.[40] Climate denialists purport this to be evidence of conflicting data as to whether or not the earth is warming. A more accurate phrase would be *global energizing,* because as indeed we are consistently measuring the warmest years on record, the effects are not equally distributed in time or space. Local weather and temperature remain highly variable.

A *heat wave* is a prolonged period of excessively hot weather, beyond the normal seasonal temperature for a particular region. The association between hot days and increased morbidity and mortality is well established.[21] In Australia between 1968 and 2010, the ratio of summer-to-winter deaths increased in association with rising annual average temperatures.[6] Studies based on hospital admissions or emergency medical presentations during heat waves report increases in temperature-related morbidity, attributed to such conditions as cardiovascular, respiratory, and kidney diseases.[39] A Harvard School of Public Health study followed a U.S. cohort older than 65 with chronic disease from 1985 to 2006 and demonstrated reduced survival associated with greater variability of temperature. Dramatic short-term fluctuations in temperature variability have been shown to increase the risk of mortality.[65] Again highlighting vulnerability in socioeconomic groups, access to ventilation and the presence or absence of air conditioning contribute to the effects of high temperatures on humans. The 2003 Northern European heat wave contributed to more than 12,000 deaths in Paris alone, the vast majority being elderly persons of the lower socioeconomic neighborhoods without climate controls or preexisting knowledge of such unprecedented extreme weather.

The "urban heat island" effect is another well-documented result of heat waves in urban settings. Poorer neighborhoods without open parks or shaded areas have a significantly higher temperature change compared to average temperature than more affluent neighborhoods.[18] Health risks during heat extremes are greater in people who are physically active. Climate change especially impacts the health of manual laborers and persons who pursue outdoor recreation.[24]

FLOODS AND STORMS

Climate energizing through warming impacts frequency and severity of many weather events, including precipitation, drought, and cyclones. One such mechanism occurs when warmer ambient

FIGURE 121-1 NOAA's GOES-13 satellite captured this visible image of the massive Hurricane Sandy on October 28, 2012, at 1302 UTC (9:02 AM EDT). *(Courtesy National Oceanic and Atmospheric Administration [NOAA].)*

temperatures increase evaporation, leading to higher absolute humidity. When more water is carried as vapor in the atmosphere, rain is less frequent, but rainstorms become more intense because the amount of water is greater in the storm clouds (Figure 121-1). Most climate models predict longer periods of drought interspersed with more intense rain events.

Floods are the most common extreme weather event. They affect and kill more people than any other form of natural disaster. The Center for Research on the Epidemiology of Disasters collects yearly data and reports that in 2011, six of the 10 largest natural disasters were flood events; 112 million people were affected and 3140 deaths were directly attributed to flooding.[19] Conservative estimates for health impact caused by storms and flooding suggest that 2.8 billion people were affected between 1980 and 2009, with more than 500,000 deaths. Worldwide, the frequency of river flooding has increased and caused greater economic losses because more population and property are present in flood plains. Flooding and storms cause death and disease through drowning, traumatic injury, hypothermia, and increased transmission of infectious disease. Flooding often causes groundwater contamination with feces, dead livestock, and chemicals leached from industrial facilities. Diarrheal disease, leptospirosis, insect vector–borne diseases, and cholera all increase in prevalence during and directly after flood events.[53]

Flooding has long-term implications on mental health. In one example involving a 2007 flood in Wales, United Kingdom, the prevalence of mental health symptoms, such as psychological distress, anxiety, and depression, was two to five times higher among individuals who reported flooding in their homes than among nonflooded individuals.[45]

HEALTH EFFECTS MEDIATED THROUGH NATURAL SYSTEMS

Global warming indirectly impacts human health through environmental and ecosystem changes. These changes include shifts caused by warmer conditions in territories and ranges of disease-carrying mosquitoes and ticks, more waterborne diseases, and increased precipitation and runoff.

VECTOR-BORNE DISEASES

Anthropogenic climate change impacts the burden of vector-borne diseases. A major determinant of the endemic range of vector-borne diseases is seasonal temperature. As the climate warms, fewer pathogens languish in the cold, so the distribution of certain diseases expands. Mosquitoes and ticks primarily transmit vector-borne infectious diseases, such as malaria and dengue. Warmer temperatures increase metabolic rates of mosquitoes and ticks, affecting their nutritional requirements and increasing the drive to feed more frequently. Transmission potential of pathogens increases accordingly. Global warming directly affects survivability of pathogens and indirectly affects the vectors and reservoirs that harbor pathogens.[5]

Malaria (see Chapter 40)

Malaria is mainly caused by one of five parasites: *Plasmodium falciparum, P. vivax, P. malariae, P. ovale,* and *P. knowlesi*. Anopheline mosquitoes transmit the parasites. In 2013, there were an estimated 283 million cases of and 584,000 deaths from malaria worldwide, mostly among children younger than 5 years.[62] Increases in average planetary temperature and associated increases in precipitation likely favor malaria transmission. Warmer conditions will prolong malaria seasons and allow mosquito migration to higher latitudes, infecting populations not traditionally at risk. In 2014, researchers were shocked to discover malaria in Alaskan birds for the first time, as well as a recrudescence in southern Italy for the first time in 50 years.[30,52]

With increased average daily temperatures, areas previously below the lower limit of the range of viability for survival of the pathogens show increased transmission. However, the increase is not linear.[43] Even modest warming may drive increases in malaria transmission if conditions are otherwise suitable.[1]

Other Viral Diseases (see Chapter 39)

The incidence of dengue viral disease ("fever") has grown dramatically in recent decades. WHO currently estimates that there may be 50 to 100 million dengue infections worldwide every year. An estimated 500,000 people with severe dengue require hospitalization each year, many of whom are children. Approximately 2.5% of affected persons expire.[63] *Aedes aegypti* and *Aedes albopictus* mosquitoes transmit the virus to humans through bites during feeding periods. These mosquitoes are climate sensitive. Over the last few decades, climate conditions in certain areas have become more suitable for *A. albopictus* mosquitoes. A new serotype of dengue affecting a previously nonimmunized population was identified in Portugal in 2012.[2] This was in part attributed to climate change, according to the IPPC Fourth Assessment Report (AR4), because of an increase in the percentage of days per year with favorable temperature for disease transmission.[22] In 2013, new dengue cases occurred in Florida and the Yunnan province of China, two regions where it had not previously been observed; one report showed that the distribution of *A. albopictus* is highly correlated with annual temperature and precipitation.[64] Another study demonstrated that dengue incidence increased in Guangzhou, China, in association with temperature, humidity, and rainfall, and that wind velocity is inversely associated with rate of disease.[32]

Both typhoons and droughts affect vector populations and increase the incidence of infections. Typhoons bring extreme rainfall, high humidity, and water pooling, generating new mosquito breeding sites. Drought causes increases in rates of disease if households store water in containers that provide suitable mosquito breeding sites.

Other Vector-Borne Diseases

Hard ticks of the family Ixodidae transmit tick-borne encephalitis virus to humans. Europe and Asia have temperate regions where the disease is endemic, and climate change has resulted in expansion of Ixodidae tick territory and prolonged the season during which the ticks transmit disease.[5] During the 1970s, tick-borne encephalitis became more prevalent in central and eastern Europe. As reported in a study describing the Czech Republic, warm spring temperatures between 1970 and 2008 encouraged transmission of the tick-borne encephalitis virus. The transmission season lengthened, and disease spread to higher altitudes.[26]

Europe, Canada, and the United States are home to ticks infected with *Borrelia burgdorferi*, which causes Lyme disease. Many studies have shown associations between tick-borne diseases and climate. In North America, based on active and passive surveillance data, there is good evidence of northward expansion of the distribution of the *Ixodes scapularis* tick vector by 2060.[42]

This list correlating increased disease risk to expected climate change is extensive and a formidable challenge to public health. For instance, hantavirus causes rates of infection correlated to increases in temperature, precipitation, and relative humidity.[28] Plague has been linked to seasonal and internal variability in climate.[33] Other vector-borne diseases linked to climate variability include chikungunya fever, transmitted by the same mosquitoes (*A. aegypti* and *A. albopictus*) that transmit dengue virus; Japanese encephalitis, transmitted by *Culex tritaeniorhynchus* mosquitoes; and Rift Valley fever, transmitted by both *Aedes* and *Culex* mosquitoes.

WATERBORNE DISEASE

Anthropogenic climate change increases exposure to climate-sensitive waterborne pathogens. Warmer climate and increased severity and frequency of storms will result in a more sustainable habitat for pathogens and greater opportunity for mixing contaminated water sources with drinking water and agriculture. Most of these pathogens are introduced to the water by human and animal feces, which is directly related to poor sanitation and exacerbated by extreme precipitation and flooding, because latrines for human and animal waste often become comingled.

Most acute exposures to waterborne pathogens result in symptoms of gastroenteritis. These illnesses are generally self-limited, with the majority of symptoms lasting less than 1 week. However, some, such as cholera, can be devastating. Chronic exposure to waterborne pathogens carries long-term health consequences. Children are particularly at risk, because each episode of illness may jeopardize healthy growth by reducing caloric uptake and nutrient absorption. This problem is most profound with severe diarrheal diseases such as typhoid and dysentery, which progress from diarrhea to systemic syndromes with high rates of death.

CLIMATE CHANGE AS A THREAT MULTIPLIER

Every wilderness medicine practitioner should understand not only the direct causality of climate change on human health, but also the indirect effects. Undernutrition, mental illness, occupational health effects, food insecurity, and increases in violence and conflict may not be primarily caused by climate change, but all will be made worse by its consequences.[9]

MENTAL HEALTH

The American Psychological Association recently released a report on the broad psychological effects of climate change.[14] Psychological stress derives from both abrupt changes experienced during natural disasters and more gradual changes in the local environment. Natural disasters cause severe psychological trauma through personal injury, loss of family and friends, and loss of personal property or livelihood.[37] People who have recently experienced an acute trauma have high levels of distress and anxiety and may report panic attacks, difficulty sleeping, low motivation, and obsessive behavior.[12] This may lead to more long-term psychopathology, such as posttraumatic stress disorder (PTSD) or major depressive disorder. PTSD has been extensively documented in survivors of the urban flood from Hurricane Katrina and is linked to higher levels of substance abuse, depression, anxiety, and suicide. Disaster events produce strains on social relationships and may lead to forced migration, which is a devastating stressor because families and friends are separated and lose their systems of social support and access to primary health care.

More gradual effects of climate change, such as prolonged droughts and increases in mean temperature, also negatively affect mental health. Increases in average temperature are associated with increased use of emergency mental health services, and warmer weather conditions increase the stress on people who already suffer from mental illness, exacerbating their disease and overwhelming their coping ability.[59] Even for people without mental illness, climate change is an additional source of stress and can affect some people deeply, causing feelings of loss, helplessness, and frustration because of inability to prevent the foreseen changes and their disastrous effects.[36]

NUTRITION

The effects of climate change on human nutrition are complex. Agricultural production, food prices, access to food, and human disease all impact nutrition and are expected to be affected by climate change. Extreme weather events, such as floods, droughts, and heat waves, exacerbate food insecurity. The IPCC concludes that climate change will have a substantial negative impact on per-capita calorie availability. Climate change will negatively impact childhood nutrition, particularly stunting of growth, and increase child deaths in developing countries.[22] Crop yield gains from warmer weather in the fields of Russia and Canada will not make up for loss of productivity in the global south, and the net result will be reduced quantity and quality of food harvested.[4] In most tropical regions, climate effects on crop yields are significant. For each degree above 30° C (86° F), African maize yields decrease by 1% under normal rain conditions and by 1.7% under drought conditions.[29] A study in Africa and South Asia revealed an 8% average yield reduction in crops of wheat, maize, sorghum, and millet as temperature increased and rainfall patterns changed.[58a]

VIOLENCE AND CONFLICT

Populations that are affected by violence and with weak civil institutions and poor governance are particularly vulnerable to the health impacts of climate change. The insecurity and full-scale natural disasters brought on by extreme weather events cause societal deterioration through exacerbation of existing or creation of new poverty situations. This adds to social strife and may escalate regional conflicts in parts of the world with little domestic capacity for resilience or adaptive response to a climate stressor.[16]

BIODIVERSITY LOSS

Destruction of species and mismanagement of natural ecosystems worldwide destabilizes the physical environment, increases vulnerability to the spread of human infectious disease, and promotes proliferation of pathogens that can affect the food supply and natural resources on which human health and well-being depend. Consider the vulnerability of agricultural monocultures. Pathogens spread more easily, and epidemics tend to be more severe, when the host plants (or animals) are more genetically homogeneous and crowded. Outbreaks of disease, insect infestations, and climatic anomalies pose a greater threat to monospecies ecosystems than to diverse ones, causing widespread crop and animal failures, undermining food security, and accelerating spread of diseases to human populations.

Biodiversity can be viewed in two ways, both of which are shrinking. *Genetic diversity* is diversity of genes within a species and functions as an information bank that determines the potential for life to evolve and adapt as the environment changes. Genetic diversity is difficult to measure on a global level, so as of now, global data do not exist for phylogenetic species variability. Species naturally go extinct at a rate of one to five species per year. Anthropogenic changes to ecosystems have increased the baseline extinction rate. The planet now loses species at a dramatically increased rate of 1000 to 10,000 species per year.[11]

Functional diversity concerns the behavior and effects of organisms in communities and ecosystems. Functional traits make up functional diversity. Functional traits are measurable aspects of an organism that reflect what it does and how it interacts with other organisms and its environment (e.g., size, diet, behavior). Functional groups are a set of species showing either similar responses to the environment or similar effects on major ecosystem processes. Functional diversity refers to the abundance of functional groups. For example, compare a section of tide pool that contains three different species of barnacle, and another section that has a starfish, sea grass, and an anemone. Each has three different species. However, the second section of tide pool has more functional diversity because each species interacts with the other organisms and the environment differently, and it also has different effect(s) on the ecosystem. Functional diversity allows for a resilient ecosystem, in this example, one that can withstand pathogens, invasive species, and extreme weather events and can provide continuous ecosystem services (e.g., clean water) for other sea life to thrive.

Some of the greatest pharmaceutical discoveries have come from nature. These include antiinflammatory, chemotherapeutic, antibiotic, and thrombolytic drugs. Our ecosystems are the raw materials for future discoveries in pharmaceuticals and biotechnology. The continued loss of biodiversity, to quote John Dingell, is "akin to burning the library without ever having read its books."[8a]

THREATS TO ECOSYSTEM SERVICES

Deforestation

Deforestation not only changes the climate, but also has direct effects on human health. As forest habitats are cleared for agriculture and urban development, human-wildlife interaction and conflict grow. Deforestation decreases the habitat available for wildlife species and can fragment these habitats into smaller patches separated by agricultural activities and human populations, thereby further promoting unhealthy interactions among pathogens, vectors, and hosts. This has been linked to expansion of "bush meat" consumption, which may have played a key role in emergence of both Ebola and human immunodeficiency virus (HIV) types 1 and 2 in Africa[46] (Figure 121-2).

This land system change likewise coincides with an upsurge in several infectious diseases, including Lyme disease, leishmaniasis, and malaria. With Lyme disease, aggressive land-use changes in the forests of North America have decreased the numbers of small-mammal predators because of habitat fragmentation. The probability that a deer tick will become infected with the *Borrelia* bacteria depends on the density of white-footed mice. As the density of white-footed mice increases because of fewer small-mammal predators, the infection rate of humans with Lyme disease increases accordingly.[8]

Deforestation also causes increases in noninfectious health risks to humans. Mercury is naturally found in rain forest soils and is used to extract gold from riverbeds. Soil erosion after deforestation and gold mining adds significant mercury loads to rivers. The mercury collects in fish, making them hazardous to eat. Mercury suppresses the human immune system and is toxic to humans in very small amounts.

FIGURE 121-2 Deforestation may have played a role in the devastating Ebola virus outbreak. While cutting away swaths of African forest, loggers may have also been inadvertently creating convenient and unexpected pathways for the virus to spread. *(Courtesy* Inhabitat.com.*)*

FIGURE 121-3 Harmful algal bloom, Kelley's Island, Ohio, Lake Erie, September 2009. *(Courtesy National Oceanic and Atmospheric Administration, Great Lakes Environmental Research Laboratory.)*

Lastly, land system change is implicated in mental health disruptions, such as a loss of personal identity and stress. Land-use changes causes a sense of loss called *solastalgia,* which is the experience of negatively perceived change to a home environment. As traditional landscapes where people live are irrevocably changed, people experience stress and negative emotions. Solastalgia is characterized by a sense of desolation and loss that is similar to the nostalgia experienced by people who are forced to migrate from their home environment. However, solastalgia relates to slow changes in the local environment.[14]

Nutrient Cycles

Nutrient cycles are natural processes in which chemical elements, such as nitrogen, phosphorus, and potassium, are continuously cycled among air, water, soil, and organisms. The ratios between elements in the environment are sensitive, so that perturbations impact biodiversity and can have downstream effects on human health. Fertilizer and fossil fuels have the greatest anthropogenic impact on phosphorus and nitrogen cycles, because these are the mainstay ingredients of chemical fertilizers used to replenish nutrients in agricultural lands. Fertilizer proliferation in nearly all parts of the world has resulted in cropland seep-off into watersheds that supply freshwater systems and ultimately the oceans. Freshwater nutrient inputs to the coastal oceans now exceed preindustrial fluxes by 10- to 15-fold in many parts of North America, Europe, and Asia.[3] The cumulative effect of concentrated pollution from different sources is known as "non–point source pollution" and is used to describe effects of watershed areas collecting agricultural pollutants in wetlands or other bodies of water.

Overconsumption of nitrogen compounds can directly affect human health. Nitrates in drinking water play a causative role in methemoglobinemia, or "blue baby syndrome," and are linked to reproductive problems and several cancers. Such nutrient loading into fresh and marine waters can also cause *eutrophication,* a complex process whereby excessive development of certain types of algae disturb aquatic ecosystems, which then become a threat to animal and human health (Figure 121-3). These *harmful algal blooms* (HABs) are the common link between algal blooms, red tides, green tides, fish kills, inedible shellfish, and blue algae. At their height, HABs can release dangerous hepatotoxins, neurotoxins, and dermatotoxins because of cyanobacteria overproliferation. People can be exposed to these

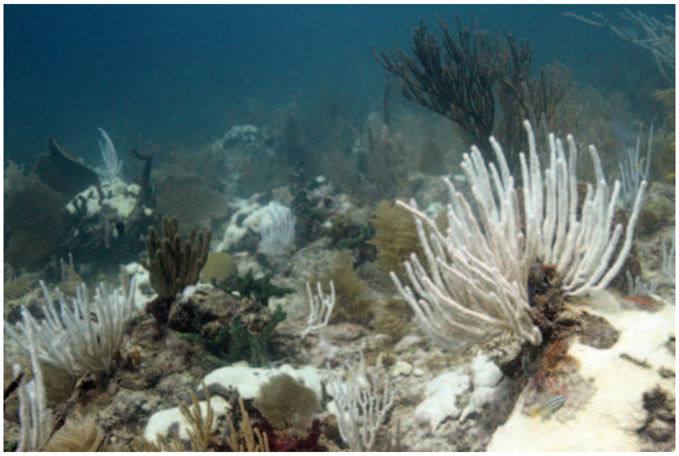


FIGURE 121-4 Most coral reefs exhibit very low annual accretion, with net carbonate production almost balanced against carbonate export, bioerosion, and dissolution. By reducing the growth rate, ocean acidification shifts the balance in favor of net carbonate loss. *(Courtesy National Oceanic and Atmospheric Administration, Coral Health and Monitoring Program.)*

toxins through consumption of contaminated drinking water or seafood, direct contact with contaminated water, and inhalation of aerosols.[61]

When algae die, they decompose, and the nutrients contained in that organic matter are converted into inorganic forms by microorganisms. This decomposition process consumes oxygen, which reduces the concentration of dissolved oxygen for healthy organisms in the ecosystem, and can instigate an anoxia-caused reduction in biodiversity.

Ocean Acidification

The IPCC Fifth Assessment Report (AR5) describes proliferation of planetary greenhouse gases over the past 150 years. By far the majority of this carbon dioxide (CO_2) is retained in the oceans. CO_2 retention in ocean waters influences carbonate chemistry, resulting in acidification (Figure 121-4). Marine organisms are particularly sensitive because aragonite, a form of calcium carbonate created by many marine organisms, dissolves in an acidic environment. This is thought to play a major role in the demise of coral reef biodiversity seen throughout the world in recent decades.[57] This affects many commercial sectors, including tourism and seafood aquaculture.[51] Ocean acidification has also been linked to less resilient coastal ecosystems in the face of extreme weather, nutrient pollution, and overfishing.[51]

FRESHWATER USE

Reduction in water availability for human uses can lead to decreases in health. About 36% of the global population lives in water-scarce regions, resulting in decreased population hygiene and concomitant increase in disease, because households, medical clinics, restaurants, and public places of convenience and hygiene are forced to use minimal water for cleaning. In the United States, the Environmental Protection Agency (EPA) has cited that areas in 36 states face water shortages. When water is scarce, people are forced to rely on drinking water sources that might not be safe. For example, in California in 2012-2013, the water quality in drought-stricken areas was found to have violated federal standards more than 1000 times. An estimated 38 million persons were exposed to contaminants such as arsenic, nitrates, radioactive minerals, and perchlorates (chemicals used in rocket fuel and explosives).[10]

About 70% of the available global freshwater supply is used to irrigate crops, much of which is used to feed livestock. As living standards improve in the developing world, meat consumption has increased, further putting pressure on available freshwater resources.[34]

New challenges to water quality have recently arisen. In the United States, emergence of chlorine-resistant pathogens, chemical contamination of water sources, aging water purification and transportation infrastructure, increased recreational water contamination, nontraditional water exposures (e.g., cooling towers at nuclear power plants), and increasing water reuse threaten public health.[13] Most municipal sanitation systems are designed to handle tons of human waste and toxic chemicals daily, yet worldwide have rarely kept pace with tremendous urban expansion. Many urban sewage networks operate at the margins of capacity and are subject to stress by even the most innocuous rainstorms. This is compounded by the fact that most urban areas do not absorb rainwater consistently, and absorb even less as arable land is put to other uses. A summer rain shower in a city can provide enough stress to allow dangerous runoff of untreated excrement and toxic material, often contaminating local water sources and public areas.

ATMOSPHERIC AEROSOL LOADING

Atmospheric aerosols, also known as particulate matter, are solid or liquid particles suspended in air, with diameters of approximately 0.002 to 100 μm. Primary atmospheric aerosols are particulates emitted directly into the atmosphere from volcanoes, sea spray, forest fires, dust storms, and vegetation in the form of pollen. Secondary atmospheric aerosols are particulates formed in the atmosphere by gas-to-particle conversion, including fossil fuel combustion, power plants, and industrial release of sulfates, nitrates, and some organic products. Atmospheric particulate matter impacts the climate and adversely affects human health, associated with approximately 7.2 million deaths per year from exacerbations of asthma, bronchitis, chronic obstructive pulmonary disease, and acute coronary syndromes.

Aerosol particles vary greatly in size, source, chemical composition, amount and distribution in space and time, and durability in the atmosphere. The smaller and lighter a particle, the longer it will remain in the air. Coarse particles are larger and heavier and are generated by industrial processes that involve crushing or grinding (e.g., construction, farming, mining). Fine particles originate from combustion sources and are formed by gaseous precursors. The size of the particle determines the risk to health. Particles that are 2 to 3 μm in size tend to deposit deeply in the terminal bronchioles and alveoli, whereas larger particles tend to deposit in upper (larger) bronchi, causing different symptoms and disease patterns. The upper bronchi are the part of the respiratory tract that has less interface with the blood supply, so symptoms mostly reflect lung inflammation and are manifested by cough, shortness of breath, and airway spasm or asthma-like symptoms. Smaller particles are especially toxic because they can reach the terminal bronchioles and even as deep as the alveoli, where they can enter the blood supply and travel to the rest of the body. The International Agency for Research on Cancer and WHO designate airborne particulates a group 1 carcinogen. Because of their unfiltered, deep penetration into the lungs and bloodstream, these particulates cause permanent DNA mutations, heart attacks, and premature deaths. In 2013, the European Study of Cohorts for Air Pollution Effects (ESCAPE) trial, a prospective cohort study involving 312,944 people in nine European countries, showed that no safe level exists for particulates, and that for every increase of 10 $\mu g/m^3$ in particulate matter, the lung cancer rate increases by 22%.[49] Symptoms caused by acute exposure to atmospheric aerosols include shortness of breath, cough, and worsened asthma.

Dangerous atmospheric aerosols include pollen released from vegetation. Allergic diseases are common. In the United States alone, approximately 50 million Americans suffer from allergic conditions in response to airborne particulate allergens, including itchy eyes, rashes, rhinorrhea, cough, shortness of breath, bronchitis, and asthma.[38] Climate change and associated increases in land temperatures and CO_2 concentrations are drivers of plant metabolism and pollen production, as well as increased fungal growth and spore release. Increases in air temperature cause earlier flowering of prairie tall grass.[55] Increasing concentrations of grass pollen lead to more frequent ambulance calls because of asthma exacerbations.[20]

CHEMICAL POLLUTION

Chemical pollution refers to new substances, new forms of existing substances, and modified life-forms that have potential for unwanted effects. Anthropogenic introduction of novel entities causes concern when these entities exhibit persistence, have mobility across widespread distributions, and impact vital Earth system processes or subsystems.[57] Of special concern are new types of engineered materials, as well as naturally occurring elements, such as heavy metals that are mobilized by human activities. Release of chlorofluorocarbons into the atmosphere is an example of a synthetic chemical previously thought to be harmless that had unexpected impacts on the ozone layer. More than 100,000 substances exist in global commerce, and if one includes nanomaterials and plastic polymers that degrade to microplastics, the list is longer. Plastic microbeads, used as an exfoliant in certain bath products and cleansers, have been identified recently as a significant pollutant of the Great Lakes. According to a study released by the 5 Gyres Institute, parts of the shore around the lakes have 466,000 particles per square kilometer, with an average of 43,000 particles per square kilometer. Most of these plastic particles are meant to wash down the drain and are less than 1 mm in size.

Water stress worldwide forces people to turn to contaminated sources, increasing their exposure to toxic substances. Water contaminated with lead leached from pipes causes lead toxicity. Lead toxicity results in nervous system damage and developmental delays in children, as well as kidney damage and anemia. Waterborne arsenic exposure usually occurs through natural sources, although environmental contaminants occur from industrial processes associated with mining, metal refining, and timber treatment. Symptoms of arsenic exposure develop over 5 to 20 years and include skin discoloration and thickening (hyperkeratosis); cancers of the skin, bladder, kidneys, and lungs; and diseases of the blood vessels of the legs and feet. The EPA monitors the water supply for a number of synthetic and volatile organic contaminants (e.g., benzene, toluene) and inorganic contaminants (e.g., arsenic, copper) that pose potential threats to human health.

Extensive use of pesticides worldwide poses a great threat to human health. California is one of the few states collecting pesticide data. Between 1991 and 2000, almost 2 billion pounds of pesticides were used in that state alone. An estimated 1.2 billion pounds are used across the United States each year. Persons most at risk for injury are those who have regular, close exposure to pesticides. This includes agricultural workers and, to a lesser extent, populations who live close to farmland. Exposure to pesticides occurs through direct handling, agricultural runoff, and residue on food products. Organophosphates are the most frequently used pesticide and can affect the neurologic system. Toxicity manifests as loss of muscle control, salivation, defecation, bradycardia, and increased bronchial secretions.

Herbicides pose threats to human health. 2,4-D is one of the most widely used herbicides worldwide and is most often used on home lawns, rangeland, and pasture. Human cells that are exposed to 2,4-D undergo genetic damage. This herbicide has been linked to decreased sperm counts in exposed men.[15]

As the amount of chemicals in our environment increases to quantities that are impossible to monitor or study, it may be prudent to consider precautionary and preventive actions to mitigate the unknown risks of pollution. Strategies include further shifting the burden of proof of safety from the consumer to the producer, as well as developing "green chemistry" strategies focused on risk reduction.

REFERENCES

Complete references used in this text are available online at expertconsult.inkling.com.

CHAPTER 122

Sustainability: Leave No Trace

NANCY V. RODWAY

"Sustainability has to be nonideological in order to be, well, sustainable.[4]" According to Tee L. Guidotti, sustainability is a culture, not an ideology, and will only succeed in that light. "The essence of sustainability is optimization of maximization of economic, social, and environmental benefits and operations performance across generations.[4]" The word implies preservation and respect for the future of Earth's resources and the health of its inhabitants. These concepts must become societal values rather than manipulable political agendas. Success in sustainability might be best achieved by technology rather than changes in human behavior, similar to engineering occupational exposure risk out of the workplace, rather than expecting worker compliance with personal protective equipment. Economic, social, and environmental sustainability must be driven by public opinion and supported by innovative technology. Public policy perhaps then will follow.

SUSTAINABILITY IN THE WILDERNESS

Decades ago, the Boy Scouts of America and Leave No Trace organization understood sustainability in the wilderness. The wilderness is not paved, and it is often delicate and easily damaged, sometimes permanently, by the human footprint. Fortunately, sustainability has long been applied to the wilderness. The Leave No Trace Center for Outdoor Ethics, located in Boulder, Colorado, is an educational, nonprofit organization dedicated to responsible enjoyment and active stewardship of the outdoors by all people worldwide. The principles of Leave No Trace (LNT) reflect a sense of stewardship and passion for the world and guide our passage, especially in untamed places.

SEVEN PRINCIPLES OF STEWARDSHIP

Seven guidelines are the official principles of Leave No Trace, Inc., and are copyrighted by the center. This copyrighted information has been reprinted with permission from the Leave No Trace Center for Outdoor Ethics (www.LNT.org):

1. Plan Ahead and Prepare
2. Travel and Camp on Durable Surfaces
3. Dispose of Waste Properly
4. Leave What You Find
5. Minimize Campfire Impacts

REFERENCES

1. Alonso D, Bouma MJ, Pascual M. Epidemic malaria and warmer temperatures in recent decades in an East African highland. Proc R Soc Lond B Biol Sci 2011;278(1712):1661–9.
2. Alves MJ, Fernandes PL, Amaro F, et al. Clinical presentation and laboratory findings for the first autochthonous cases of dengue fever in Madeira Island, Portugal, October 2012. Eur Surveill 2013;18(6):20398.
3. Backer L. Effects of climate change on noninfectious waterborne threats. In: Luber G, Lemery J, editors. Global climate change and human health: From science to practice. San Francisco: Jossey Bass; 2015. p. 171–94.
4. Battisti DS, Naylor RL. Historical warnings of future food insecurity with unprecedented seasonal heat. Science 2009;323(5911):240–4.
5. Beard C, Garofalo J, Gage K. Climate and its impacts on vector-borne and zoonotic diseases. In: Luber G, Lemery J, editors. Global climate change and human health: From science to practice. San Francisco: Jossey Bass; 2015. p. 221–65.
6. Bennett CM, Dear KG, McMichael A. Shifts in the seasonal distribution of deaths in Australia, 1968-2007. Int J Biometeorol 2014;58(5):835–42.
7. Brouwer R, Akter S, Brander L, et al. Socioeconomic vulnerability and adaptation to environmental risk: A case study of climate change and flooding in Bangladesh. Risk Anal 2007;27(2):313–26.
8. Brownstein JS, Skelly DK, Holford TR, Fish D. Forest fragmentation predicts local scale heterogeneity of Lyme disease risk. Oecologia 2005;146(3):469–75.
8a. Burgess BB. Fate of the wild. Athens, GA: University of Georgia Press; 2003.
9. Burkle F. Lecture #2: Global health challenges. Foundations for Global Health Responders, Online course available at Coursera.org. <https://class.coursera.org/ghresponder-001>.
10. California Environmental Protection Agency. <http://www.calepa.ca.gov/>.
11. Cardinale BJ, Duffy JE, Gonzalez A, et al. Biodiversity loss and its impact on humanity. Nature 2012;486(7401):59–67.
12. Carroll B, Morbey H, Balogh R, et al. Flooded homes, broken bonds, the meaning of home: Psychological processes and their impact on psychological health in a disaster. Health Place 2009;15(2):540–7.
13. Centers for Disease Control and Prevention. Coping with water scarcity. <http://www.cdc.gov/ncidod/water_scarcity.htm>.
14. Clayton S, Manning CM, Hodge C. Beyond storms and droughts: The psychological impacts of climate change. Washington, DC: American Psychological Association; 2014. p. 23–5.
15. Cox C. Herbicide factsheet: 2,4-D. J Pesticide Reform 2005;25(4):10–15.
16. Faisal F, Sheffield P. Mitigation: International institutions and global governance. In: Luber G, Lemery J, editors. Global climate change and human health: From science to practice. San Francisco: Jossey Bass; 2015. p. 579–600.
17. Gamble JL, Hurley BJ, Schultz PA, et al. Climate change and older Americans: State of the science. Environ Health Perspect 2013;121(1):15–22.
18. Golden JS. The built environment induced urban heat island effect in rapidly urbanizing arid regions—a sustainable urban engineering complexity. Environ Sci 2004;1(4):321–49.
19. Guha-Sapir D, Vos F, Below R. Annual disaster statistical review 2010: The numbers and trends. Brussels, Belgium: Center for Research on the Epidemiology of Disasters (CRED), University catholic de Louvain; 2010. p. 42.
20. Heguy L, Garneau M, Goldberg MS, et al. Associations between grass and weed pollen and emergency department visits for asthma among children in Montreal. Environ Res 2008;106(2):203–11.
21. Honda Y, Kondo M, Mc Gregor G, et al. Heat-related mortality risk model for climate change impact projection. Environ Health Prev Med 2013;19:56–63.
22. Intergovernmental Panel on Climate Change. Fourth Assessment Report, Working Group 2. <https://www.ipcc.ch/publications_and_data/ar4/wg2/en/ch8s8-4-1-2.html>.
23. Joint Statement of Sixteen National Academies of Science: The science of climate change. The Royal Society. May 18, 2001.
24. Kjellstrom T, Holmer CI, Lemke B. Workplace heat stress, health and productivity—an increasing challenge for low and middle-income countries during climate change. Global Health Action 2009;2(1):6.
25. Kjellstrom T, Lemke B, Otto M. Mapping occupational heat exposure and effects in South-East Asia:ongoing time trends 1980-2009 and future estimates to 2050. Ind Health 2013;51:56–67.
26. Kritz B, Maly M, Benes C, et al. Epidemiology of tick-borne encephalitis in the Czech Republic 1970-2008. Vector-Borne Zoonot Dis 2012;12(11):994–9.
27. Lemery J, Williams C, Farmer P. The great procrastination. Health Hum Rights 2014;16(1):1–3.
28. Liu X, Jiang B, Gu W, et al. Temporal trend and climate factors of hemorrhagic fever with renal syndrome epidemic in Shenyang City, China. BMC Infect Dis 2011;11:331.
29. Lobell DB, Banziger M, Magorokosho C, et al. Nonlinear heat effects on African maize as evidenced by historical yield trials. Nat Clim Chang 2012;1(1):42–5.
30. Loiseau C, Harrigan RJ, Cornel AJ, et al. First evidence and predictions of *Plasmodium* transmission in Alaskan bird populations. PLoS ONE 2012;7(9):e44729.
31. Lozano-Fuentes S, Hayden MH, Ochoa-Martinez C, et al. The dengue virus mosquito vector *Aedes aegypti* at high elevation in Mexico. Am J Trop Med Hyg 2012;87(5):4623–48.
32. Lu L, Lin H. Time series analysis of dengue fever and weather in Guangzhou, China. BMC Public Health 2009;9:395.
33. MacMillan K, Monaghan AJ, Apangu T, et al. Climate predictors of the spatial distribution of human plague cases in the west Nile Region of Uganda. Am J Trop Med Hyg 2012;86(2):514–23.
34. Mekonnen MM, Hoekstra AY. A global assessment of the water footprint of farm animal products. Ecosystems 2012;15(3):401–15.
35. Michon P, Cole-Tobian JL, Dabod E, et al. The risk of malarial infections and disease in Papua New Guinean children. Am J Trop Med Hyg 2007;76(6):997–1008.
36. Moser SC. Navigating the political and emotional terrain of adaptation: Community engagement when climate change comes home. In: Moser SC, Boykoff MT, editors. Successful adaptation to climate change: Linking science and policy in a rapidly changing world. New York: Routledge; 2013. p. 289–305.
37. Neria P, Shultz JM. Mental health effects of Hurricane Sandy: Characteristics, potential aftermath, and response. JAMA 2012;308(24):2571–2.
38. NIH Medline Plus: Achoooooo! Nothing to sneeze at. National Institutes of Health. <https://www.nlm.nih.gov/medlineplus/magazine/issues/summer06/articles/summer06pg20-24.html>.
39. Nitschke M, Tucker G, Hansen A, et al. Impact of two recent extreme heat episodes on morbidity and mortality in Adelaide, South Australia: A case-series analysis. Environ Health 2011;10(1):42–51.
40. NOAA website. National Oceanic and Atmospheric Administration. <http://www.ncdc.noaa.gov/sotc/summary-info/global/201412>.
41. Nunn PD. Responding to the challenges of climate change in the Pacific Islands: Management and technological imperatives. Climate Res 2009;40:211–31.
42. Ogden N, St-Onge HL, Barker IK, et al. Risk maps for range expansion of the Lyme disease vector, *Ixodes scapularis,* in Canada now and with climate change. Int J Health Geogr 2008;7:24.
43. Paaijmans KP, Read AF, Thomas MB. Understanding the link between malaria risk and climate. Proc Natl Acad Sci USA 2009;106(33):13844–9.
44. Pachauri RK, Allen MR, Barros VR, et al. Climate change 2014: Synthesis report. Intergovernmental Panel on Climate Change (Fifth Assessment AR5), Geneva, 2014.
45. Paranjothy S, Gallacher J, Amlot R, et al. Psychosocial impact of the summer 2007 floods in England. BMC Public Health 2011;11(1):145.
46. Patz JA, Daszak P, Tabor GM, et al. Unhealthy landscapes: Policy recommendations on land use change and infectious disease emergence. Environ Health Perspect 2004;112(10):1092–8.
47. Peduzzi P, Chatenouz B, Dao H. Global trends in tropical cyclone risk. Nat Clim Chang 2012;2(4):289–94.
48. Perkins SE, Alexander LV, Nairn J. Increasing frequency, intensity and duration of observed global heat waves and warm spells. Geophys Res Lett 2012;39.
49. Raaschou-Nelsen O, Anderson JZ, Beelen R, et al. Air pollution and lung cancer incidence in 17 European cohorts: Prospective analyses from the European Study of Cohorts for Air Pollution Effects (ESCAPE). Lancet Oncol 2013;14(9):813–22.
50. Robinson M. Foreword. Health Hum Rights 2014;16(1):1–3.
51. Roundtable on Environmental Health Sciences, Research, and Medicine; Board on Population Health and Public Health Practice; Institute of Medicine. Understanding the connections between coastal waters and ocean ecosystem services and human health: Workshop summary. Washington, DC: National Academies Press; 2014. Stressors impacting coastal and ocean ecosystem services and human health. <http://www.ncbi.nlm.nih.gov/books/NBK209264/>.
52. Sainz-Elipe S, Latorre JM, Escosa R, et al. Malaria resurgence risk in southern Europe: Climate assessment in an historically endemic area of rice fields at the Mediterranean shore of Spain. Malar J 2010;9:221.
53. Semenza J. Changes in hydrology and its impacts on waterborne disease. In: Luber G, Lemery J, editors. Global climate change and human health: From science to practice. San Francisco: Jossey Bass; 2015. p. 103–38.
54. Shepherd A, Ivins ER, Barletta VR, et al. A reconciled estimate of ice sheet mass balance. Science 2012;338:1183–9.

55. Sherry RA, Zhou X, Gu S, et al. Divergence of reproductive phenology under climate warming. Proc Natl Acad Sci USA 2007;104(1):198–202.
56. Smith KR, Woodward A, Campbell-Lendrum D, et al. Human health: impacts, adaptation, and co-benefits. In: Climate change 2014: Impacts, adaptation, and vulnerability. Part A. Global and sectoral aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, UK, New York: Cambridge University Press; 2014. p. 736.
57. Steffen W, Richardson K, Rockstrom J, et al. Planetary boundaries: Guiding human development on a changing planet. Science 2015; 347(6223):1–17.
58. Teixeira EI, Fischer G, Van Velthuizen H. Global hot-spots of heat stress on agricultural crops due to climate change. Agric Forest Meteorol 2011;170:206–15.

58a. Tirado C. Addressing the challenges of climate change to food security, safety, and nutrition. In: Luber G, Lemery J, editors. Climate change and human health: from science to practice. San Francisco: Jossey-Bass; 2015. p. 267–97.

59. Vida S, Durocher M, Ouarda T, et al. Relationship between ambient temperature and humidity and visits to mental health emergency departments in Quebec. Psychiatr Serv 2012;63(11):1150–3.
60. World Health Organization. Gender, climate change and health. Public Health and Environment Department (PHE), Health Security and Environment Cluster (HSE). Geneva: World Health Organization (WHO); 2011. p. 249.
61. World Health Organization. Eutrophication and health. Copenhagen, Denmark: WHO Regional Office for Europe; 2002. p. 30.
62. World Health Organization. World malaria report 2014. Global Malaria Programme. Geneva: WHO; 2014. p. 249.
63. World Health Organization. 2014 Dengue and severe dengue. Geneva: WHO.; 2014 <http://www.who.int/mediacentre/factsheets/fs117/en/>.
64. Wu F, Liu Q, Lu L, et al. Distribution of *Aedes albopictus* (Diptera: Culicidae) in northwestern China. Vector-Borne Zoonot Dis 2011;11(8): 1181–6.
65. Zanobetti A, O’Neill MS, Gronlund CJ, et al. Summer temperature variability and long-term survival among elderly people with chronic disease. Proc Natl Acad Sci USA 2012;10(17):6608–13.

6. Respect Wildlife
7. Be Considerate of Other Visitors

Plan Ahead and Prepare

Before departing for an expedition, trip, or even a hike, research the environs and become familiar with the regulations for use. Acquire permits if needed. Plan party size accordingly, and limit size or split the group if necessary to minimize impact. Hike and camp separately if necessary. Avoid high-use times on popular trails, or do not travel at all if poor conditions, such as when a trail is muddy, would cause significant adverse impact. Use proper gear, such as a camp stove, and plan meals ahead to minimize waste. Repackage food before departure in reusable containers or plastic bags that can be easily packed out. Register at the trailhead or with the ranger, and be responsible and aware of personal and party limitations to minimize the chance of rescue. Use a map and compass to eliminate the need for rock cairns or markings on the trail that can mar the landscape for other travelers.

Travel and Camp on Durable Surfaces

Wherever traveling and camping, move on surfaces that are resistant to impact. These include rocky outcroppings, sand, gravel, dry grasses, snow, and water. Stay on well-traveled trails and hike in the center of the trail in single file. Do not shortcut. When boating, launch the craft from a durable area and camp at least 60 m (200 feet, 70 adult steps) from the waterfront. Good campsites, whether in the mountains, beach, desert, or plains, are found, not made. When campsites are not apparent and the area is pristine, try to disperse the impact rather than camp in a tight group.

Dispose of Waste Properly

For human waste, use outhouses where available. If necessary to use a cat-hole, dig it 6 to 8 inches (15 to 20 cm) deep, and choose a site far from water sources. Pack out toilet paper. Don't burn it. Disguise the hole as much as possible. Pack out feminine hygiene products, because they decompose slowly. Treat pet waste as human waste. Urine is generally environmentally tolerated, but nevertheless, urinate far from camps and trails. Aim to urinate on rocks or bare ground to discourage animals from eating tasty and salty urine-soaked foliage. On the water, a portable toilet has become standard practice and may be required by law. With regard to waste from cooking, attempt to plan meals to minimize leftovers that are tempting for wildlife and therefore potentially dangerous. Clean pots with hot water and a scant amount of soap, and scatter the dishwater, after it has been strained to remove food particles, at least 60 m (200 feet) from any water source.

Leave What You Find

Artifacts are often protected by law; leave them where they are found. Do not collect rocks or other portables. Take caution to avoid transporting plant species from one location to another on pack animals, on boots, or in tire treads.

Minimize Campfire Impacts

Avoid campfires unless they are essential for comfort or food preparation. If a campfire is unavoidable, use a fire ring and gather sticks and branches that have fallen on the ground. Do not cut down plants for fuel. Campfires on the beach are a little different; dig a shallow depression in the sand or gravel along the shoreline and, once the fire is cool, scatter the ashes and refill the depression.

Respect Wildlife

Observe wildlife from a safe distance, and do not approach them if you are not an expert. Do not feed wildlife. Avoid wildlife outright during mating season, nesting season, and when they are rearing young. When traveling with pets, keep them under control.

Be Considerate of Other Visitors

Be considerate of other visitors, natives, and native lands. Yield to other users on the trail and, when encountering pack animals, step to the downhill side of the trail. Avoid loud talk, music, and other cacophony.

SUSTAINABILITY IN SPECIAL ENVIRONS

The Mountains

Climbers, when traveling horizontally, can apply many of the aforementioned principles. Approach the route on an established trail, using a trail guide to minimize impact. Once vertical, certain caveats become more task and environ specific. Use only removable protection and as little chalk as possible. Avoid "scrubbing" or "gardening" the route, removing vegetation only when necessary for safety reasons. Do not climb near archeologically sensitive sites or animal habitats such as bird nests.

Toileting can be tricky. Urinating is generally not a problem, but defecating can be gravitationally and environmentally troublesome. In the past, it was allowable to drop waste in a paper bag over the edge of the rock. That is no longer acceptable for obvious reasons. In a pinch, feces can be smeared on the rocks (away from the route), but the preferred method is to pack waste out in a "poop tube," which is a piece of PVC pipe with a screw top that can be attached to the outside of a pack or haul sack, then emptied at the end of a climb or trip. Purchase a 30- to 60-cm (1- to 2-foot, depending on duration of your trip) length of PVC pipe approximately 10 cm (4 inches) in diameter and threaded on one end. Tightly glue a cap on the unthreaded end using PVC cement, and fit a threaded cap on the other end. Fashion a loop out of duct tape or other lashing to attach the screw cap to the piping to prevent it from becoming lost. The loop will also accommodate a carabiner to allow attachment to the outside of a pack. Defecate into small brown paper bags, add cat litter, place the bag into the tube, and then deposit the collection into a vault toilet or dumping station at the end of the trip. Do not flush the paper bags.[9]

Snow

Traveling and camping in the winter are difficult for humans and stressful for wildlife stressed by scarce food supplies. Avoid skiing or camping near game trails or in areas with obvious animal activity, to limit pressure on hungry animals. Campfires are not recommended, given the dearth of accessible firewood and temptation to harvest green wood fuel. When camping, make every effort to "fluff up" the trampled snow for the benefits of subsequent visitors, and do not leave visible "yellow snow" near well-traveled areas. Digging a cat-hole in the snow when winter camping is tempting, but simply leaves the frozen trophy to thaw on the exposed ground in the spring. Pack out all waste.[9]

Water

Waterways are often overused and are shared by recreational and nonrecreational users on motorized and nonmotorized crafts. In addition, water sports enthusiasts have varying levels of respect for scenic rivers, oceans, lakes, and delicate riparian environments. Some users litter beaches with trash. To limit impact in coastal environments, travelers should recreate within the intertidal zone because it is the most durable. Whenever possible, conduct all beach activities in this zone. Camp in an established campsite above the high-tide line of the intertidal zone, and tread on durable surfaces such as trails or rock. If fires are permitted and driftwood is available, build a campfire, if necessary, below high tide. Urinate below the high-tide line away from fellow campers and tidal pools. Use a cat-hole above the high-tide line or pack feces out if the environment is especially fragile. Launch and land watercraft on sand or gravel, avoiding dirt and vegetation. When on the water, do not approach marine wildlife; give all marine animals at least a 90-m (100-yard) berth. Do not dump waste into waterways.[9]

On rivers and freshwater, much of the aforementioned applies. Camp, where possible, in the river's floodplain. Bury human waste in a cat-hole at least 60 m (200 feet) from shore, or pack it out. For large groups such as rafting parties, consider a toilet tank or other latrine.[9]

Tundra

Tundra, the treeless vast soil of polar regions, is visited and traveled most often in the summer, the season when it is most vulnerable. During summer, the surface of tundra thaws to the depth of the permafrost, making it mushy. The thaw is the time when plants and burrowing animal life are most active in this layer. Trampling on summer tundra can be very destructive. When the thin layer of ground cover plants is destroyed, crystals in the underlying permafrost can melt because of increased sun exposure. This is called *thermokarsting* and can leave permanent scars, such as footprints and tire treads. Hiking and camping on durable surfaces are critical. If trails are not available, travel on shallow streambeds or snow, or, as a last resort, walk on tundra grasses rather than on lichen beds. Do not hike single file. For waste disposal, do not dig a cat-hole. Rather, smear feces on a rock or pack them out. Campfires are inappropriate.[9]

Alpine tundra, as opposed to the arctic type, occurs above tree line and has a short growing season (similar to arctic tundra) but has no permafrost. It is delicate but cannot undergo thermokarsting because of the absence of permafrost. This alpine landscape is home to a handful of hearty species of vegetation and adapted mammals. When traveling through alpine tundra, stay on the trail, because damage caused by shortcutting in this fragile environment can remain for several hundred years. Camping is generally discouraged because of the risk of lightning strike at this altitude, as well as soil sensitivity. If camping becomes necessary, try to use an established site or camp on a durable surface, such as rock, snow, or mineral soil. Avoid campfires. Tether trash and camping items to avoid their being blown away by high winds. With regard to human waste, many popular alpine destinations now require that visitors pack it out. However, a patient hiker can descend below tree line to dig a cat-hole if desperate.[9]

The Desert

Many desert soils are covered with a dark cryptobiotic crust, or *biocrust,* composed of cyanobacteria, algae, mosses, and lichen held together by organic materials. These crusts often cover a majority of the desert floor and help to stabilize the soil, fix nitrogen, and retain moisture. If disturbed, biocrusts may not regenerate for a century or more. In the desert, therefore, travelers must remain on designated trails to minimize damage to the biocrust or walk on durable surfaces, such as slickrock, gravel, or sand washes. Camp on durable surfaces or established campsites. Avoid campfires in this dry, treeless environment. Water is limited, so pack plenty. Wandering off the trail in search of water is detrimental for the soil and for thirsty desert creatures. Do not use precious water sources for bathing, because soaps and body oils contaminate the environment. In the desert, cat-holes are the preferred method for waste disposal, but keep them 60 m (200 feet) from any water source, because feces decompose slowly in arid climates.

SUSTAINABILITY AND THE HUMAN FOOTPRINT

According to the World Wildlife Fund, in the 21st century, humankind's footprint exceeds the earth's regenerative capacity by 30%.[5] Wildlife populations have declined by one-third over the last 35 years, and humans are consuming resources at a rate that far exceeds natural regenerative capacity. Biodiversity has declined as species have been overexploited. In 2005, the single largest human footprint was carbon in the form of carbon dioxide (CO_2) from combustion of fossil fuels; ambient CO_2 has grown more than 10-fold since 1961. At the present rate, humankind will need two planets to maintain the present level of consumption by the year 2030.[5] Two generations ago, humankind was an ecologic creditor. At present, two-thirds of our species live in countries that consume more natural resources than exist within their national borders. Therefore, these countries must depend on nations, often Third World countries, with fewer environmental restrictions for resources.

The earth provides food, water, material goods, and fuel. These resources have present economic value. Less salient and less marketable services include nutrient recycling, soil formation, pollination, pest control, and water purification, as well as the aesthetic, spiritual, and recreational provisions of the earth. Some natural resources with market value (e.g., fossil fuels) have been regulated, whereas others (e.g., the atmosphere) have been undervalued or regarded as a common good and therefore have had little to no market oversight, leading to exploitation and overuse. *Sustainability* is the science of managing humankind's footprint on Earth to achieve a balance between what humans use and what the earth can replenish.

ENERGY

The world is powered by fossil fuels that release CO_2 when burned. CO_2 is naturally present in the atmosphere as a trace component; it is released as a product of respiration by plants and animals and in small amounts from volcanoes and geysers. It is one of the greenhouse gases, which serve to keep the earth warm by absorbing and emitting infrared radiation. Without greenhouse gases, the earth would be much colder. The three main greenhouse gases—CO_2, methane (CO_4), and nitrous oxide (N_2O)—affect the atmosphere as functions of their chemistry and rate of decay. Methane, released by fossil fuel combustion, waste dumps, rice paddies, and livestock, accounts for 14.3% of greenhouse gas emissions but is greater than 20 times more effective at trapping heat than is CO_2. Nitrous oxide from fertilizers and industrial processes accounts for only 7.9% of emissions but is about 300 times more effective at trapping heat, ton per ton, than is CO_2.[16] In 1750, before industrialization, atmospheric CO_2 concentration was approximately 280 parts per million (ppm). By 2005, it had reached 379 ppm; the rate of increase from 1995 to 2005 was the highest in recorded history. More CO_2 in the atmosphere generally means more greenhouse gases and a warmer planet. CO_2 is the greatest contributor to global warming, accounting for 43% to 56% of all greenhouse gases, depending on the reference. The single largest contributor to atmospheric CO_2 is consumption of fossil fuels (e.g., coal, oil, natural gas) for energy in transportation, industry, and forestry.[16]

Oil is the most carbon-dense fossil fuel. Natural gas (methane) is the "cleanest" of the carbon-based fuels, because it has more hydrogen atoms per carbon molecule. For each unit of energy, methane produces 52.6 kg (117 lb) of emitted CO_2, compared with 73.8 kg (164 lb) released by burning oil. To achieve a similar amount of heat, one would need almost three times more weight of wood, which would release 88 kg (195 lb) of CO_2. Coal burns the dirtiest, producing 103 kg (227 lb) of CO_2 to achieve the same unit of energy.[3] Although some of these fuels are used primarily and directly for energy, such as in natural-gas stoves, fossil fuels are also burned in power plants to create electricity. Various sources of fuels are combusted or otherwise converted into heat to generate steam, which spins turbines that produce electricity. Electricity then enters the "grid," a large network of power lines, power stations, and transmission subsystems, to be distributed to homes and businesses (Box 122-1).

Worldwide, a majority of the electricity produced comes from coal. Coal combustion is the largest contributor to CO_2 worldwide, but its production and use damage more than just the atmosphere. Coal is extracted by underground, open-surface pit, and "mountaintop" mining. In mountaintop mining, mountaintop debris is deposited in the adjacent valley, destroying vegetation, soil, habitats, and the landscape aesthetics. Acid, along with heavy metals such as mercury, selenium, and arsenic, from this debris seeps into waterways and groundwater. Coal combustion is the largest source of human-made mercury pollution. Acid rain is produced by burning high-sulfur coal.[3]

Power stations can use any type of fuel (uranium, solar energy, biomass, oil, or methane) to power turbines to generate electricity, or the turbines can be turned directly by water and wind power. Although oil and natural gas burn cleaner than does coal, renewable sources of fuel can turn turbines and create electricity, so once the infrastructure is in place, the fuel is essentially free. According to the International Energy Agency, an intergovernmental energy policy advisor founded in the oil crisis of 1974 and located in France, renewable energy sources currently have the technologic potential to supply almost 20 times the current global energy demand, but presently account for no more than 17% of global energy consumption.[7] Biomass and hydropower provide less than 15% of global energy need, whereas wind and solar power fuel provide approximately 2%.

BOX 122-1 The Grid

Turbines that generate electricity from coal, natural gas, biomass, nuclear, wind, and solar power plants must distribute that power from the generating facility to end users. Many power plants are not located close to population centers, so they must transmit electricity via overhead (and occasionally underground) high-voltage transmission lines to substations closer to cities for voltage step-down and further distribution. Energy storage at the substations is inefficient, so electricity is best distributed in real time. A sophisticated system of controls is therefore necessary to ensure that electric generation matches demand. If supply and demand are mismatched, electricity generation and distribution can become overloaded, causing a blackout. To prevent this, generating and distributing stations are all interconnected in a "grid" or "power grid" to allow for redundancy in the system, creating a series of transmission triangles rather than a branching hub. These individual triangles are then connected regionally and nationally. The grid allows for generally uninterrupted electricity delivery through periods of high and low demand and high and low power generation, as can occur with intermittent renewable fuel sources. Finally, the miles of transmission lines of the grid act to pool and store electricity from various sources, both renewable and nonrenewable.

The network nature of the grid is ripe for computer management to improve efficiency. A "smart grid" delivers electricity from suppliers to consumers controlled by two-way digital technology. Such a system could alert the consumer, via smart meter devices such as a glowing orb, to high-use periods, giving direct feedback to limit electricity use during expensive peak demand periods. Alternatively, the system could automatically turn off selected high-demand appliances during peak periods, which can be cost and carbon saving, turning them back on as demand lessens. A smart grid can charge an electric vehicle at night, a time of low electricity demand. Renewable energy sources will need a smarter grid. As the world converts to renewable energy, which is mostly intermittent power, a smart system will be necessary to limit demand during peak periods. Some of this can be accomplished by pricing energy as a function of demand and allowing market forces to work.

A "home grid" extends some of these capabilities into the home to allow the individual homeowner to cut electricity cost and the individual's global footprint. Alternatively, individual homes can generate their own electricity "off the grid" or sell surplus to the grid. Many municipalities, regions, and countries already have established a smart or "smarter" grid.

RENEWABLE ENERGY

Renewable energies are essential contributors to the global energy menu and have the potential to reduce reliance on fossil fuel and to mitigate greenhouse gases, thereby lessening humans' ecologic footprint. Regions of the world differ in the types of renewable energy used, but overall, hydropower is the global renewable energy source of choice for electricity, and biomass for energy production.

BIOMASS

Biomass is an advanced form of photosynthetic solar power and a promising renewable replacement for fossil fuels in power plants. The energy generation is a relatively simple chemical process. Biomass is mashed and fermented with yeast to make ethanol. Ethanol is then burned for fuel. Initially, food crops such as corn were regarded as ideal biomass fuel options. However, the fossil fuel demands of industrial agriculture negated the benefits of decreased emissions. Corn was replaced by nonfood sources of biomass when the economic and social impacts of using food for fuel became obvious. Many sources of renewable (and sustainable) biocellulose work as sources of biomass. These include grasses (especially *Miscanthus* [elephant grass]), waste paper, corn silage, and bagasse waste from sugar cane production. One acre of sugar cane produces 2500 L (650 gal) of fuel; 1 acre of corn produces 1500 L (400 gal); and 1 acre of *Miscanthus* grass produces 4750 L (1250 gal). Soybeans can be used to make 174 L (46 gal) of biodiesel. Oil palms can be converted into 2300 L (610 gal) of biodiesel per acre.[3]

WIND POWER

Wind power, like biomass, is a form of solar energy, because atmospheric temperature differentials from the sun create wind. It is so abundant that it could fuel the entire planet five times over, and it is the fastest-growing renewable energy source. Wind power has the advantage of being rapidly installed, aesthetic, and scalable to the needs of the locale. An average windmill can generate enough electricity to power 400 average American homes; a smaller windmill 11 to 43 m (35 to 140 feet) tall can pay for itself after 6 years. The United States, Germany, and Spain generate the most electricity from wind power; however, India and China may soon surpass these countries because of their current investments in wind farms. England boasts the largest offshore wind farm in the North Sea, where winds are strong and reliable.[3]

A typical windmill stands 50 to 100 m (160 to 325 feet) tall and has blades that range from 27 to 45 m (90 to 150 feet) in length, making construction and transportation a challenge, especially for offshore wind farms. Wind farms are regarded as hazardous for birds. In reality, cats kill 3000 birds for every one struck by a windmill, and tall buildings flatten 19,000 birds for every windmill kill.[3] Despite this relative safety, engineers are perfecting sensors to halt windmill operation when birds are nearby. The largest drawback to wind power is its intermittency; power isn't generated if the wind doesn't blow.

SOLAR POWER

The sun radiates enough energy, were it to be captured properly to the earth in 1 hour, to power the entire planet for a year. Solar energy is the most common form used by persons trying to "live off the grid." Capturing solar power can be challenging because it is limited by clouds and darkness of night and is intermittent. Technically speaking, there are two mechanisms to harness solar energy. Solar rays on a large scale can be focused by curved mirrors to heat liquids that turn turbines in power plants. Alternatively, and usually on a smaller scale, photovoltaic cells convert sunlight directly into energy using semiconductor devices. Photovoltaic cells consist of a thin layer of silicon atoms that release free electrons when exposed to solar energy. They work in the presence of intermittent sunlight and can be deployed in small clusters or in large arrays. Free electrons flow out of the photovoltaic cell as electric current, which is converted to alternating current by an inverter for use in residential homes and other applications.

Solar power can also be harnessed in a passive manner by intelligent design of residential and commercial buildings. Orientation of the building can take advantage of winter sun and minimize summer glare. The roof can be colored to reflect or absorb intense heat and can serve as a solar water heater. Building materials, such as stone, can be chosen to absorb heat, whereas proper ventilation can circulate both cool and warm air to limit the use of fossil fuels.

GEOTHERMAL ENERGY

Geothermal energy has enough potential stored energy to satisfy the world's needs many times over, according to the United Nations World Energy Assessment Report.[15] Geothermal energy creates virtually no CO_2 emissions and is not intermittent, an advantage over other renewable sources. Geothermal activity is greatest where the tectonic plates meet, such as the Ring of Fire surrounding the Pacific Ocean. In addition, there are other naturally occurring hot spots where magma has found its way to the surface and created springs and geysers. Both these natural formations can serve as direct sources of hydrothermal energy, where natural steam is used to turn turbines. There is, however, the potential for geothermal energy in unsuspected locations. In many areas, rock below the ground surface is hot but dry. If rock temperature exceeds 149° C (300° F) and this bedrock is close enough to the surface to be cost-efficient, water can be injected into the ground in an Enhanced Geothermal System, and the resultant steam used to generate electricity. For personal consumption, a homeowner can install a geothermal heat pump to reduce the cost of heating and cooling a building. A hole is drilled 12 to 60 m (40 to 200 feet) below the surface, where the earth's temperature is a stable 16° C (60° F). A loop of copper pipes is installed, through which refrigerant pumped from the house circulates in the loop, exchanging heat from the home with the earth. In the summer, warm air in the home is absorbed by the refrigerant, then circulates underground to bring cool air back to the surface. The process is reversed in winter.

NUCLEAR ENERGY

Nuclear power, which generates heat through a controlled fission chain reaction using uranium, is an option to reduce carbon emissions. Uranium is the heaviest naturally occurring compound, containing 92 protons. When split, energy is released and the free neutrons collide with nearby uranium atoms, splitting them as well. "Control rods" of cadmium or other elements absorb some of the neutrons, limiting and controlling the reaction. The generated heat boils water into steam that turns an electric turbine. One pound of uranium contains as much energy as 3 million lb of coal.[3]

The United States is the current leader in nuclear power, with 100 active nuclear reactors, followed by France with 58 and Japan with 48. Nuclear engineers are aging and academic programs closing. The cost of building and maintaining a nuclear power plant has skyrocketed, and new reactors worldwide remain unfinished. According to the United National World Energy Assessment, 17% of global electricity production comes from nuclear power, behind coal (38.3%) and gas (18.1%).[15]

Nuclear energy has fallen out of public favor because of concerns about the consequences of long-term storage of radioactive waste, the public's hesitation to have this waste stored in its "backyard," and safety concerns after the accidents at Three Mile Island in 1979 (United States), Chernobyl in 1986 (USSR/Russia), and Fukushima Daiichi in 2011 (Japan). According to the U.S. Nuclear Regulatory Commission (NRC) and the International Atomic Energy Association (IAEA), the accident at Chernobyl released more than 100 times the radiation released by the two atomic bombs dropped by the United States on Japan in 1945 in World War II.[3]

Radiation is a natural phenomenon. Humans emit radiation from potassium-40 in the body and are exposed to naturally occurring radiation from elements and rocks (e.g., granite). Most living things are able to genetically withstand a certain amount of radiation. Radioactive waste, however, contains a number of radioisotopes that emit ionizing radiation that can be harmful to humans and the environment. Nuclear fuel, nuclear weapons, and health care industries produce nuclear waste. Waste generated by hospitals, such as contaminated towels, filters, and rags, is generally low-level waste that can be incinerated and buried in landfills, which poses insignificant long-term risk. Iodine-131, used in diagnosis and treatment of thyroid conditions, has a half-life of 8 days and is essentially gone from the body and environment in approximately 3 months. Plutonium-239, used for nuclear power and weapons, has a half-life of more than 24,000 years. Such high-level waste originates from spent reactor fuel and waste materials from reprocessing spent fuel rods. High-level waste is thermally hot and highly radioactive and remains so for many years. This waste is generally stored above ground or underwater for 3 to 5 years to allow it to cool before definitive disposal. The waste is contained, then relocated and disposed in a permanent, dry geologic site, far from human contact, with the surrounding rock providing a natural radiation barrier. Once filled, the disposal site is closed and sealed. Geologic disposal, regarded as the safest solution to radioactive waste, obviates the need for long-term storage facility maintenance and lessens the risk of terrorist acquisition. Despite the impressive safety record in the nuclear waste storage and disposal industry, public concern and opposition continue. Public opinion supports long-term storage, which is a security risk, rather than permanent disposal.

The International Atomic Energy Agency (IAEA), headquartered in Vienna, Austria, was created in 1957 as an independent, intergovernmental, and science-based organization in the United Nations family. The IAEA publishes safety standards for transport and storage of radioactive waste and also maintains a databank, the Energy and Environment Data Reference Bank, which is a compilation of country-specific energy and environment-related indicators, such as CO_2 emissions per capita and overall energy statistics (http://www.iaea.org/inisnkm/nkm/aws/eedrb/). In 2005, the IAEA was the recipient of the Nobel Peace Prize for its efforts at ensuring that nuclear energy is used for peaceful purposes.

SUSTAINABLE LIVING

Renewable energy choices and decreased dependence on fossil fuels are the core of environmental sustainability; however, simple lifestyle changes can also lessen the individual human footprint. In the 2009 book, *Our Choice*, Al Gore reveals the CO_2 excesses of the Western diet.[3] Industrial agriculture uses 10 calories of energy from fossil fuel to produce 1 calorie of food, which does not include fuel burned to transport grain, meat, vegetables, and fruit. In the livestock industry, 3.18 kg (7 lb) of plant protein and 23,000 L (6000 gal) of water are required to produce 0.45 kg (1 lb) of beef. In addition, large livestock operations, such as those in Canada and California, release significant amounts of methane, a potent greenhouse gas, into the atmosphere. Natural gas, required for production of nitrogen fertilizer for industrial agriculture, releases 4.6 tons of CO_2 for every ton of fertilizer manufactured.[3] Nitrogen excess degrades the carbon content of soil as hungry soil bacteria consume fertilizer and release CO_2. Residual fertilizer enters the waterways, triggering robust algal blooms that starve water of oxygen, kill fish, and leave a dead zone. Cane sugar, another Western staple, is water intensive, requiring 1500 L (400 gal) of freshwater to produce 1 kg (2.2 lb) of sugar. Food choices can be more "sustainable" and carbon neutral if humans consume less meat, more fruits and vegetables, and purchase foodstuffs from local sources to minimize fuel used in transportation. Gardening is a green alternative, especially if natural fertilizers (manure or garden spoilage) are used.

Clothing choices can be sustainable. Cotton production, similar to beef and sugar, has a large water footprint, requiring 2900 L (800 gal) of water to produce one cotton shirt.[5]

SUSTAINABLE HOSPITALS

According to the U.S. Department of Energy, hospitals are among the most energy-intensive facilities, producing 2.5 times the CO_2 emissions of commercial office buildings and releasing more than 13.5 kg (30 lb) of CO_2 emissions per square foot.[17] Besides consuming large amounts of energy, hospitals use toxic elements and chemicals for diagnosis, treatment, and decontamination, potentially threatening health of the environment, patients, visitors, and health care workers. The 21st century has seen an emphasis on sustainable and healthy health care with such initiatives as the Global Health Security Initiative (GHSI; http://www.ghsi.ca/), Health Care Without Harm (www.noharm.org), Practice Greenhealth (www.practicegreenhealth.org), and The Center for Health Design (www.healthdesign.org). The focus of these initiatives is to improve architectural design to improve energy efficiency and to eliminate waste, improve management of medical waste, eliminate mercury in hospitals, and provide sustainable and organic options in hospital cafeterias. In addition, green hospitals are using environment-friendly cleaning products and more eco-friendly building practices, which can improve working conditions for employees and improve air quality for patients and visitors.

Medical waste is an environmental challenge. The vast majority of medical waste is incinerated, releasing CO_2, mercury, and dioxins, among other byproducts. Because paper waste makes up about one-half of hospital waste, many hospitals are replacing paper surgical gowns with reusable ones and medical charts with electronic medical records. Sharps containers can be reused, saving money and oil. There are biodegradable bedpans made of recycled phone books and beeswax.

SUSTAINABLE TRAVEL

Traveling in a carbon-neutral or carbon-friendly fashion seems straightforward. One should walk or bicycle when possible and use public transportation, such as a train or bus, where available, favoring trains for short trips and buses for long distances. Air travel is the most carbon intense, with short flights emitting more carbon per traveler per mile than do longer ones; nonstop flights are easier on travelers and the environment. The eco-conscious traveler can purchase a carbon offset to neutralize environmental impact. Using an online carbon calculator, a traveler pays a fee, based on calculated carbon emissions, to a service organization that typically plants trees, captures methane, or builds a windmill to offset the CO_2 emitted. The growing popularity of this practice has created Internet scams, so the carbon-offset industry is striving for transparency and legitimacy. Trusted carbon-offset sellers include Terrapass, which has a simple carbon footprint calculator and carbon gift certificates; Native Energy; and the Climate Trust, which is focused on eco-friendly businesses. The Gold Standard (www.cdmgoldstandard.org), a nonprofit organization under Swiss law that operates a certification program for carbon credits, is supported by multiple nongovernmental organizations (NGOs), World Wildlife Fund International, Greenpeace International, and others. It is regarded as the international certifying body for premium-quality carbon credits.

Critics of the concept of purchasing carbon offsets believe that the presence of an "easy carbon out" only encourages carbon use when we should be preventing its release. In addition, how can the purchaser ensure that the tree was planted and survived, and/or that the project funded by the carbon offsets was not previously scheduled? The Gold Standard and critics discourage reliance on tree planting for many reasons, including difficulty with verification.

HYDRAULIC FRACTURING

Hydraulic fracturing, or fracking, is an issue that cannot be ignored in any discussion of sustainability. Fracking is a method for extraction of oil or natural gas from tight formations of shale or other rock located thousands of feet underground (Figure 122-1). It has made the United States one of the world's leading producers of natural gas and has contributed to declining natural-gas prices.[14] Natural gas is regarded as a cleaner "bridge fuel" between the carbon-heavy fossil fuels of coal and oil and the carbon-neutral renewable energy sources such as wind power. When burned, natural gas emits less CO_2 per unit of energy than when oil or coal is burned. However, if the gas is leaked directly into the atmosphere, methane (natural gas) is 86 times more potent than CO_2 over a period of 20 years. The global-warming impact of methane falls to 34 over a period of 100 years because methane has a life span of only 12.4 years in the atmosphere, much shorter than that of CO_2.[18]

Natural gas, or methane, can be found in pockets and loose rock formations such as sandstone, where it is easily accessed by conventional drilling methods. Unfortunately, conventional oil and gas reservoirs in the United States have long ago been pumped dry. Newer technologies with horizontal drilling methods can extract natural gas from tight rock formations, such as shale, where methane is less accessible. Rock formations amenable to fracking exist throughout the world. However, one of the largest shale gas formations is the Marcellus shale region that underlies Ohio, Pennsylvania, New York, Maryland, Virginia, and West Virginia in the United States. Fracking, the newer unconventional horizontal drilling method, accesses reservoirs previously inaccessible or too expensive to drill. In this process, a vertical well is drilled into the rock formation many hundreds to thousands of feet underground, generally far below the water table. Once the vertical drill reaches the shale or other rock, it is then directed horizontally for thousands of feet, along the horizontal length of the targeted rock formation. The vertical portion of the well is cased with a steel lining to prevent groundwater contamination. The horizontal portion of the wellbore is lined with perforated piping. The well is then "fracked." A mixture of large amounts of water, chemicals, and sand is injected under high pressure into the well, creating fractures in the shale that are propped open by the injected sand. Fracked wells typically produce oil and gas for 3 years.[14] Along with marketable fuels, fracked wells produce wastewater that contains some of the injected chemicals, dissolved clays, salts, heavy metals, and radioactive substances. Once a fracked well becomes economically unproductive, it is "capped" to prevent leakage of methane into the atmosphere.[14]

WATER POLLUTION

Unconventional drilling requires millions of gallons of water to frack a well and creates millions of gallons of wastewater for disposal, a serious concern when fracking occurs in water-starved areas.[14] Even if water is abundant, such as in the Marcellus shale region, wastewater produced from fracking can contaminate groundwater and surface water. The chemical toxicity of wastewater is often unknown because drilling companies protect the identity of chemicals used in fracking as a "trade secret" despite that dangerous and carcinogenic chemicals such as toluene, benzene, and xylene have been identified.[2] In Ohio, the state passed a bill that, in the event of an emergent medical exposure, physicians can acquire proprietary information regarding chemicals used in fracking; however, the physician must keep the identity of the chemical confidential.[13] The radioactivity of fracking wastewater limits disposal options; the safest method is injection into deep injection wells. Transportation of water to the disposal site is risky because truck traffic can be hazardous.[11] Wastewater from fracking can also be reused on site but more often is placed in lined temporary holding pits created locally until permanent disposal is available. Occasionally, wastewater is transported to specialized water treatment facilities because regional septic wastewater treatment plants cannot accommodate radioactive wastewater.[14] In one instance, tributaries of the Ohio River were contaminated with barium, strontium, and bromides when wastewater was treated at a municipal treatment plant.[6] In West Virginia, wastewater is legally sprayed on local roads or grounds. All these disposal methods are vulnerable to leakage and have the potential for subversion. Cases of groundwater contamination have been documented in Ohio and Pennsylvania.[14] Surface water contamination, through accident or subversion, has the potential to poison fish, wildlife, and local livestock; these contaminated animals can potentially enter the food chain.[14] Methane from fracked wells can contaminate local drinking water, although the gas itself poses a significantly greater risk to the atmosphere. Methane-contaminated water is not hazardous to drink but poses a risk of combustion. A study of private wells near fracking sites showed that 75% of wells within 1 km (0.6 mile) of hydraulic fracturing sites in the Marcellus shale in Pennsylvania were contaminated with methane isotopically identical to the fracked gas.[6]

The true extent of contamination of surface and groundwater by fracking wastewater is unknown. The Energy Policy Act of 2005 created a loophole that exempted the oil and gas extraction industry from the National Environmental Policy Act, and thus fracking companies are not legally obligated to consider environmental impacts of oil and gas extraction. This loophole also exempts the fracking industry from most provisions of the Safe Drinking Water Act of 1974. As a result, much of the fracking industry goes unregulated. The work of regulating hydraulic fracturing has fallen to the states and localities, with disparate results. Water quality data before drilling are often unavailable, because fracking companies are not required to perform these tests, making the causation of water contamination difficult. When proof of water contamination is attributable to fracking, legal proceedings and settlements often include confidentiality

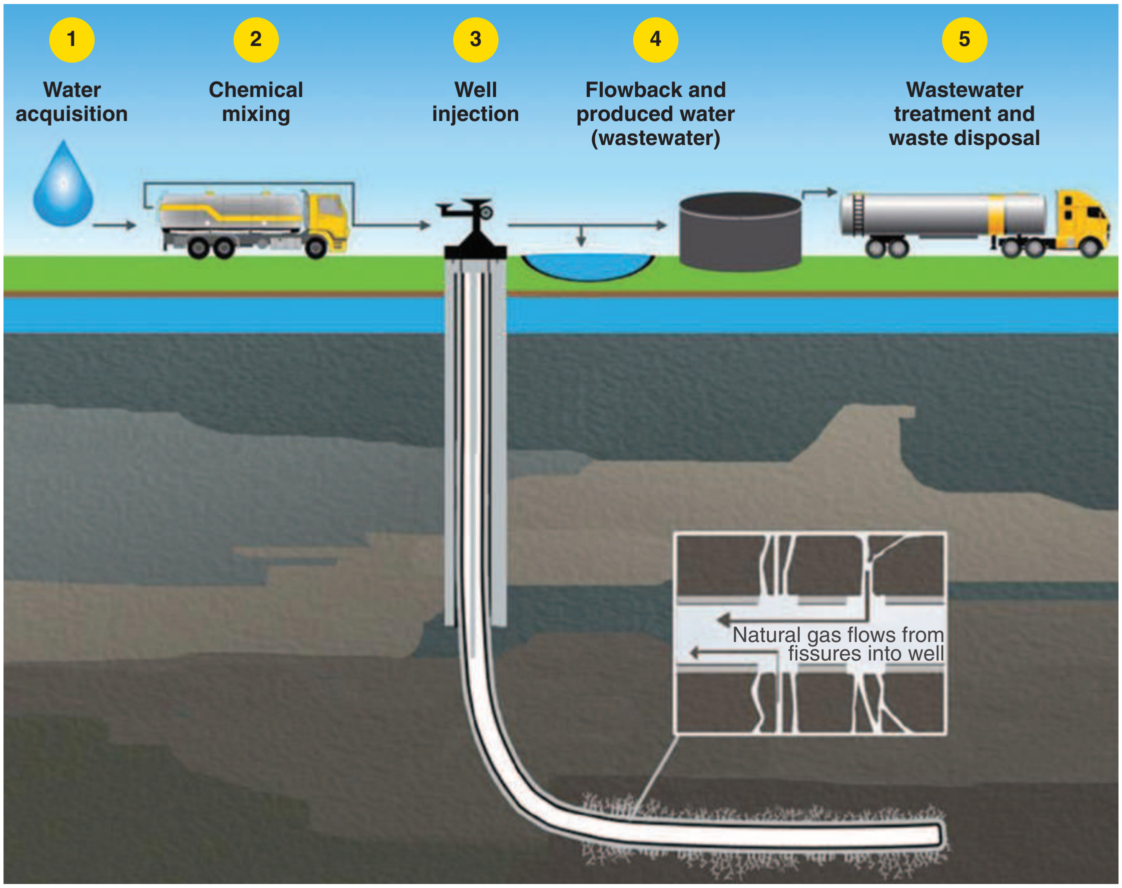


FIGURE 122-1 **Stage 1: Water Acquisition**[1]
Large volumes of water are withdrawn from groundwater[2] and surface water[3] resources to be used in the hydraulic fracturing process.
Potential Impacts on Drinking Water Resources
Change in the quantity of water available for drinking
Change in drinking water quality
Stage 2: Chemical Mixing
Once delivered to the well site, the acquired water is combined with chemical additives[4] and proppant[5] to make the hydraulic fracturing fluid.
Potential Impacts on Drinking Water Resources
Release to surface and groundwater through on-site spills and leaks
Stage 3: Well Injection
Pressurized hydraulic fracturing fluid is injected into the well, creating cracks in the geologic formation that allow oil or gas to escape through the well to be collected at the surface.
Potential Impacts on Drinking Water Resources
Release of hydraulic fracturing fluids to groundwater caused by inadequate well construction or operation
Movement of hydraulic fracturing fluids from the target formation to drinking water aquifers through local human-made or natural features (e.g., abandoned wells, existing faults)
Movement into drinking water aquifers of natural substances found underground, such as metals or radioactive materials, which are mobilized during hydraulic fracturing activities
Stage 4: Flowback[6] and Produced Water[7] (Hydraulic Fracturing Wastewaters)
When pressure in the well is released, hydraulic fracturing fluid, formation water, and natural gas begin to flow back up the well. This combination of fluids, containing hydraulic fracturing chemical additives and naturally occurring substances, must be stored on-site, typically in tanks or pits, before treatment, recycling, or disposal.
Potential Impacts on Drinking Water Resources
Release to surface water or groundwater through spills or leakage from on-site storage
Stage 5: Wastewater Treatment and Waste Disposal
Wastewater is dealt with in one of several ways, including, but not limited to, disposal by underground injection, treatment followed by disposal to surface water bodies, or recycling (with or without treatment) for use in future hydraulic fracturing operations.
Potential Impacts on Drinking Water Resources
Contaminants reaching drinking water caused by surface water discharge and inadequate treatment of wastewater
Byproducts formed at drinking water treatment facilities by reaction of hydraulic fracturing contaminants with disinfectants

[1] Recently, some companies have begun recycling wastewater from previous hydraulic fracturing activities, rather than acquiring water from ground or surface resources.

[2] Groundwater is the supply of freshwater found beneath the earth's surface, usually in aquifers, which supply wells and springs. It provides a major source of drinking water.

[3] Surface water resources include any water naturally open to the atmosphere, such as rivers, lakes, reservoirs, ponds, streams, impoundments, seas, and estuaries. It provides a major source of drinking water.

[4] Chemical additives are used for a variety of purposes (see examples in Table 4, p. 29, of the hydraulic fracturing Study Plan). A list of publicly known chemical additives found in hydraulic fracturing fluids is provided in Appendix E, Table E1, of the hydraulic fracturing Study Plan.

[5] Proppant is a granular substance such as sand that is used to keep the underground cracks open once the hydraulic fracturing fluid is withdrawn.

[6] Flowback.

[7] Produced water.

(Courtesy U.S. Environmental Protection Agency (EPA). http://www2.epa.gov/hfstudy/hydraulic-fracturing-water-cycle.)

agreements, limiting public discourse and availability of objective data.

AIR QUALITY

Air quality is typically measured in terms of six "criteria pollutants"—particulate matter, ozone, carbon monoxide, lead, nitrogen oxides, and sulfur dioxide—chosen for their impact on human cardiorespiratory health and carcinogenicity. In areas of fracking, particularly during the drilling and completion phase, particulate matter and ozone levels are increased because of truck traffic and emissions from drilling.[14] In addition, volatile hydrocarbons from fracking fluids, such as benzene, toluene, and xylene, are released into the air, causing at least noxious odors.[2] Toxic air pollutants near fracking sites in Texas have included formaldehyde, chloroform, and carbon tetrachloride.[1] Objective data on air pollution from fracking are limited because drilling sites are restricted with only remote data ("fence studies") available. A 2013 report from the Office of the Inspector General of the United States concluded: "Limited data from direct measurements, poor quality emission factors, and incomplete NEI [National Emissions Inventory] data hamper EPA's ability to assess air quality impacts from oil and gas production activities. With limited data, human health risks are uncertain, states may design incorrect or ineffective emission control strategies, and the EPA's decisions about regulating this industry may be misinformed."[14]

SEISMIC ACTIVITY

Hydraulic fracturing, the particular process of initially injecting the chemicals/sand/water into a well to fracture the rock formation, causes small earthquakes, generally undetectable at the surface.[14] Disposal of wastewater into deep injection wells is associated with more significant seismic activity. The largest earthquake caused by deep water injection, rated at 5.7 magnitude, occurred in November 2011 in Oklahoma.[14] Improvements in well site-selection can mitigate this risk, particularly when fracking is considered in earthquake-prone areas, such as near fault lines.

SAND MINING

Hydraulic fracturing accounts for 41% of all the silica sand used in the United States, triggering a silica sand surface strip-mining boom, especially in the upper Midwest.[14] The sudden expansion of silica strip mining creates an environmental concern as silica dust is harmful to workers at mines, workers at well sites, and residents living near fracked wells. In 2012, the National Institute for Occupational Safety and Health (NIOSH) issued a hazard alert for workers at fracking sites after discovering elevated levels of silica in the air near many wells.[12] Silicosis is a pulmonary disease caused by inhaling silica dust that can have a latency period of more than 10 years.

GLOBAL WARMING

Combustion of methane releases less CO_2 into the atmosphere per energy generated than combustion of other fossil fuels, such as oil or coal. Although less carbon friendly than renewable energy sources, methane is regarded as the most environmentally safe fossil fuel. However, if methane gas is leaked directly into the atmosphere before burning, it is very damaging. One molecule of methane creates an atmospheric global-warming footprint that is 86 times greater than one molecule of CO_2.[18] Using present practices, a fracked well typically releases 3.6% to 7.9% of "fugitive methane" into the atmosphere. A Cornell study of the carbon footprint of shale gas used for consumer heat generation estimated that, with the 3.6% to 7.9% rate of fugitive methane, a fracked well creates a larger CO_2 footprint than does an oil well, with a 1.7% to 6% methane leak, over a 20-year period.[6] When methane is burned for electricity, excluding the consumer component, the CO_2 footprint produced by shale gas is somewhat better because of the efficiency of gas power plants over coal power plants, but methane is still more damaging to the atmosphere than coal or oil. The longer the estimated time, the smaller is the difference, because methane remains in the atmosphere for a shorter period than CO_2 but is a much more potent greenhouse gas. According to the Cornell study, shale gas creates a larger estimated carbon footprint in timescales of less than about 50 years.[6] The National Oceanic and Atmospheric Administration (NOAA) found that methane leakage rates are even higher than previously estimated. This concern alone can hasten the global-warming tipping point (the point in time when interventions can no longer stabilize the global-warming process) to within the next 15 to 35 years.[2] The carbon footprint of hydraulic fracturing can be easily mitigated by trapping or flaring fugitive methane; present U.S. federal regulatory requirements do not mandate this.[14]

NATIONAL PARKS IN THE UNITED STATES

Drilling for methane and oil has already occurred in 12 national parks and threatens the borders of many others, especially parks that are atop shale formations.[10] The surface footprint of a fracking drilling well site, approximately 2.5 acres, is much larger than conventional wells to allow room for chemical storage tanks, holding pits, and roads for heavy truck traffic.[10,14] The sheer size of the footprint has great potential to fragment habitats and impact native wildlife species within and around a national park.[10] This can potentially affect biodiversity and stability of plant and animal species, and alter migration patterns and grazing habitats. Flaring of methane gas from fracked wells near and within national parks, although good for the atmosphere, causes significant light obscuration in lands that were previous nighttime viewing destinations, such as the Theodore Roosevelt National Park in North Dakota.[10] Noise pollution from fracking is disruptive to visitors. Water extraction for drilling can impact the amount and quality of water in rivers coursing through national parks, in addition to concerns about fracking wastewater. Visitors heading east from Glacier National Park encounter a sign warning against poisonous gases from fracking operations.[10]

CONCLUSION

The shale gas boom has created a market where the United States will rapidly become a net exporter of natural gas.[2] However, the typical boom-and-bust cycles of energy extraction have historically created abandoned oil and gas wells and abandoned coal mines. These leave taxpayers with cleanup bills and are an environmental risk until the sites are remediated.[12] Although most states in the United States have small bonding requirements for well plugging (some as low as $100 per well), few states have established policies that require drillers to have bonds, trust funds, or insurance policies to cover well reclamation once the well has lost its profitability.[12] These funds would cover restoration of the environment, compensation of victims for damage to property and health, provision of alternate sources of water if needed, and full restoration of damaged public infrastructure (e.g., road damage from heavy truck traffic). With such low bonding rates, the oil and gas industry has little incentive to reclaim sites. Between 2001 and 2008, 127 mines in West Virginia and 227 mines in Pennsylvania were abandoned and the posted bonds forfeited; this may increase as falling fuel prices financially pressure drilling companies.[12]

The National Resources Defense Council opposes expanding fracking without additional safeguards to protect human health and the environment.[11] The U.S. Environmental Protection Agency (EPA) is now working on a comprehensive study of the environmental impacts of hydraulic fracturing. Objective scientific data regarding the environmental safety of unconventional drilling are difficult to find. Legal settlements for groundwater contamination of private wells often include confidentiality agreements. Academic studies have been sponsored by the fracking industry without appropriate attribution, coined "frackademia."[14] Policy makers and regulators are subjected to financial and environmental pressures from both corporate lobbyists in favor of fracking and voters and landowners strongly opposed to fracking. Direct measurement of air and water quality data at and around fracking well sites is regularly restricted to "fence studies" where the fracked site is known, but the industry does not routinely reveal the location of fracking sites, wastewater pits, or injection wells. The overall paucity of objective environmental data regarding the impact of hydraulic fracturing on human and environmental health has created a biased informational landscape, impairing public debate as well as policy formation. Geisinger Health Systems, sitting atop the Marcellus shale in Pennsylvania, is presently conducting a comprehensive epidemiologic survey on the health implications of fracking using 2.6 million available electronic health records.[2]

However, the present informational vacuum cannot adequately predict the short-term and long-term impacts of fracking. In a 2013 review of the impact of fracking on public health, Finkel and Hays[2] summarize the data: "no sound epidemiologic study has been done" to quantify fracking's impact on human and environmental health. They conclude, "Natural gas has been in shale formations for millions of years; it isn't going anywhere."

THE CHALLENGE

A mere two generations ago, humankind was an ecologic creditor, using fewer natural resources than were generated by the planet. By 2010, fueled by population growth and individual consumption, we have amassed environmental debt. We must learn to live within our ecologic means. Each country has a measurable ecologic footprint equivalent to the sum of all the land and water required to support its consumption and absorb its waste. According to the 2008 *Living Planet Report,* this global sum exceeded the planet's available supply in the 1980s, and by 2005, demand was 30% greater than was biocapacity.[5] The United States and China have the largest individual footprints, with each country consuming 21% of the planet's biocapacity. Much of China's footprint is caused by its large population. India is a distant third at 7%. These three countries have among the highest natural biocapacities in the world but are now in ecologic debt. Many developed countries have ecologic footprints that far exceed their biocapacity and increasingly depend on the remaining creditor countries for resources.

Scientists cannot accurately predict precisely when the earth will reach an ecologic "point of no return." According to the World Wildlife Fund and the *Living Planet Report* of 2005, we have the technology to return the earth to biostability by controlling population, limiting individual consumerism, and decreasing the amount of resource use and waste production.

Global population in 2015 was more than 7,291,000,000 (http://www.worldometers.info/). A nation's ecologic footprint is a function of its population, natural resources used, and waste produced. Lowest-income countries have had the greatest increases in population. Increases in middle-income population and consumerism have led to the highest per-person footprint, accounting for 39% of the total per-capita footprint. High-income countries account for 36% of the footprint, primarily from increases in per-person carbon footprint rather than increases in population.[5] To achieve sustainability, lower-income nations would benefit from family-planning services, education, and empowerment of women regarding childbearing choices. In middle- and high-income countries, family planning may be helpful; however, renewable energy sources and investment in resource-efficient cities are also important. In addition, material consumerism in richer countries, along with the accompanying electronic waste, has reached unprecedented levels. A culture of nonmaterial personal rewards could help ease this particular ecologic stress.

The largest gap between biocapacity and ecologic footprint is caused by energy consumption. Energy production from burning fossil fuels accounts for 45% of the global ecologic footprint.[5] Energy in all forms is a global issue. Its unfettered production, distribution, and use are unsustainable. The energy status quo is a threat not only to the environment but also to social equity and national security. Energy externalities are not only environmental (acid rain, global warming, radiation exposure) but social (black lung disease, asthma exacerbation, malignancies). External costs vary by energy source, being greatest for coal, followed by oil, and then by renewable energy sources.[15] The "free market," which has neglected these long-term costs in the past, can include them, pricing energy as a function of energy's total cost, rather than its subsidized cost. This may encourage consumers to purchase greener products and use renewable energy sources. A multifaceted approach to sustainability in the home and workplace, on the road, and in the wilderness, along with significant political will, is needed, because no single approach is adequate to sustain the earth and democratize energy for all people.

Renewable energies, including biomass, hydropower, wind, solar and geothermal, account for a few percent of the total global energy market because of technologic reasons and a disabling policy environment that subsidizes fossil fuel–based energy. Leadership to advance innovation in energy may come from Third World countries with less entrenched infrastructures. Once the infrastructure is in place, renewable fuel is essentially free. "Success" in the future will be measured in financial, social, and environmental terms. The addition of "environmental success" to our cultural lexicon will require leadership, sense of urgency, and global effort.

REFERENCES AND SELECTED RESOURCES

Complete references and selected resources used in this text are available online at expertconsult.inkling.com.

SELECTED RESOURCES

Leave No Trace. <http://www.lnt.org>.
Our Choice website by Al Gore. <http://ourchoicethebook.com>.
Intergovernmental Panel on Climate Change. <http://www.ipcc.ch/>.
The United Nations Environmental Programme. <http://www.unep.org/>.
UNEP e-book Climate in Peril. <http://www.grida.no/publications/climate-in-peril/ebook.aspx>.

Carbon Offsets

The Gold Standard. <http://www.cdmgoldstandard.org/>.
Native Energy. <http://www.nativeenergy.com/>.
Terrapass. <http://www.terrapass.com/>.

Nuclear

International Atomic Energy Agency. <http://www.iaea.org/>.
IAEI publishes recommendations for radiation exposure of patients during diagnostic procedures, such as computed tomography, dentistry, interventional cardiology procedures, and mammograms. These are found at <http://rpop.iaea.org/RPoP/RPoP/Content/index.htm#>.

Green Hospitals

The Global Health Security Initiative (GHSI). <http://www.ghsi.ca/>.
Health Care Without Harm. <www.noharm.org>.
Practice Greenhealth. <www.practicegreenhealth.org>.
The Center for Health Design. <www.healthdesign.org>.

REFERENCES

1. Allen DT. Atmospheric emissions and air quality impacts from natural gas production and use. Annu Rev Chem Biomol Eng 2014;5:55–75.
2. Finkel ML, Hays J. The implications of unconventional drilling for natural gas: A global public health concern. Public Health 2013;127:889–93.
3. Gore A. Our choice. Emmaus, Pa: Rodale; 2009.
4. Guidotti TL. Health and sustainability, an introduction. New York: Oxford University Press; 2015.
5. Hails C, editor. Living planet report. Gland, Switzerland: World Wide Fund for Nature, World Wildlife Fund; 2008.
6. Howarth RW, Ingraffea A. Comment: Should fracking stop? Nature 2011;477:271–3.
7. International Energy Agency statistics, 2010. <www.iea.org/stats/index.asp>.
8. Reference deleted in proofs.
9. McGivney A. Leave no trace: A guide to the new wilderness etiquette. Seattle: The Mountaineers; 1998.
10. National Parks Conservation Association Center for Park Research. National Parks and hydraulic fracturing: Balancing energy needs, nature, and America's national heritage. <http://www.npca.org/assets/pdf/Fracking_Report.pdf>.
11. National Resources Defense Council. In fracking's wake: New rules are needed to protect our health and environment from contaminated wastewater. <http://www.nrdc.org/energy/files/fracking-wastewater-fullreport.pdf>.
12. PennEnvironment Research and Policy Center. Who pays the cost of fracking? Weak bonding rules for oil and gas drilling leave the public at risk. <http://pennenvironmentcenter.org/sites/environment/files/reports/Who%20Pays%20the%20Cost%20of%20Fracking.pdf>.
13. Thompson H. Fracking boom spurs environmental audit. Nature 2012;485:556–7.
14. Union of Concerned Scientists. Toward an evidence-based fracking debate: Science, democracy, and community right to know in unconventional oil and gas development. <http://www.ucsusa.org/HF report>.
15. United Nations Development Programme. World energy assessment: Energy and the challenge of sustainability. New York: United Nations; 2004.
16. United Nations Development Programme. Climate in peril: A popular guide to the latest IPCC reports. Arendal, Norway: GRID-Arendal and SMI Books; 2009.
17. US Department of Energy. 2009, April 29. Retrieved October 30, 2010, from DOE announces the launch of the hospital energy alliance to increase the energy efficiency in the heathcare sector, <www.energy.gov/news/7363.htm>.
18. Wikipedia. Global warming. <http://en.wikipedia.org/wiki/Global-warming_potential>; [accessed 01.02.15].

CHAPTER 123

Brief Introduction to Oceanography

TOM GARRISON

OCEAN STATISTICS

The ocean may be defined as the vast body of saline water that occupies the depressions of the earth's surface. More than 97% of the water on or near the earth's surface is contained in the ocean; less than 3% is held in land ice, groundwater, and all the freshwater lakes and rivers.

Traditionally, we have divided the ocean into artificial compartments called *oceans* and *seas* by using the boundaries of continents and imaginary lines such as the equator. In fact, the ocean has few dependable natural divisions and is only one great mass of water. The Pacific and Atlantic Oceans and the Mediterranean and Baltic Seas, so named for our convenience, are in reality only temporary features of a single world ocean. In this chapter, I refer to the ocean as a single entity, with subtly different characteristics at different locations but with very few natural partitions. Such a view emphasizes the interdependence of the ocean with land, life, water, atmospheric and oceanic circulation, and natural and human-made environments.

Earth's ocean exists because of a fortuitous combination of circumstances. Our planet's orbit is roughly circular around a relatively stable star. Earth is large enough to hold an atmosphere, but not so large that its gravity would overwhelm. Its neighborhood is tranquil—supernovae have not seared its surface with ionizing radiation. Our planet generates enough warmth to recycle its interior and generate the raw materials of atmosphere and ocean, but is not so hot that lava fills vast lowlands or roasts complex molecules. Best of all, our distance from the Sun allows the earth's abundant surface water to exist in the liquid state. Ours is a clement ocean world; surely *Oceanus* would be a better name for our watery home.

The ocean moderates temperature and dramatically influences weather. The ocean borders most of the planet's largest cities. It is a primary shipping and transportation route that provides much of our food. From its floor is pumped more than one-third the world's supply of petroleum and natural gas. The dry land on which almost all of human history has unfolded is hardly visible from space, because nearly three-quarters of the planet is covered by water.

BRIEF APPRECIATION OF THE OCEAN'S HISTORY AND MODERN OCEAN TOOLS

The ocean did not prevent the spread of humanity. By the time European explorers set out to "discover" the world, native peoples met them at almost every landfall. Ocean transportation offers people the benefits of mobility and greater access to food supplies. Any coastal culture skilled at raft building or small-boat navigation would have economic and nutritional advantages over less-skilled competitors. Thus, from the earliest period of human history, understanding and appreciating the ocean and its lifeforms benefited those people patient enough to learn.

Systematic application of marine science began at the Library of Alexandria in Egypt. Founded during the third century BC at the behest of Alexander the Great, the library and adjacent museum could be considered the first university in the world. When any ship entered the harbor, the books (actually, scrolls) it contained were by law removed and copied; the copies were returned to the owner and the originals kept for the library. Caravans arriving by land were also searched. Manuscripts describing the Mediterranean coast were of great interest, and traders quickly realized the competitive benefit of this information.

The second librarian at Alexandria (from 235 to 192 BC) was the Greek astronomer, philosopher, and poet Eratosthenes of Cyrene. This remarkable man was the first to calculate the circumference of the earth using geometry. The Greek Pythagoreans had realized that Earth was spherical by the sixth century BC, but Eratosthenes was the first to estimate its true size.

Scientific oceanography began with the departure of *HMS Challenger* from Plymouth, England, in 1872. Conceived by Wyville Thomson, a professor of natural history at Scotland's University of Edinburgh, and his Canadian-born student, John Murray, *HMS Challenger*'s 4-year cruise was the first research expedition devoted completely to marine science, and it also holds the record as the longest such voyage. Other scientific expeditions had been launched previously, such as the voyages of Captain James Cook, RN, and the United States Exploring Expedition under Charles Wilkes, but these were hybrid military and scientific undertakings. The *HMS Challenger* voyage is notable for being the first purely scientific oceanographic endeavor. Stimulated by their own curiosity and with the inspiration of Charles Darwin's voyage in *HMS Beagle*, Thomson and Murray convinced the Royal Society and British government to provide a Royal Navy ship and trained crew for a prolonged and arduous voyage of exploration across the oceans of the world. They coined a word for their enterprise: *oceanography*. Although the term literally implies only marking or charting, it has come to refer to the science of the ocean.

The demands of scientific oceanography have become greater than the capability of any single voyage. Modern oceanography depends on an interlocking suite of terrestrial and space-based sensors. Among the most interesting are the radar altimeters borne by *TOPEX/Poseidon/Jason,* as the project is known, a train of satellites orbiting 1336 km (830.2 miles) above Earth in a pattern that allows coverage of 95% of the ice-free ocean every 10 days. Experiments that are occurring as part of this 5-year program include sensing water vapor over the ocean, determining the precise location of ocean currents, and determining wind speed and direction. The most revolutionary devices are the satellites' *TOPography Experiment,* which use radar positioning devices to allow researchers to determine position to within 1 cm (0.4 inch) of Earth's center. Computers can then determine the height of the sea surface with unprecedented accuracy.

Disregarding waves, tides, and currents, researchers have found the ocean surface can vary from the ideal smooth (ellipsoid) shape by as much as 200 m (656 feet). The reason is that the pull of gravity varies across Earth's surface depending on the nearness or farness of massive parts of the earth. An undersea mountain or ridge "pulls" water toward it from the sides, forming a mound of water over itself. For example, a typical undersea volcano with a height of 2000 m (6562 feet) above the seabed and a radius of 20 km (12.4 miles) would produce a 2-m (6.6-feet) rise in the ocean surface. This mound cannot be seen with the unaided eye because the slope of the surface is very gradual. The large features of the seabed are amazingly and accurately reproduced in the subtle standing irregularities of the sea surface. Hundreds of previously unknown features have been discovered using the data provided by this project.

Small robot submersibles were much in the news during the *Deepwater Horizon* oil spill in the Gulf of Mexico in 2010. These nimble devices can manipulate valves, lift and reposition equipment, and act as remote sets of eyes for decision-makers. Scientists use them to probe submerged geologic features, examine shipwrecks, and measure water.

Perhaps the most imaginative new technology being incorporated into submersibles is *telepresence,* the extension of a person's senses by remote manipulators. A scientist might wear a helmet containing small stereo television screens and earphones, and place his hands in special gloves equipped with tactile feedback units. Movements of his head and hands would be duplicated by a robot on the seafloor, and sensations "felt" by the robot would be relayed back to the scientist through the TVs, earphones, and gloves. He or she would thus have the sensation of being on the seafloor and could take samples, manipulate tools, or just look around. Other researchers could watch or participate at distant locations via a high-speed data link.

Personal investigation is still an important option. As amazing as are robots, satellites, and multibeam systems, sometimes there is no substitute for actually *seeing*—focusing a well-trained set of eyes on the ocean floor.

The most difficult problem is to reach extreme depths, but amazingly, scientists have visited the bottom of the deepest ocean basin. On 23 January 1960, U.S. Navy lieutenant Don Walsh and Dr. Jacques Piccard descended to a depth of 11,022 m (6.85 miles) into the Challenger Deep, an area of the Mariana Trench discovered in 1951 by the British oceanographic research vessel *Challenger II.* The vehicle used in the descent was *Trieste*, a deep-diving submersible designed like a blimp with a very strong and thick (and cramped) steel crew sphere suspended below. A blimp uses helium gas for buoyancy, but a gas would be compressed by water pressure; so gasoline, which is relatively incompressible, was used to provide lift. The trip was repeated by film director James Cameron in 2012 in a vehicle of different design.

We have come a long way since the 1960s. *Alvin*, the best-known and oldest of the deep-diving manned submarines now in operation, has made more than 4500 dives since its commissioning in 1964. Recently refurbished and certified to even greater depths, *Alvin* will continue to explore the Mid-Atlantic Ridge and other seabed features of interest to geologists and biologists. *Alvin*'s abilities have been surpassed by a new class of manned vehicles, the most capable of which are Japan's *Shinkai* 6500 and China's *Jaiolong*. These submarines can reach depths greater than 7000 m (22,966 feet).

WATER CHARACTERISTICS

Water is so familiar and abundant that we do not always appreciate its unusual characteristics. Two major concepts are reviewed here. The first is the influence of water on global temperatures. Liquid water's thermal characteristics prevent broad swings of temperature during day and night and, through a longer span, during winter and summer. Heat is stored in the ocean during the day and released at night. A much greater amount of heat is stored through the summer and given off during the winter. Liquid water has an important thermostatic balancing effect—an oceanless Earth would be much colder in winter and much hotter in summer than the moderate climates we experience. The second concept is the influence of density on ocean structure. Ocean structure and large-scale movement depend on changes in the density of seawater, with this density dependent on temperature and salt content.

Perhaps the most important physical properties of water are related to its behavior as it absorbs or loses heat. Water's unusual thermal characteristics prevent wide temperature variation from day to night and from winter to summer, permit vast amounts of heat to flow from equatorial to polar regions, and power Earth's great storms, wind waves, and ocean currents.

Heat and temperature are related concepts but are not the same. *Heat* is energy produced by random vibration of atoms or molecules. On average, water molecules in hot water vibrate more rapidly than do water molecules in cold water. Heat is a measure of how many molecules are vibrating and how rapidly they are vibrating. *Temperature* records only how rapidly the molecules of a substance are vibrating. Temperature is an object's response to the input (or removal) of heat. The amount of heat required to bring a substance to a certain temperature varies with the nature of that substance.

Heat capacity is a measure of the heat required to raise the temperature of 1 gram of a substance by 1°C. Different substances have different heat capacities, and not all substances respond to identical inputs of heat by rising in temperature the same number of degrees (Table 123-1). Heat capacity is measured in calories per gram per degree centigrade.

Because of the great strength and large number of the hydrogen bonds between water molecules, more heat energy must be added to speed up molecular movement and raise water's temperature than would be necessary in a substance held together by weaker bonds. Liquid water's heat capacity is therefore among the highest of all known substances. This means that water can absorb (or release) large amounts of heat while changing relatively little in temperature.

The uniqueness of water becomes even more apparent when one considers the effect of temperature change on water's density. Most substances become denser as they become colder. Pure water generally becomes denser as heat is removed and its temperature falls, but water's density behaves in an unexpected way as its temperature approaches the freezing point. As the water continues to cool, its framework of hydrogen bonds becomes more rigid; this causes the liquid to expand slightly because the molecules are held slightly farther apart. Water becomes slightly less dense as cooling continues, until 0°C (32°F) is reached; this is the point at which water begins to freeze and change state by crystallizing into ice. At this point, the density of the water decreases abruptly. Ice is therefore lighter than an equal volume of water. Ice increases in density as it becomes colder than 0°C; however, no matter how cold it becomes, ice never reaches the density of liquid water. Because it is less dense than water, ice "freezes over" as a floating layer instead of "freezing under" as do the solid forms of virtually all other liquids.

Progressive transition from liquid water to ice crystals requires continued removal of heat energy; the change in state does not occur instantly throughout the mass when the cooling water reaches 0°C (32°F). Removal of heat does not stop when some of the water in the freezing water mass reaches the freezing point, but the decline in temperature stops. Although heat continues to be removed, the water will not become colder until the water mass has changed state from liquid (water) to solid (ice). Heat may therefore be removed from water when it is changing state (i.e., when it is freezing) without the water dropping in

TABLE 123-1 Heat Capacity of Common Substances*

Substance	Heat Capacity† in Calories/gram/°C
Silver	0.06
Granite/sand	0.20
Aluminum	0.22
Alcohol (ethyl)	0.30
Gasoline	0.50
Acetone	0.51
Ice (not freezing or thawing)	0.51
Pure liquid water	1.00
Ammonia (liquid)	1.13

*Heat capacity is a measure of the heat required to raise the temperature of 1 gram of a substance by 1°C.

†Different substances have different heat capacities. Note how little heat is required to raise the temperature of 1 gram of silver by 1°C. Of all common substances, only liquid ammonia has a higher heat capacity than liquid water.

temperature. Indeed, continued removal of heat is what makes the change in state possible. Heat is released as hydrogen bonds form to make ice, and that heat must be removed to allow more ice to form. This heat is called the *latent* heat of fusion (from the Latin *latere,* meaning "to be hidden").

The implications of this odd thermal behavior are striking. More than 18,000 km^3 (11,185 miles^3) of polar ice that covers as many as 20 million km^2 (12.4 million miles^2) of surface thaws and refreezes in the southern hemisphere each year; this is an area of ocean larger than South America. The annual change in sea ice cover is less in the Arctic, averaging about 5 million km^2 (3.1 million miles^2). Incoming solar heat melts ice in the local polar summer, but the ocean's temperature does not change. The situation reverses during the winter: the water freezes, and again the temperature does not change. Models suggest that without this thermostatic effect—or if ice "froze under" rather than "froze over"—Earth would be a much different planet, perhaps roiled by near-transonic winds peaking about 1 month after the equinoxes at equatorial latitudes.

The total quantity or concentration of dissolved inorganic solids in water is its *salinity*. The ocean's salinity varies from about 3.3% to 3.7% by mass, depending on such factors as evaporation, precipitation, and freshwater runoff from the continents. However, the average salinity is usually given as 3.5%. Most of the dissolved solids in seawater are salts that have been separated into ions. Sodium (Na^+) and chloride (Cl^-) are the most abundant of these ions.

The many ions present in seawater react with each other and with water molecules in complex ways to modify the physical properties of pure water:

- The heat capacity of water decreases with increasing salinity. In other words, less heat is necessary to raise the temperature of seawater than is required to raise the temperature of freshwater by the same amount.
- Dissolved salts disrupt the webwork of hydrogen bonding in water. As salinity increases, the freezing point of water becomes lower; the salts act as a form of antifreeze. Sea ice therefore forms at a lower temperature than does ice in freshwater lakes.
- Because dissolved salts tend to attract water molecules, seawater evaporates more slowly than does freshwater. Swimmers usually notice that freshwater evaporates quickly and completely from their skin, but seawater lingers.
- *Osmotic pressure,* which is the pressure exerted on a biologic membrane when the salinity of the environment is different from that within cells, rises with increasing salinity. This is a key factor related to the transport of water into and out of cells.

These four properties, which vary with the quantity of solutes dissolved in water, are called water's *colligative* properties (Latin *colligatus,* "to bind together"). Because colligative properties are the properties of solutions, the more concentrated (saline) the solute, the more important these properties become. Because it is not a solution, pure water has no colligative properties.

Because about 3.5% of seawater consists of dissolved substances, boiling away 100 kg of seawater theoretically produces a residue with a mass of 3.5 kg. Because variations of 0.1% are significant, however, oceanographers prefer to use the parts-per-thousand notation (‰) rather than percent (%, parts-per-hundred notation) when discussing these materials. The seven ions listed below oxygen and hydrogen in Table 123-2 make up more than 99% of this residual material; sodium and chloride make up 85% of the total. When seawater evaporates, its ionic components combine in many different ways to form table salt, Epsom salts, and other mineral salts.

Seawater also contains minor constituents. The ocean is sort of an "Earth tea"; almost every element present in Earth's crust and atmosphere is also present in the oceans, although sometimes in extremely small amounts. Only 14 elements have concentrations in seawater of more than 1 part per million. Elements present in amounts less than 0.001‰ (1 part per million) are known as *trace elements.*

Remembering the effectiveness of water as a solvent, one might think that the ocean's saltiness has resulted from the ability of rain, groundwater, or crashing surf to dissolve crustal rock. Much of the sea's dissolved material originated in that way, but is crustal rock the source of all the ocean's solutes? An easy way to find out would be to investigate the composition of salts in river water and compare this to that of the ocean as a whole. If crustal rock is the only source, the salts in the ocean should be like those of concentrated river water; however, they are not. River water is usually a dilute solution of bicarbonate and calcium ions, whereas the principal ions in seawater are sodium and chloride. The magnesium content of seawater would be higher if seawater were simply concentrated river water. The proportions of salts in isolated salty inland lakes (e.g., Utah's Great Salt Lake, the Dead Sea) are much different from the proportions of salts in the ocean. Thus, weathering and erosion of crustal rock cannot be the only source of sea salts.

TABLE 123-2 Major Constituents of Seawater at 34.4‰ Salinity

Constituent	Concentration in Parts per Thousand (‰) or Grams per Kilogram (g/kg)	Percent by Mass
Water Itself		
Oxygen	857.8	85.8
Hydrogen	107.2	10.7
Most Abundant Ions		
Chloride (Cl^-)	18.980	1.9
Sodium (Na^+)	10.556	1.1
Sulfate (SO_4^{2-})	2.649	0.3
Magnesium (Mg^{2+})	1.272	0.1
Calcium (Ca^{2+})	0.400	0.04
Potassium (K^+)	0.380	0.04
Bicarbonate (HCO_3^-)	0.140	0.01
Total	999.377 g/kg	99.99%

The components of ocean water with proportions that are not accounted for by the weathering of surface rocks are called *excess volatiles*. The sources of these excess volatiles are Earth's deeper layers. The upper mantle appears to contain more of the substances found in seawater (including the water itself) than are found in surface rocks, and their proportions are about the same as found in the ocean. Convection currents slowly churn Earth's mantle, causing the movement of tectonic plates. Because of this activity, some deeply trapped volatile substances escape to the exterior, outgassing through volcanoes and rift vents. These excess volatiles include carbon dioxide (CO_2), chlorine, sulfur, hydrogen, fluorine, nitrogen, and water vapor. This material, along with residue from surface weathering, accounts for the chemical constituents of today's ocean.

Some of the ocean's solutes are hybrids of the two processes of weathering and outgassing. Table salt (sodium chloride) is an example of this. Sodium ions come from the weathering of crustal rocks, whereas chloride ions come from the mantle by way of volcanic vents and outgassing from midocean rifts. As for the lower-than-expected quantity of magnesium and sulfate ions in the ocean, research at a spreading center east of the Galápagos Islands suggests that the chemical composition of seawater percolating through midocean rifts is altered by contact with fresh crust. The water that circulates through new ocean floor at these sites is stripped of magnesium as well as of a few other elements. The magnesium seems to be incorporated into mineral deposits, but calcium is added as hot water dissolves adjacent rocks.

Recent research has shown that temperature and density gradients inside seamounts also drive great quantities of water into close association with hot geologic bits. The ocean contains about 15,000 seamounts, and the volume of seawater circulated through them may exceed the amounts associated with ridges. Astonishingly, all the water in the ocean is thought to cycle through the seabed at rift zones every 1 to 2 million years.

OCEAN STRUCTURE

Heat combines with salinity to define ocean structure. A liter of seawater weighs between 2% and 3% more than a liter of pure water because of the solids (often called *salts*) dissolved in seawater. The density of seawater is thus between 1.020 and 1.030 g/cm^3 compared with 1.000 g/cm^3 for pure water at the same temperature. Cold, salty water is denser than warm, less salty water. Seawater's density increases with increasing salinity, increasing pressure, and decreasing temperature. Figure 123-1 shows the relationship between temperature, salinity, and density. Note that two samples of water can have the same density at different combinations of temperature and salinity.

Much of the ocean is divided into three density zones: the surface zone, pycnocline, and deep zone. The *surface zone,* or mixed layer, is the upper layer of ocean. Temperature and salinity are relatively constant with depth in the surface zone because of the action of waves and currents. The surface zone consists of water in contact with the atmosphere and exposed to sunlight; it contains the ocean's least dense water and accounts for only about 2% of total ocean volume. This layer typically extends to a depth of about 150 m (500 feet), but depending on local conditions, may reach a depth of 1000 m (3300 feet) or may be absent entirely.

The *pycnocline* (Greek *pyknos,* "strong," and Latin *clinare,* "to slope" or "to lean") is a zone in which density increases with increasing depth. This zone isolates surface water from the denser layer below. The pycnocline contains about 18% of all ocean water.

The *deep zone* lies below the pycnocline at depths of more than about 1000 m (3300 feet) in midlatitudes (40 degrees south to 40 degrees north). There is little additional change in water density with increasing depth through this zone. This deep zone contains about 80% of all ocean water.

The pycnocline's rapid density increase with depth is mainly the result of a decrease in water temperature. The surface zone is well mixed, with little decrease in temperature with depth. In the next layer, temperature drops rapidly with depth. Beneath it lies the deep zone of cold, stable water. The middle layer, the zone in which temperature changes rapidly with depth, is called the *thermocline.* Falling temperature is the major contributor to the formation of the pycnocline.

Thermoclines are not identical in form in all areas or latitudes. Surface temperature is proportional to available sunlight. More solar energy is available in the tropics than in the polar regions, so the water there is warmer. The ocean's sunlit upper layer is thicker in the tropics, both because the solar angle there is more nearly vertical and because water in the open tropical ocean contains fewer suspended particles and is therefore clearer than water in open temperate or polar regions. Because the ocean is heated to a greater depth, the tropical thermocline is deeper and much more pronounced than thermoclines at higher latitudes. The transition to the colder, denser water below is more abrupt in the tropics than at high latitudes.

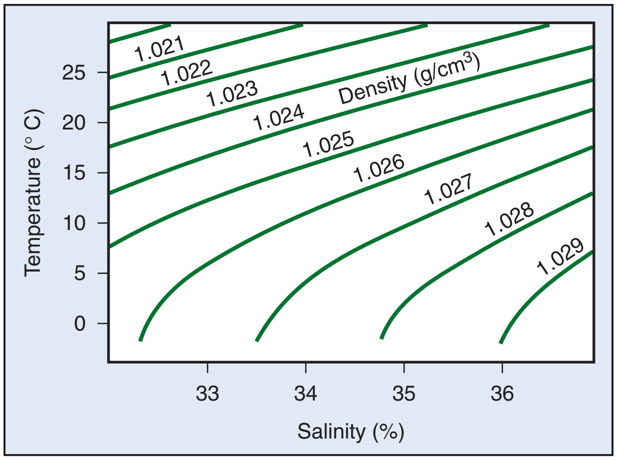


FIGURE 123-1 The complex relationships between temperature, salinity, and density of seawater. Note that two samples of water can have the same density at different combinations of temperature and salinity.

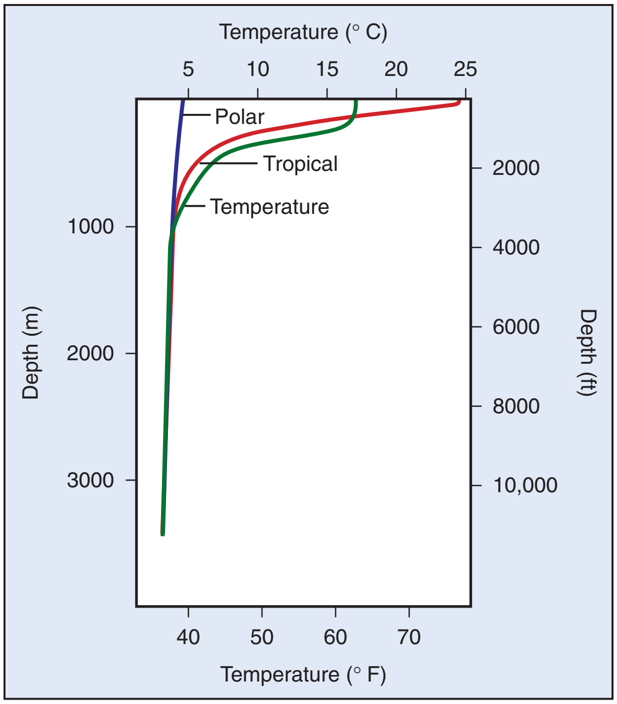


FIGURE 123-2 Typical temperature profiles at polar, tropical, and middle (temperate) latitudes. Note that polar waters lack a thermocline.

Polar waters, which receive relatively little solar warmth, are not stratified by temperature and generally lack a thermocline because surface water in the polar regions is almost as cold as water at great depths.

Figure 123-2 contrasts polar, tropical, and temperate thermal profiles, showing that the thermocline is primarily a mid- and low-latitude phenomenon. Thermocline depth and intensity vary with season, local conditions (e.g., storms), currents, and many other factors.

The vertical movement of large volumes of water from the surface to great depths (and vice versa) is possible only where surface-water density is similar to deep-water density. The great difference in temperature—and therefore density—between surface water and deep water in the tropics makes the water column very stable and prevents an exchange of surface and deep water. This stability is maintained even though the surface of the tropical ocean is in constant horizontal motion, churned by tropical cyclones and stirred by currents.

Vertical movement of water in the northern polar ocean is also limited. There, however, the stratification is caused largely by a salinity difference between surface water and water at great depths. The surface of the Arctic Ocean receives a large volume of freshwater runoff from Siberian and Canadian rivers. Continental masses block the formation of large currents, and the landlocked northern ocean communicates sluggishly with other ocean areas, so the surface water tends not to mix with deeper water or to flow to lower latitudes.

By contrast, the southern polar ocean is only weakly stratified. The cold temperature of southern ocean surface water closely matches that of deep water, so no thermocline divides surface water from deep water (see Figure 123-2). The absence of confining continental margins and mixing at the boundaries of the Antarctic Circumpolar Current minimize salinity differences. Turbulence and weak stratification encourage a huge volume of deep-water upwelling, which contributes to high surface nutrient levels and high biologic productivity.

OCEAN CIRCULATION

Layering by density traps dense water masses at great depths, where they are not exposed to daily heating and cooling, surface circulation driven by winds and storms, or light. The pycnocline effectively isolates 80% of the world ocean's water from the 20% involved in surface circulation. Dense water masses form near polar continental shelves (as cold water freezes and excludes salt) or in enclosed areas such as the Mediterranean Sea (where evaporation exceeds precipitation and river input, raising salinity). These heavy-water masses sink, sometimes overlapping one another and often retaining their identity for long periods. Separate water masses below the pycnocline tend not to merge, because little energy is available for mixing in these quiet depths.

However, water does circulate in the ocean. The mass flow of ocean water in currents occurs in two forms: (1) surface currents are wind-driven movements of water at or near the ocean's surface; and (2) thermohaline currents (so named because they depend on density differences caused by variations in water's temperature and salinity) are the slow, deep currents that affect the vast bulk of seawater beneath the pycnocline (see later). Both have very important influences on Earth's temperature, climate, and biologic productivity, and will change as Earth's climate varies.

A small fraction of the water in the world ocean is involved in surface currents, comprised of water that flows horizontally in the uppermost 400 m (1300 feet) of the ocean's surface, driven mainly by wind friction. Most surface currents move water above the pycnocline.

The primary force responsible for surface currents is wind. Surface winds form global patterns within latitude bands. Most of Earth's surface wind energy is concentrated in each hemisphere's trade winds (i.e., easterlies) and westerlies. Waves on the sea surface transfer some of the energy from the moving air to the water using friction. This tug of wind on the ocean surface begins a mass flow of water, and the water flowing beneath the wind forms a surface current.

Because of the Coriolis effect, northern hemisphere surface currents flow to the right and southern hemisphere currents flow to the left of the wind direction. Continents and basin topography often block continuous flow and help deflect the moving water into a circular pattern, clockwise in the northern hemisphere and counterclockwise in the southern hemisphere. This flow around the periphery of an ocean basin is called a *gyre* (Greek *gyros,* "a circle").

There are six great current circuits in the world ocean, two in the northern hemisphere and four in the southern hemisphere. Five are geostrophic gyres, gyres that flow around the periphery of an ocean basin: the North Atlantic gyre, South Atlantic gyre, North Pacific gyre, South Pacific gyre, and Indian Ocean gyre. Although it is a closed circuit, the sixth and largest current is technically not a gyre because it does not flow around the periphery of an ocean basin. The West Wind Drift, or Antarctic Circumpolar Current, as this exception is called, flows endlessly eastward around Antarctica, driven by powerful, nearly ceaseless westerly winds. This greatest of all the surface ocean currents is never deflected by a continent. Figure 123-3 shows the major surface currents of the world ocean.

Assisted by the winds, surface currents distribute tropical heat worldwide. Warm water flows to higher latitudes, transfers heat to the air and cools, moves back to low latitudes, and absorbs heat again; the cycle then repeats. The greatest amount of heat transfer occurs at midlatitudes, where about 10^{15} calories of heat are transferred each second; this is 1 million times as much power as is consumed by the entire world's human population in the same length of time.

This combination of water flow and heat transfer from and to water influences climate and weather in several ways. For example, during the winter, Edinburgh, Dublin, and London are bathed in eastward-moving air only recently in contact with the

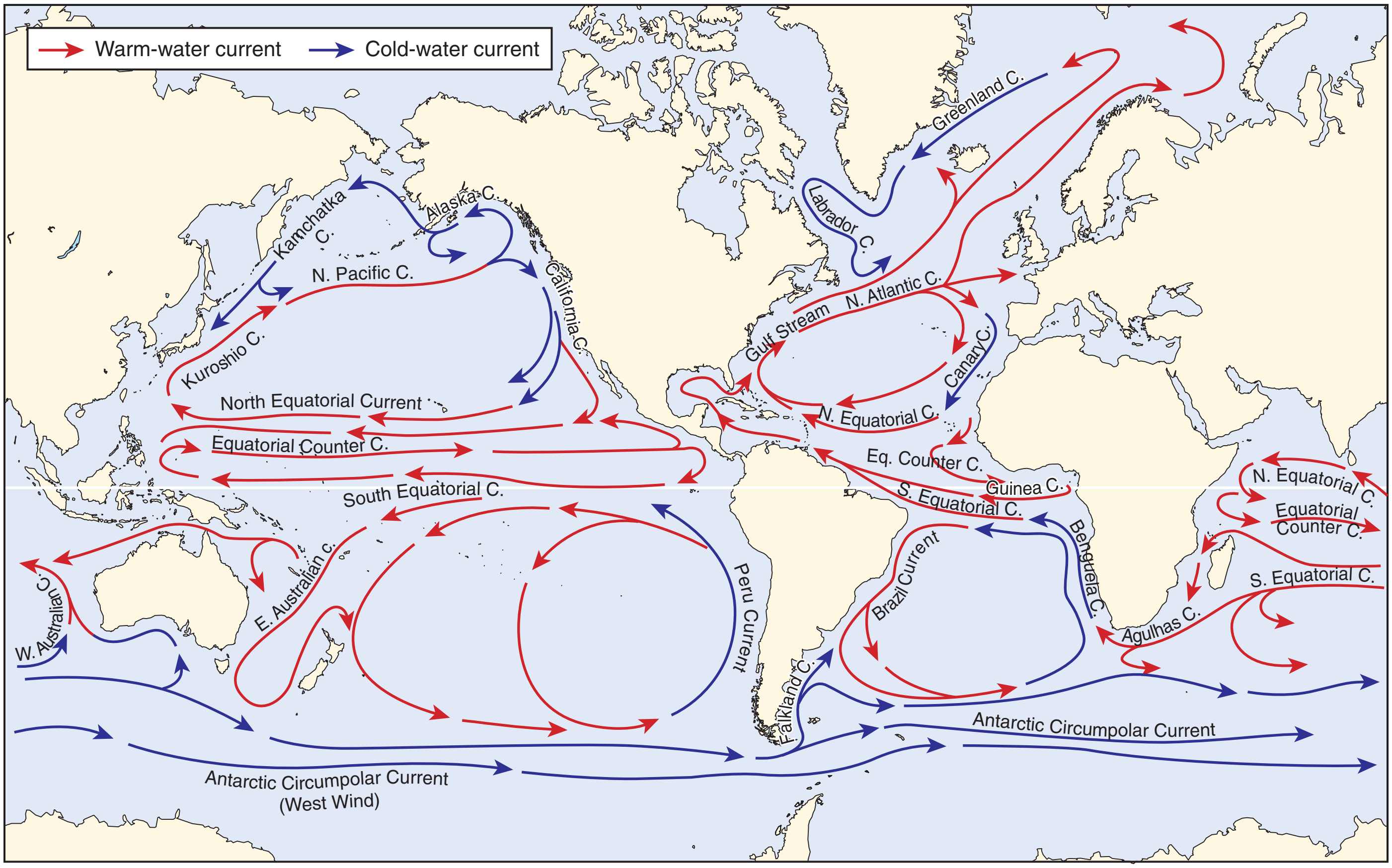


FIGURE 123-3 A chart showing the names and usual directions of the world ocean's major surface currents. The powerful western boundary currents flow along the western boundaries of ocean basins in both hemispheres.

relatively warm North Atlantic Current. Scotland, Ireland, and England have a maritime climate. These places are warmed in part by the energy of tropical sunlight transported to high latitudes by the Gulf Stream.

At lower latitudes on an ocean's eastern boundary the situation is often reversed. Mark Twain supposedly said that the coldest winter he ever spent was a summer in San Francisco. Summer months in that West Coast city are cool, foggy, and mild. Alternatively, Washington, DC—on nearly the same line of latitude as San Francisco but on the western boundary of an ocean basin—is known for its all-but-intolerable August heat and humidity. The California Current, carrying cold water from the north, comes close to the coast at San Francisco. Air normally flows clockwise in summer around an offshore zone of high atmospheric pressure. Wind approaching the California coast loses heat to the cold sea and comes ashore to chill San Francisco. Summer air often flows around a similar high off the East Coast (i.e., the Bermuda High). Therefore, winds that approach Washington, DC, blow from the south and east. Heat and moisture from the Gulf Stream contribute to the capital's oppressive summers. Alternatively, during the winter, Washington, DC, is colder than San Francisco, because westerly winds approaching Washington, DC, are chilled by the cold continent over which they cross.

Surface currents affect the uppermost layer of the world ocean (i.e., about 10% of its volume), but horizontal and vertical currents also exist below the pycnocline in the ocean's deeper waters. Because density is largely a function of water temperature and salinity, the movement of water as a result of differences in density is called *thermohaline circulation* (Greek *therme,* "heat," and *halos,* "salt"). The entire ocean is involved in slow thermohaline circulation, a process responsible for the large-scale vertical movement of ocean water and circulation of the global ocean as a whole.

Formation and downwelling of deep water occurs in the polar regions. Antarctic Bottom Water, the most distinctive of the deep-water masses, is characterized by salinity of 34.65‰, temperature of −0.5°C (30°F), and density of 1.0279 g/cm^3. This water is noted for its extreme density (it is the densest in the world ocean), the great amount of it produced near Antarctic coasts, and its ability to migrate north along the seafloor.

Most Antarctic Bottom Water forms near the Antarctic coast south of South America during winter. Salt is concentrated in pockets between crystals of pure water and then squeezed out of the freezing mass to form a frigid brine. Between 20 and 50 million m^3 of this brine form every second. The water's great density causes it to sink toward the continental shelf, where it mixes with nearly equal parts of water from the southern Antarctic Circumpolar Current.

The mixture settles along the edge of Antarctica's continental shelf, descends along the slope, and spreads along the deep-sea bed, creeping north in slow sheets. Antarctic Bottom Water flows many times more slowly than the water in surface currents: in the Pacific, it may take 1000 years for this water to reach the equator; 600 years later, it may be as far away as the Aleutian Islands at 50 degrees N latitude. Antarctic Bottom Water also flows into the Atlantic Ocean basin, where it flows north at a faster rate than in the Pacific. Antarctic Bottom Water has been identified as high as 40 degrees N latitude on the Atlantic floor, a journey that will have taken some 750 years to complete.

Similar water masses form in the North Atlantic and even at the Gibraltar outlet of the Mediterranean Sea. However, none is as dense as Antarctic Bottom Water. Oxygen is delivered to organisms in the deepest ocean basins by these slowly creeping water masses.

The great quantities of dense water sinking at polar ocean basin edges must be offset by equal quantities of water rising elsewhere. Figure 123-4 shows an idealized model of thermohaline flow.

Note that water sinks relatively rapidly in a small area where the ocean is very cold, but rises much more gradually across a very large area in the warmer temperate and tropical zones. It then slowly returns poleward near the surface to repeat the cycle. The continual diffuse upwelling of deep water maintains the existence of the permanent thermocline found everywhere at low latitudes and midlatitudes. This slow, upward movement is estimated to be about 1 cm (0.4 inch) per day over most of the ocean. If this rise were to stop, downward movement of heat would cause the thermocline to descend and would reduce its steepness. In a sense, the thermocline is "held up" by the continual, slow, upward movement of water.

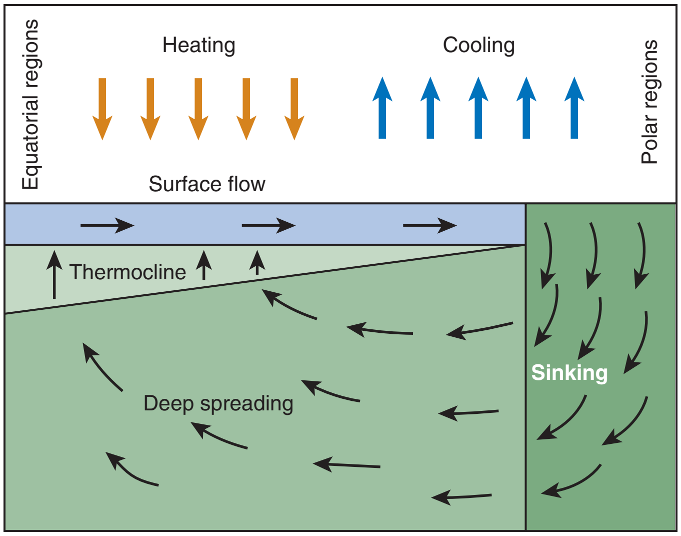


FIGURE 123-4 The classic model of a pure thermohaline circulation, caused by heating in lower latitudes and cooling in higher latitudes. Note that the low- and mid-latitude thermocline is "held up" by continuous replacement from below.

OCEAN MOVEMENT: WAVES, TIDES, AND TSUNAMIS

Wind waves, ocean tides, and tsunamis are expressions of energy moving across the ocean surface. Waves are disturbances caused by the movement of energy from a source through a medium (i.e., solid, liquid, or gas). As the energy of the disturbance travels, the medium through which it passes moves in specific ways. Sometimes this movement is visible as crests or ridges in the medium. The traveling crests produce the appearance of movement we see in a wave. In an ocean wave, a ribbon of energy is moving at the speed of the wave, but water is not. In a sense, the wave is an illusion.

Picture a resting seagull as it bobs on the wavy ocean surface far from shore. The gull moves in circles: up and forward as the tops of the waves move to its position, down and backward as the tops move past. Each circle is equal in diameter to the wave's height. Energy in waves flows past the resting bird, but the gull and its patch of water move only a very short distance forward in each up-and-forward, down-and-back wave cycle. The water on which the bird rests does not move continuously across the sea surface, as is suggested by the wave illusion. To clarify the important idea of wave as illusion, imagine yourself at a sports stadium where spectators are doing "the wave." Your role in wave propagation is simple: you stand up and sit down in precise synchronization with your neighbors. Although you move only a few feet vertically, the wave of which you were a part circles the arena at high speed. You and all the other participants stay in place, but the wave moves faster than anyone can run.

Transfer of energy from water particle to water particle in these circular paths or orbits transmits wave energy across the ocean surface and causes the waveform to move. This type of wave is known as an *orbital wave*; it is a wave in which particles of the medium (water) move in closed circles as the wave passes. Orbital ocean waves occur at the boundary between two fluid media (i.e., air and water) and between layers of water of different densities. Because the waveform moves forward, these waves are a type of progressive wave.

The progressive wave that moved the gull was probably caused by wind. Other forces can generate much greater progressive waves in which water molecules move through much larger circular or elliptical orbits. Some of these waves are so large that they do not appear to us as waves at all, but rather as the slow sloshing of water in a harbor or bay, as dangerous flooding surges of water, or as rhythmic and predictable ocean tides.

Ocean waves are classified by the disturbing force that creates them, the extent to which the disturbing force continues to influence the waves after they are formed, the restoring force that tries to flatten them, and their wavelength. (Wave height is not often used for classification, because it varies greatly depending on water depth, interference between waves, and other factors.)

Wind blowing across the ocean surface provides the disturbing force for *capillary waves* and *wind waves*. Arrival of a storm surge or seismic sea wave in an enclosed harbor or bay, or a sudden change in atmospheric pressure, is the disturbing force for the resonant rocking of water known as a *seiche*. Landslides, volcanic eruptions, and faulting of the seafloor associated with earthquakes are the disturbing forces for *seismic sea waves,* which are also known as *tsunamis*. The disturbing forces for tides are changes in the magnitude and direction of gravitational forces among Earth, Moon, and Sun in combination with Earth's rotation.

A wave that is formed and then propagates across the sea surface without the further influence of the force that formed it is known as a *free wave*. When wind waves move away from the storm that created them, or when the storm ceases, they continue without the injection of additional wind energy. Likewise, tsunamis—waves caused by submerged landslides or earthquakes—continue to move across the ocean surface long after the landslide or earthquake has stopped moving.

By contrast, a *forced wave* is maintained by its disturbing force. Tides are forced waves that depend on the gravitational attraction of the moon and sun.

Restoring force is the dominant force that returns the water surface to flatness after a wave has formed in it. If the restoring force of a wave were quickly and fully successful, a disturbed sea surface would immediately become smooth, and the energy of the embryo wave would be dissipated as heat. However, that is not what happens. Waves continue after they form because the restoring force overcompensates and causes oscillation. The situation is analogous to a weight bobbing at the bottom of a very flexible spring, constantly moving up and down past its normal resting point.

The restoring force for very small water waves (i.e., those with wavelengths of <1.73 cm [0.68 inch]) is *cohesion,* the property that enables individual water molecules to stick to each other by means of hydrogen bonds. The same force that makes tea creep up on the sides of a teacup tugs the tiny wave troughs and crests toward flatness.

All waves with wavelengths of more than 1.73 cm (0.68 inch) depend mostly on gravity to provide the restoring force. Gravity pulls the crests downward, but inertia of the water causes the crests to overshoot and become troughs. The repetitive nature of this movement, like the spring weight moving up and down, gives rise to the circular orbits of individual water molecules in an ocean wave. These larger waves are called *gravity waves*. Because the circular motion of water molecules in a wave is nearly free of friction, gravity waves can travel across thousands of miles of ocean surface without disappearing, eventually breaking on a distant shore.

Wavelength is an important measure of wave size. Table 123-3 lists the causes and typical wavelengths of capillary waves, wind waves, seiches, seismic sea waves, and tides.

Most characteristics of ocean waves depend on the relationship between their wavelength and water depth. Wavelength determines the size of the orbits of water molecules within a wave, but water depth determines the shape of the orbits. The paths of water molecules in a wind wave are circular only when the wave is traveling in deep water. A wave cannot "feel" the bottom when it moves through water deeper than half its wavelength because too little wave energy is contained in the small circles below that depth. Waves moving through water deeper than half their wavelengths are known as *deep-water waves*. A wave has no way of "knowing" how deep the water is, only that it is in water deeper than about half its wavelength. For example, a wind wave with a 20-m (66-feet) wavelength will act as a deep-water wave if it is passing through water deeper than 10 m (33 feet).

The situation is different for wind-generated waves close to shore. The orbits of water molecules in waves moving through shallow water are flattened by proximity of the bottom. Water just above the seafloor cannot move in a circular path, only forward and backward. Waves in water shallower than 0.05 their original wavelength are known as *shallow-water waves*. For example, a wave with a 20-m (66-feet) wavelength will act as a shallow-water wave if the water is less than 1 m (3.3 feet) deep.

Transitional waves travel through water deeper than 0.05 their original wavelength but shallower than one-half their original wavelength. For the previously discussed example, this would be water between 1 m (3.3 feet) and 10 m (33 feet) deep.

Of the five wave types listed in Table 123-3, only capillary waves and wind waves can be deep-water waves. To understand why, remember that most of the ocean floor is deeper than 125 m (400 feet), one-half the wavelength of very large wind waves. The wavelengths of the larger waves are much longer; the wavelength of seismic sea waves usually exceeds 100 km (62 miles). No ocean is 50 km (31 miles) deep, so seiches, seismic sea waves, and tides are forever in water that to them is shallow or transitional in depth. Their huge orbit circles flatten against a distant bottom that is always less than one-half a wavelength away.

In general, the longer the wavelength, the faster the wave energy will move through the water. For deep-water waves this relationship is shown in the formula:

$$C = L/T$$

where C represents speed (celerity), L is wavelength, and T is time or period in seconds.

The speed of all ocean waves is controlled by gravity, wavelength, and water depth. The speed of a deep-water wave may also be approximated by the formula:

$$C = \sqrt{gL}/2\pi$$

where g is acceleration due to gravity, which is 9.8 m/sec^2. Because g and π (3.14) are constants,

$$C = 1.251\sqrt{L}$$

TABLE 123-3 Disturbing Forces, Wavelengths, and Restoring Forces of Ocean Waves

Wave Type	Disturbing Force	Typical Wavelength	Restoring Force
Capillary wave	Usually wind	Up to 1.73 cm (0.68 inch)	Cohesion of water molecules
Wine wave	Wind over ocean	60-150 m (200-500 feet)	Gravity
Seiche	Change in atmospheric pressure, storm surge, or tsunami	Large and variable; a function of ocean basin size	Gravity
Seismic sea wave (tsunami)	Faulting of seafloor, volcanic eruption, or landslide	200 km (125 miles)	Gravity
Tide	Gravitational attraction or rotation of Earth	Half of Earth's circumference	Gravity

where C is measured in m/sec and L in meters. Note in both instances that wave speed is proportional to wavelength.

Wavelength is difficult to determine at sea, but period is comparatively easy to find. For example, an observer simply times the movement of waves past the bow of a stopped ship. If the period (T) is known, speed (S) can be calculated from the relationship:

$$\begin{aligned} C\ (\text{in m/sec}) &= gT/2\pi \\ &= 9.8\ \text{m/sec}^2 \times T\ (\text{in sec})/2 \times (3.14) \\ &= 1.56T \end{aligned}$$

where g is acceleration caused by gravity.

The speed of shallow-water waves is described by the equation:

$$C = \sqrt{gd} \text{ or } C = 3.1\sqrt{d}$$

where C is speed in m/sec, g is acceleration from gravity (9.8 m/sec^2), and d is the depth of the water in meters. The period of a wave remains unchanged regardless of the depth of water through which it is moving. However, as deep-water waves enter the shallows and touch bottom, their speed is reduced and their crests "bunch up," so their wavelengths shorten.

Comparing deep-water wind waves with shallow-water seismic sea waves is like comparing apples with oranges, but the following paragraphs demonstrate the general relationship between wavelength and wave speed: the longer the wavelength, the greater the speed. Remember that energy (not the water mass itself) is moving through the water at the astonishing speed of 760 km/hour (472.2 miles/hr) in seismic sea waves; this is the same speed as a jet airliner.

Long-wavelength, shallow-water progressive waves caused by rapid displacement of ocean water are called *tsunamis;* this is a descriptive Japanese term combining *tsu,* "harbor," with *nami,* "wave." Tsunamis that are caused by the sudden, vertical movement of Earth along fault lines (i.e., the same forces that cause earthquakes) are properly called *seismic sea waves.* Tsunamis can also be caused by landslides, icebergs falling from glaciers, volcanic eruptions, asteroid impacts, and other direct displacements of the water surface. Note that all seismic sea waves are tsunamis, but not all tsunamis are seismic sea waves.

Displacement of surface water by small seismic fractures causes "small" tsunamis. Although less energy is released by landslides than by most seismic fractures, the resulting sea waves are still very destructive for people or structures near their point of origin. This is especially true if the wave is formed within a confined area.

Seismic sea waves originate on the seafloor when Earth movement along fault lines displaces seawater. When the seismic sea wave in the Indian Ocean occurred on December 26, 2004, it ruptured along a submerged fault line, lifting the sea surface as much as 10 m (33 feet). Gravity pulled the crest downward, but the momentum of the water caused the crest to overshoot and become a trough. The oscillating ocean surface generated progressive waves that radiated from the epicenter in all directions. Waves would also have formed if the fault movement were downward. In that case, a depression in the water surface would have propagated outward as a trough; the trough would have been followed by smaller crests and troughs caused by surface oscillation.

It seems strange to refer to tsunamis—with wavelengths of up to 200 km (124.3 miles)—as shallow-water waves. However, one-half their wavelengths would be 100 km (62.1 miles), and even the deepest ocean trenches do not exceed 11 km (6.8 miles) in depth. Therefore, these immense waves never find themselves in water deeper than one-half their wavelengths. As with any shallow-water wave, seismic sea waves are affected by the contour of the bottom and are usually refracted, sometimes in unexpected ways. Detailed analysis of the 2004 event showed that the midocean ridges acted as topographic waveguides. These shallow-water waves were in constant contact with the seabed and appear to have followed the Southwest Indian Ridge below the southern tip of Africa to the Mid-Atlantic Ridge.

We are familiar with the steepness of a wind wave and the short period of a few seconds between its crests. Tsunamis are much different. After a tsunami is generated, its steepness (i.e., ratio of height to wavelength) is extremely low. This lack of steepness, in combination with the wave's very long period (i.e., 5 to 20 minutes), enables it to pass unnoticed beneath ships at sea. A ship on the open ocean that encounters a tsunami with a 16-minute period would rise slowly and imperceptibly for about 8 minutes, to a crest only 0.3 to 0.6 m (1 to 2 feet) above average sea level; it would then ease into the following trough 8 minutes later. With all the wind waves around, such a movement would not be noticed.

However, as the tsunami crest approaches shore, the situation changes rapidly and often dramatically. The period of the wave remains constant, its velocity drops, and the wave height greatly increases. As the crest arrives at the coast, observers would see water surge ashore in the same way as a very high and fast tide. In confined coastal waters relatively close to their points of origin, tsunamis can reach a height of 30 m (100 feet). The wave is a fast, onrushing flood of water and not the huge, plunging breaker of popular movies and folklore. Tsunamis can be catastrophic even if the wave crests are not that high; imagine a smaller wave inundating a flat, low-lying coast. The combination of wave height and "run up" (i.e., the distance the waves move ashore) determines a tsunami's lethality.

The wave energy spreads through an enlarging circumference as a tsunami expands from its point of origin. People onshore near the generating shock have reason to be concerned because the energy will not have dissipated much. Because of its low elevation and proximity to the earthquake epicenter, the Indonesian city of Banda Aceh was essentially demolished during the December 2004 event.

The same seismic sea wave reached the coast of India about 3 hours later. By this time, the wave circumference was enormous and its energy more dispersed. Even so, successive waves surged onto Sri Lankan, Indian, and African beaches at regular intervals for more than 2 hours.

Note that the destruction was not caused by one wave, but by a series of waves spaced at regular intervals. Some energy from the main tsunami wave was distributed into smaller waves ahead of or behind the main wave as it moved. If the epicenter of the displacement responsible for a tsunami is far away, sea level at shore will rise and fall as these waves arrive. The interval between crests (i.e., the wave period) is usually about 15 minutes. Coastal residents far from a tsunami's origin can be lulled into thinking the waves are over; they return to the coastline only to be injured or killed by the next crest. This behavior contributed to the enormous loss of life around the Indian Ocean.

Destructive tsunamis strike somewhere in the world an average of once each year. An earthquake along the Peru-Chile Trench on May 22, 1960, killed more than 4000 people; the associated tsunami reached Japan, 14,500 km (9010 miles) away, killing 180 people and causing $50 million in structural damage. Los Angeles and San Diego harbors in California were badly disrupted by seiches excited by the tsunami.

The catastrophic seismic sea wave that struck northern Japan on 11 March 2011 was the result of a rupture along a submerged fault that lifted the sea surface as much as 6 m (20 feet) in places. Gravity pulled the crest downward, but the momentum of the water caused the crest to overshoot and become a trough. The oscillating ocean surface generated progressive waves that radiated from the epicenter in all directions. Because of its low elevation and proximity to the earthquake epicenter, parts of the Miyagi and Iwate prefectures were inundated by the March 2011 wave. The Sendai region, adjacent to the epicenter, suffered the greatest damage. More than 20,000 people were killed or declared missing in the surrounding area. Nuclear power plants in the area were seriously damaged and leaked radioactive substances into air, land, and ocean. The economic loss will almost certainly amount to about 3% of a year's production by the world's third-largest economy, more than US$310 billion.

Modern tsunami warning systems depend on seabed seismometers and submerged devices and satellites that watch the shape of the sea surface.

CONDITIONS FOR OCEANIC LIFE

Marine organisms depend on the ocean's chemical composition and physical characteristics for life support. Any aspect of the physical environment that affects living organisms is called a *physical factor*. Living in the ocean often has advantages over living on land; for example, physical conditions in the sea are usually milder and less variable than physical conditions on land. The most important physical factors for marine organisms are light, temperature, dissolved nutrients, salinity, dissolved gases, acid-base balance, and hydrostatic pressure.

These physical factors work in concert to provide the physical environment for oceanic life. Additional biologic factors (i.e., biologically generated aspects of the environment) also affect living organisms. These biologic factors include diffusion, osmosis, active transport, and surface-to-volume ratio.

Too much or little of a single factor can adversely affect the function of an organism. Such a factor is called a *limiting factor,* a physical or biologic necessity with a presence that, in inappropriate amounts, limits the normal action of the organism. For example, in an ocean area where everything is perfect for photosynthesis (i.e., warmth, nutrients, adequate CO_2) except for light, no photosynthesis would occur, because light is the limiting factor. If light was present but nitrates absent, nitrate nutrients would be the limiting factor.

Marine organisms are often subject to great pressure from the constant weight of water above them, but hydrostatic pressure presents very little difficulty for them. In fact, the situation in the ocean is parallel to that on land. Land animals live in air pressurized by the weight of the atmosphere above them (i.e., 1 kg/cm^2).

Pressures inside and outside an organism are essentially the same, both in the ocean and at the bottom of the atmosphere. Thus, marine organisms do not require heavy shells to keep from being crushed by hydrostatic pressure. Great pressure has chemical effects: gases become more soluble at high pressure, some enzymes are inactivated, and metabolic rates for a given temperature tend to be slightly higher. However, these effects are felt only at great depth. Unless marine organisms have gas-filled spaces in their bodies (e.g., lungs, swim bladders), a moderate change in pressure has little effect.

All cells of every organism are enclosed by membranes, which are essentially complex films through which a few select substances can move. A cell's membranes are greatly affected by salinity of the surrounding water.

Salinity of seawater can vary in places as a result of rainfall, evaporation, runoff of water and salts from land, and other factors. Surface salinity varies most, with lows of 6% or less along the coast of the inner Baltic Sea in early summer, to year-around highs exceeding 40% in the Red Sea. Salinity is less variable with increasing depth, with the ocean typically becoming slightly saltier with depth.

Change in salinity can physically damage cell membranes, and concentrated salts can alter protein structures. Salinity can affect the specific gravity and density of seawater and therefore buoyancy of an organism. Salinity is also important because it can cause water to enter or leave a cell through the membrane, changing the cell's overall water balance. Seawater is nearly identical in salinity to the interior of all but the most advanced forms of marine life, so maintaining salt balance—and therefore water balance—is easy for most marine species.

Almost all marine organisms require dissolved gases, particularly CO_2 and oxygen, to stay alive. Oxygen does not easily dissolve in water, and as a result, about 100 times as much gaseous oxygen exists in the atmosphere as in the ocean. However, CO_2, which is essential to primary productivity, is much more soluble and reactive in seawater than is oxygen (Table 123-4).

Although up to 1000 times as much CO_2 as oxygen can dissolve in water, normal values at the ocean surface average around 50 mL/L for CO_2 and approximately 6 mL/L for oxygen. At present, the ocean holds about 60 times as much CO_2 as the atmosphere. Because of this abundance, marine plants almost never run out of CO_2.

TABLE 123-4 The Solubility of Gases in Seawater Decreases as Temperature Rises*

	Solubility (mL/L at Atmospheric Pressure and Salinity of 33‰)		
Temperature	**Nitrogen**	**Oxygen**	**Carbon Dioxide**
0° C (32° F)	14.47	8.14	8700.0
10° C (50° F)	11.59	6.42	8030.0
20° C (68° F)	9.65	5.26	7350.0
30° C (86° F)	8.26	4.41	6600.0

Modified from Walton-Smith FG: *CRC handbook of the marine sciences.* Cleveland, Ohio, 1974, CRC Press.

*Note that the data shown represent values at saturation.

Deep water tends to contain more CO_2 than does surface water. Why should this be? Table 123-4 shows the relationship between water temperature and its ability to dissolve gases; note that colder water contains more gas at saturation. You may recall that the deepest and most dense seawater masses are formed at the surface in the cold polar regions, and more CO_2 can dissolve in that low-temperature environment. The dense water sinks, taking its large load of CO_2 to the bottom, and the pressure at depth helps keep it in solution. CO_2 also builds in deep water because only heterotrophs (animals) live and metabolize there, and because CO_2 is produced as decomposers consume falling organic matter. No photosynthetic primary producers are present in the dark depths to use this excess CO_2, because there is not enough sunlight for photosynthesis to occur.

Rapid photosynthesis at the surface lowers CO_2 concentrations and increases the quantity of dissolved oxygen. Oxygen is least plentiful just below the limit of photosynthesis because of respiration by many small animals at middle depths.

Low oxygen levels can sometimes be a problem at the ocean surface. Plants produce more oxygen than they use, but they produce it only during daylight hours. The continuing respiration of plants at night will sometimes remove much of the oxygen from the surrounding water. In extreme cases, this oxygen depletion may lead to the death of the plants and animals in the area, a phenomenon most noticeable in enclosed coastal waters during spring and fall plankton blooms.

The greatest variability in levels of dissolved gas is found at the surface near shore. Less dramatic changes occur in the open sea.

Ocean temperature varies with depth and latitude. The average temperature of the world ocean is only a few degrees above freezing. Warmer water is found only in the lighted surface zones of the temperate and tropical ocean and in deep, warm, chemosynthetic communities. Although temperature ranges of the ocean are considerable, they are much narrower than comparable ranges on land.

What are the implications of the ocean's temperature for living things? The rate at which chemical reactions occur in an organism largely depends on the molecular vibration known as *heat*. Because agitation brings reactants together, warmer temperatures increase the rate at which chemical reactions occur. Thus, an organism's metabolic rate increases with temperature. The metabolic rate approximately doubles with each temperature rise of 10° C (18° F). The interior temperature of an organism is directly related to the rate at which it moves, reacts, and lives.

The great majority of marine organisms are ectothermic, which means that they have an internal temperature that stays very close to that of their surroundings. A few complex animals (e.g., mammals, birds, some of the larger and faster fishes) are endothermic; they have a stable, high internal temperature.

In general, the warmer the environment of an ectotherm within its tolerance range, the more rapidly its metabolic processes will proceed. Tropical fish in a heated aquarium eat more food and require more oxygen than goldfish of the same size living in an unheated but otherwise identical aquarium. Tropical fish generally grow faster, have a faster heartbeat, reproduce more rapidly, swim more swiftly, and live shorter lives. The

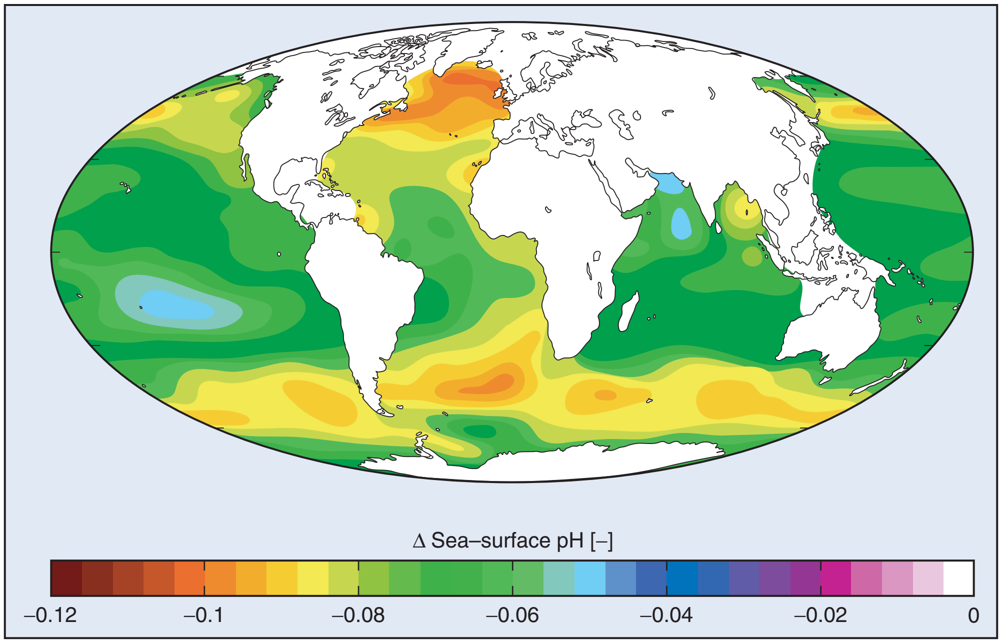


FIGURE 123-5 The ocean is becoming more acidic as it absorbs additional carbon dioxide from the atmosphere. A less alkaline environment will make it more difficult for organisms to build hard structures containing calcium (e.g., shells, coral). The chart shows changes in sea surface pH between the 1700s and the late 1990s.

upper limit of temperature that an ectotherm can tolerate is often not much higher than its optimum temperature. The lower limit is usually more forgiving because molecules are merely slowed.

Compared with ectotherms, endotherms can tolerate a tremendous range of external temperatures (e.g., a whale migrating from polar waters to the tropics, an emperor penguin incubating an egg at −51° C [−60° F]). However, their internal temperatures vary only slightly. Sophisticated thermal-regulation mechanisms make it possible for endotherms to live in a variety of habitats, but they pay a price. Their high metabolic rates make proportionally high demands on food supply and gas transport, but the benefit of having a biochemistry fine-tuned to a single efficient temperature is worth the regulatory difficulties involved.

Another physical condition that affects ocean life is the acid-base balance of seawater. The complex chemistry of Earth's life-forms depends on precisely shaped enzymes, which are large protein molecules that accelerate the rate of chemical reactions. When strong acids or bases distort the shapes of these vital enzymes, the chemical reactions they govern may not function normally.

Seawater's average pH is about 8. The dissolved substances in seawater act to buffer pH changes, preventing broad swings of pH when acids or bases are introduced. The normal pH range of seawater is much less variable than that of soil, and terrestrial organisms are sometimes limited by the presence of harsh alkali soils that damage cell components.

Although seawater remains slightly alkaline, it is subject to some variation. When dissolved in water, some CO_2 becomes carbonic acid. In areas of rapid plant growth, pH will rise because the plants use CO_2 for photosynthesis. Because temperatures are generally warmer at the surface, less CO_2 can dissolve. Surface pH in warm, productive water is usually around 8.5.

At middle depths and in deep water, more CO_2 may be present. Its source is animal and bacterial respiration. With cold temperatures, high pressure, and no photosynthetic plants to remove it, this CO_2 will lower the pH of water and make it more acid with increasing depth. Thus, deep, cold seawater below 4500 m (14,764 feet) has a pH around 7.5. This lower pH can dissolve calcium-containing marine sediments. A drop to pH 7 can occur at the deep ocean floor when bottom bacteria consume oxygen and produce hydrogen sulfide.

Concentrations of CO_2 are rising in the atmosphere. Much, perhaps most, of this increase comes from burning fossil fuels to support human industry and economic growth. Some researchers believe that this rapid increase in CO_2 could overwhelm the carbonate buffer system in surface waters and cause surface oceanic pH to drop. Even slightly more acidic seawater would interfere with formation of calcareous materials, such as coral skeletons, plankton tests, and some other hard parts of marine organisms. Figure 123-5 shows recent changes in oceanic pH.

MARINE PRIMARY PRODUCTIVITY

Primary productivity involves synthesis of organic materials from inorganic substances by photosynthesis or chemosynthesis. Primary productivity is expressed in grams of carbon bound into organic material per square meter of ocean surface area per year: $gC/m^2/yr$. The immediate organic material produced is the carbohydrate glucose, and dissolved CO_2 provides carbon for the glucose.

A *nutrient* is a compound required for an organism to produce organic matter. Some nutrients help form the structural parts of organisms, some make up the chemicals that directly manipulate energy, and some have other functions. A few of these necessary nutrients are always present in seawater, but most are not readily available.

The main inorganic nutrients required for primary productivity include nitrogen (as nitrate) and phosphorus (as phosphate). As any gardener knows, plants require fertilizer, mainly nitrates and phosphates, to grow. Ocean gardeners have more trouble raising crops than do their terrestrial counterparts, because the most fertile ocean water contains only about 0.0001 of the available nitrogen of topsoil. Phosphorus is even more scarce in the ocean, but fortunately, living things need less of it, because they have only about 1 atom of phosphorus for every 16 atoms of nitrogen.

Nitrogen and phosphorus are often depleted by autotrophs during times of high productivity and rapid reproduction. Also in short supply during rapid growth are dissolved silicates (used for shells and other hard parts) and trace elements such as iron and copper (used in enzymes, vitamins, and other large molecules). Marine plants must recycle these nutrients.

On land, most photosynthesis proceeds at or just above ground level. However, seawater, unlike soil, is relatively transparent, which allows photosynthesis to proceed for some distance below the ocean surface. At the same time, incoming sunlight must run a gauntlet of difficulties before it can be absorbed by chlorophyll in marine autotrophs.

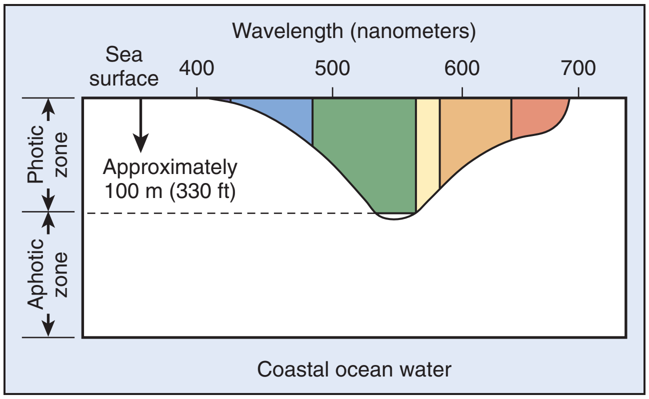


FIGURE 123-6 Because of the suspended particles often present in coastal waters, light cannot penetrate very far; approximately 100 m (330 feet) is typical. The sunlit upper zone is called the *photic zone,* and the dark ocean beneath is called the *aphotic zone.*

Most sunlight approaching at a low angle (e.g., near sunrise or sunset, in the polar regions) reflects off the water surface and does not enter the ocean. Light that penetrates the surface is selectively absorbed; water is more transparent to some colors of light than to others. In clear water, blue light penetrates to the greatest depth, whereas red light is absorbed near the surface. Light energy absorbed by water turns to heat.

The number and characteristics of particles in the water also limit the depth to which light penetrates. These particles—which may include suspended sediments, dust-like bits of once-living tissue, or the organisms themselves—scatter and absorb light. High concentrations of particles quickly absorb most blue and ultraviolet light. This absorption, in combination with the reflection of green light by chlorophyll within the producers, changes the color of productive coastal waters to green.

How far down does light penetrate? Figure 123-6 shows the depths reached by light of various wavelengths (i.e., colors) in the ocean. The *photic zone* is the uppermost layer of seawater lit by the sun. Because of the abundant small organisms and light-scattering particles, the photic zone near the coasts usually extends to about 100 m (330 feet), and in midlatitude waters it reaches down to about 150 m (500 feet). In clear tropical waters in the open ocean, instruments much more sensitive than human eyes have detected light at much greater depths; the present record is 590 m (1935 feet) in the tropical Pacific. The *aphotic zone,* the permanently dark layer of seawater beneath the photic zone, extends below the sunlit surface to the seabed. The vast bulk of the ocean is never brightened by sunlight.

Photosynthesis proceeds slowly at low light levels. Most of the biologic productivity of the ocean occurs in the upper part of the photic zone called the *euphotic zone* (Greek *eu,* "good"). This is the zone in which marine autotrophs can capture enough sunlight energy for plant primary production by photosynthesis to exceed the loss of carbohydrates by respiration. Although it is difficult to generalize about the ocean as a whole, the euphotic zone typically extends to a depth of approximately 70 m (230 feet) in the midlatitudes (as averaged over an entire year). The upper productive layer of ocean is a very thin skin indeed; the water within this zone amounts to less than 1% of world ocean volume, yet almost all marine life depends on this fine, illuminated band.

Phytoplankton, which are minute drifting photosynthetic organisms, produce between 90% and 96% of the surface ocean's carbohydrates. Seaweeds, which are larger marine photosynthesizers, contribute only 2% to 5% of the ocean's primary productivity. Chemosynthetic organisms probably account for 2% to 5% of the total primary productivity in the water column. Although estimates vary widely, recent studies suggest that total ocean productivity ranges from 75 to 150 gC/m^2/yr. For comparison, a well-tended alfalfa field produces about 1600 gC/m^2/yr.

How does marine productivity compare with terrestrial productivity? Recent research suggests the global net productivity in marine ecosystems is 35 to 50 billion metric tons of carbon bound into carbohydrates per year; global terrestrial productivity is roughly similar at 50 to 70 billion metric tons per year. However, the total producer biomass (i.e., the mass of living tissue) in the ocean is only 1 to 2 billion metric tons, compared with 600 to 1000 billion metric tons of living biomass on land. As the rapid turnover time indicates, nutrients cycle from producer to consumer and back much more quickly in marine ecosystems.

A primary producer's mass is assumed to be about 10 times the mass of the carbon it has bound into carbohydrates. Thus, a primary productivity of 100 gC/m^2/yr represents the yearly growth of about 1000 grams from primary producers for each square meter of ocean surface. Since 35 to 50 billion metric tons of carbon are bound into carbohydrates in the ocean each year, between 350 and 500 billion metric tons of marine plants and plantlike organisms are produced annually. Each year, the producers' metabolic activity and the consumers that graze on them consume this vast bulk. The component atoms are then reassembled by photosynthesis into carbohydrates in a continuous solar-powered cycle.

The extent of primary productivity by chemosynthesis within the seabed itself has been a surprise to researchers. High bacterial populations are present in some marine sediments to a depth of hundreds of meters. Samples have been taken at 842 m (2762 feet) below the seafloor in sediments 14 million years old. These bacteria are thriving in extreme conditions at these depths; they have high diversity and are well adapted to life in the subsurface. In fact, a single gram of rock may harbor 10 million bacteria.

These specialized organisms are usually called *extremophiles,* because they are capable of life under extreme conditions (Figure 123-7). Bacteria and similar organisms known as *archaea* have been found in fractured rocks more than 3 km (1.8 miles) below the surface of Africa at the same depth. As with bacteria, a single gram of rock may harbor 10 million archaea. Specialized organisms have been seen in hot oil reservoirs below the North Sea and the North Slope of Alaska (where they cause oil to "sour") and in volcanic rock 1220 m (4000 feet) below the surface of the island of Hawaii. Bacterial biomass in sediments and solid rock


© 2007 Thomson Higher Education

FIGURE 123-7 Deep-living chemosynthetic bacteria cultured from the minute spaces between mineral crystals in solid rock.

may represent at least 10% of the total known Earth surface biomass (although more than a few biologists believe this estimate to be low by at least an order of magnitude). Some of these organisms can tolerate the extreme temperatures found at hydrothermal vents.

Photosynthetic and chemosynthetic organisms can be called either *primary producers* or *autotrophs,* because they make their own food. The bodies of autotrophs are rich sources of chemical energy for any organisms capable of consuming them. *Heterotrophs* are organisms such as animals that must consume food from other organisms because they are unable to synthesize their own food molecules. Some heterotrophs consume autotrophs, and some consume other heterotrophs.

We can label organisms by their positions in a “who eats whom” feeding hierarchy called a *trophic pyramid.* The primary producers at the bottom of a trophic pyramid are mostly chlorophyll-containing photosynthesizers. The animal heterotrophs that eat them are called *primary consumers* or *herbivores;* the animals that eat these primary consumers are called *secondary consumers,* and so on, to the *top consumer* or *top carnivore.*

The mass of consumers becomes smaller as energy flows toward the top of the pyramid. In other words, there are many small primary producers at the base and very few large top consumers at the apex. Only about 10% of the energy from the organisms consumed is stored in the consumers as flesh, so each level is about one-tenth the mass of the level directly below. The rest of the energy is lost as waste heat as organisms live and work to maintain themselves.

Pyramid constructs can lead to the misconception that one type of fish eats only one other type of fish (and so on). Real communities are more accurately described as *food webs*; these are groups of organisms linked by complex feeding relationships in which the flow of energy can be followed from primary producers through consumers. Organisms in a food web almost always have some choices of food species.

IMPORTANT PLANKTONIC AUTOTROPHS

Autotrophic plankton that generate glucose by photosynthesis are generally called *phytoplankton* (Greek *phyton,* “plant”). A huge and nearly invisible mass of phytoplankton drifts within the euphotic zone, which is the productive sunlit surface layer of the world ocean. Although the water within the euphotic zone amounts to less than 2% of world ocean volume, most pelagic marine life depends on this fine, illuminated band.

There are at least eight major types of phytoplankton, the most prominent being the diatoms and dinoflagellates. As noted earlier, recent research suggests that very small producers, most of which are forms of cyanobacteria and archaea, may be responsible for much more oceanic primary productivity than are their larger and better-known counterparts.

Apart from cyanobacteria, the most productive photosynthetic organisms in the plankton are the diatoms. *Diatoms* evolved comparatively recently and began to dominate phytoplanktonic productivity during the Cretaceous Period, about 100 million years ago. Their abundance and photosynthetic efficiency increased the proportion of free oxygen in Earth’s atmosphere. More than 5600 species of diatoms are known to exist. The larger species are barely visible to the unaided eye. Most are round, but some are elongated, branched, or triangular (Figure 123-8).

For more effective light absorption, *chlorophyll,* which is the main photosynthetic pigment, is accompanied in diatoms by accessory pigments. These yellow or brown pigments give most diatoms a yellow-green or tan appearance. Diatoms store energy as fatty acids and oils, which are compounds that are lighter than their equivalent volume of water and assist with flotation. Buoyancy is a potential problem for diatoms because the weight of their heavy silica frustule seems at odds with their need to stay near the sunlit ocean surface. Oil floats, glass sinks, and a balanced amount of both reduces cell density and lightens the load. Not all diatoms, however, need to float. Many nonplanktonic species lie on shallow bottoms, where light and nutrients are able to support photosynthesis. These benthic species are almost always elongated (pennate) in shape.

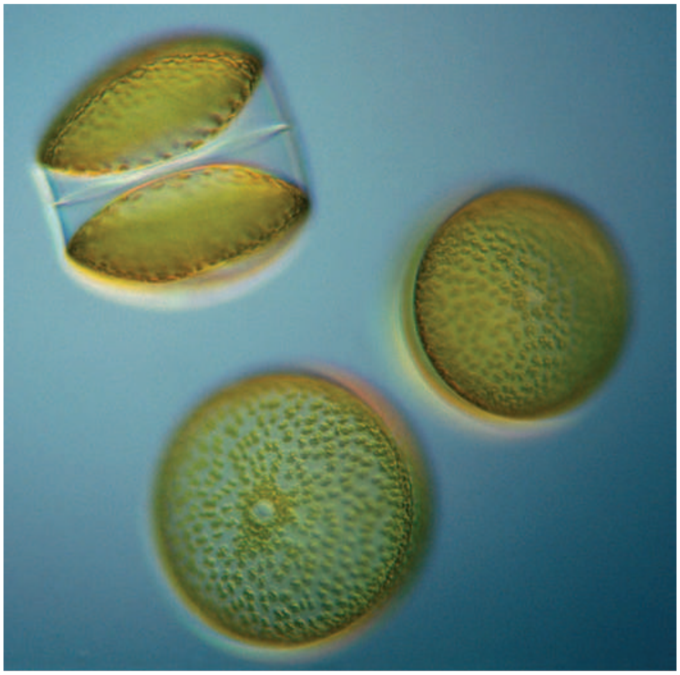


FIGURE 123-8 Diatoms of the genus *Coscinodiscus.* One is shown reproducing.

Like diatoms, *dinoflagellates* are single-celled autotrophs (Figure 123-9). Dinoflagellates are not as productive as diatoms, and they appear to have evolved much earlier. A few species live within the tissues of other organisms (e.g., the zooxanthellae of coral animals), but the majority of dinoflagellates live free in the water. Most have two whip-like flagella in channels grooved in their protective outer cell wall of cellulose. One flagellum drives the organism forward, whereas the other causes it to rotate in the water; thus the name, derived from the Greek word *dino,* meaning “whirling,” and the Latin *flagellum,* “whip.” These flagella allow dinoflagellates to adjust their orientation and vertical position to make the best photosynthetic use of available light or to move vertically in the water column to obtain nutrients.

Dinoflagellates are widely distributed, solitary organisms that reproduce by simple fission; they rarely form colonies. During reproduction, the cellulose covering that surrounds most species splits, and the single cell divides in half. Each daughter cell subsequently replaces the missing portion of covering. Under favorable conditions, the organisms can reproduce once a day, growing in number but not in size.

Some species of dinoflagellates can become so numerous that the water turns rusty red because light is reflecting from the accessory pigments within each cell. These species are usually responsible for the phenomenon referred to as a *red tide,* or

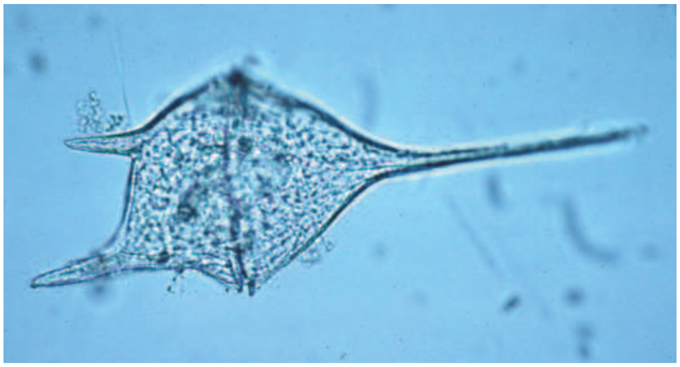


FIGURE 123-9 A dinoflagellate of the genus *Ceratium.* *(Courtesy Tom Garrison.)*

more generally, as a *harmful algal bloom* (HAB). During times of such rapid growth, which is usually in the spring, the concentration of microscopic planktonic organisms may briefly reach 6 to 8 million per liter.

HARMFUL ALGAL BLOOMS

An HAB occurs when high concentrations of phytoplankton adversely affect the physiology of nearby organisms. HABs do not always turn water red, and the organisms that cause them are not always visible. A number of factors are thought to contribute to HABs, including warm surface temperature, reduced salinity, optimal nutrient and light conditions, and a mechanism (e.g., gentle onshore winds) that physically concentrates the dinoflagellates.

Although the dinoflagellates responsible for most red tides are comparatively simple organisms, some have the ability to synthesize potent toxins as byproducts of metabolism. Among the most effective poisons known, these toxins may affect nearby marine life if ingested, or may even indirectly poison humans through the food chain. Some of the toxins are similar in chemical structure to the muscle relaxant curare, but much more potent (see Chapter 77).

The number and severity of HABs appear to be increasing. Perhaps this is not surprising; coastal waters receive industrial, agricultural, and domestic wastes, rich in nitrogen and other plant nutrients that stimulate algal growth. In addition, the long-distance transport of algal species in the ballast water of cargo vessels can introduce alien species into coastal waters, where they may thrive in the absence of the organisms that naturally consume them. Australia has recently issued strict guidelines for discharging ballast in the country's ports.

BIOGEOCHEMICAL CYCLES

The atoms and small molecules that make up the biochemicals—and thus the bodies—of organisms move between the living and nonliving realms in biogeochemical cycles. Living organisms are supported and sustained by huge, nonliving chemical reserves, and there is large-scale transport of elements between the reserves and the organisms themselves. Sometimes the environment contains enough of a required element to sustain life; sometimes the element is in short supply and is thus limiting. The tropical and temperate ocean is usually highly stratified, with a warm, less dense layer of water (the mixed zone) separated from the cold, dense, deep zone by a strong pycnocline. In the surface mixed layer, the atoms and small molecules that make up the bodies of organisms may cycle rapidly for a time among predators, prey, scavengers, and decomposers. When these organisms die, their bodies can sink below the sunlit upper sea and pycnocline to become isolated from the rapid biologic activity of the surface. Regions of upwelling are critical to returning these substances to the surface, if only for a short reprieve, before their eventual incorporation into deep sediments, from which only the very slow progress of tectonic cycles will liberate them.

As you read about the biogeochemical cycles described later in this chapter, remember that the elements and small molecules forming the tissues of an organism are always on the move. They may cycle rapidly in and out of living things, or they may be trapped in Earth for vast spans of time. However, the nature of the cycles dictates what will live where, which creatures will be successful, and ultimately, what will be the composition of the ocean and atmosphere.

The largest biogeochemical cycle is the global carbon cycle. Because of its ability to form long chains to which other atoms can attach, carbon is the basic building block of all life on Earth. Carbon enters the atmosphere as CO_2 through respiration of living organisms, volcanic eruptions that release carbon from rocks deep in Earth's crust, burning of fossil fuels, and other sources. When levels of atmospheric CO_2 are high, Earth's surface temperature rises as a result of the greenhouse effect.

Large and small plants and plantlike organisms capture sunlight and use this energy to incorporate, or fix, CO_2 into organic molecules. Some of these molecules are used as food, and some become structural components. When an animal eats a plant or plantlike organism, the carbon can do one of the following: (1) be incorporated into the animal's body for growth; (2) be respired by the animal (i.e., taken apart to harvest the energy); or (3) be excreted back into the seawater as what is called *dissolved organic carbon* (DOC). Typically, about 45% of the carbon from an ingested plant is used for growth, 45% is used for respiration, and 10% is lost as DOC. The end product of respiration is CO_2, a gas eventually lost to the atmosphere. Most of the DOC is rapidly used by bacteria, which are in turn eaten by protozoa, which are eaten by zooplankton, which are then eaten by fish; this is called the *microbial loop*. Eventually, the organisms (or at least their hard parts, containing calcium carbonate) sink below the mixed layer and begin the long fall toward the seabed. Most of the carbon in this calcium carbonate is turned into CO_2 by bacteria long before it hits the bottom, but a small percentage (<1%) reaches the sediments and is buried. The carbonate sediments can be uplifted over geologic time and weathered so that the carbon is eventually returned to the biologically active upper sea.

Because of the large amount of CO_2 available in the ocean, and because CO_2 from the atmosphere dissolves readily in seawater, marine organisms almost never suffer from a deficit of available carbon. For life in the sea, the critical bottlenecks lie elsewhere, mainly in the nitrogen, phosphorus, and iron cycles.

Nitrogen is a critical component of proteins, chlorophyll, and nucleic acids. Like carbon, nitrogen may be found in the bodies of organisms, as a dissolved gas, and as dissolved organic matter known as *dissolved organic nitrogen*.

One might think nitrogen would be abundantly available in the ocean, because nitrogen accounts for 48% of the dissolved gas in seawater by volume. However, most organisms cannot use free nitrogen in the atmosphere and ocean directly. It must first be bound with oxygen or hydrogen, or *fixed,* to usable chemical forms by specialized organisms, usually bacteria or cyanobacteria. Thus, oceanic regions are frequently nitrogen limited; growth of plants and plantlike organisms is often held back by lack of available nitrogen.

Forms of nitrogen available for uptake by living things are ammonium and nitrate, an ion formed by oxidation of ammonium and nitrite. Nitrate runoff from soil is an especially rich source of this often limiting nutrient, which explains why coastal water tends to support greater plankton populations than does oceanic water. Nitrogen is assimilated by small plants and plantlike organisms and then recycled as animals excrete ammonium and urea. These reduced forms of nitrogen are then oxidized back into nitrate, through nitrite, by nitrifying bacteria. In the deep ocean, most nitrogen is in the form of nitrate. In anoxic sediments and certain low-oxygen regions of the ocean, denitrifying bacteria use nitrate in respiration and convert nitrate back to nitrite and nitrogen gas, which is lost to the atmosphere. The other major loss occurs when nitrogen-containing organisms and debris are buried in ocean sediments.

Iron is used in minute quantities in the reactions of photosynthesis, in certain enzymes crucial to nitrogen fixation, and in the structure of proteins. Other essential trace metals, such as zinc, copper, and manganese, are also used by organisms in small quantities, primarily in enzymes. Although in absolute terms, organisms require only tiny quantities of iron, the concentration of iron in seawater relative to the concentration of nutrients, such as nitrogen and phosphorus, can sometimes be so low that phytoplankton growth is limited by the availability of iron. Although iron is one of the most abundant elements in Earth's crust, it is nearly insoluble in oxygenated seawater, and the little dissolved iron that is present is highly reactive, sticks to falling particles, and sinks to the bottom of the water column.

In general, the biogeochemical cycles of the trace metals follow the pattern of uptake and recycling in the surface ocean and regeneration, sometimes over long periods of time, at depth. However, much remains to be learned about the interactions between living organisms and trace metals. Iron and other trace metals exist in many chemical forms in seawater. Discovering what these forms are, how transformation between forms occurs, and availability of different forms to marine organisms are major foci of current research in trace metal biochemistry.

MARINE ENVIRONMENTAL ISSUES

Human demand has exceeded Earth's ability to regenerate resources since at least the early 1980s. Since 1961, human demand on Earth's organisms and raw materials has more than doubled and now exceeds Earth's natural replacement capacity by at least 20%. Our present rate of consumption is clearly unsustainable. The ocean's great volume and relentless motion dissipate and distribute natural and synthetic substances. For this reason, humans have long used the sea as a dump for wastes. The ocean's ability to absorb is not inexhaustible, however, and the ocean is being severely affected by human activity.

A *pollutant* causes damage by interfering directly or indirectly with the mechanical or biochemical processes of an organism. Many pollutants are harmful to human health. Some pollution-induced changes may be instantly lethal; other changes may weaken an organism over weeks or months, alter the dynamics of the population of which it is a part, or gradually unbalance the entire community.

An organism's response to a particular pollutant depends on its sensitivity to the combination of *quantity* and *toxicity* of that pollutant. Some pollutants are toxic to organisms in tiny concentrations. For example, photosynthetic ability of some species of diatoms is diminished when chlorinated hydrocarbon compounds are present in parts-per-trillion quantities. Other pollutants may seem harmless, as when fertilizers flowing from agricultural land stimulate plant growth in estuaries. However, these pollutants may be hazardous to certain organisms but not others. For example, crude oil interferes with the delicate feeding structures of zooplankton and coats the feathers of birds, but simultaneously serves as a feast for certain bacteria.

Pollutants also vary in their *persistence*; some reside in the environment for thousands of years, whereas others last only minutes. Pollutants may break down into harmless substances spontaneously or through physical processes (e.g., shattering of large molecules by sunlight). Sometimes pollutants are removed from the environment through biologic activity. For example, some marine organisms escape permanent damage by metabolizing hazardous substances to harmless ones. Indeed, many pollutants are ultimately biodegradable; that is, they can be broken down by natural processes into simpler compounds. However, many pollutants resist attack by water, air, sunlight, or living organisms, because the synthetic compounds of which they are composed resemble nothing in nature.

Determining the ways in which pollutants are changing the ocean and the atmosphere is often difficult for researchers. Environmental impact cannot always be predicted or explained. As a result, marine scientists vary widely in their opinions about what pollutants are doing to the ocean and atmosphere and what to do about it. Environmental issues are frequently emotional, and media reports tend to sensationalize short-term incidents (e.g., oil spills) rather than more serious, long-term problems (e.g., climate change, effects of long-lived chlorinated hydrocarbon compounds). Figure 123-10 summarizes sources of marine pollution.

OIL POLLUTION

Public perception equates marine pollution with oil spills. Oil is a natural part of the marine environment. Oil seeps have been leaking large quantities of oil into the sea for millions of years; indeed, natural seeps are the largest source of oil in the ocean. The amount of oil entering the sea has increased in recent years, however, because of our growing dependence on marine transportation for petroleum products, offshore drilling, nearshore refining, and street runoff carrying waste oil from automobiles.

The world's accelerating thirst for oil is currently running at about 3800 L *per second*, slightly more than half of which is transported to market in large tankers. In the 1990s, approximately 1.3 million metric tons of oil entered the world ocean each year. Natural seeps accounted for almost half this annual input, or 600,000 metric tons. About 8% of the total was associated with marine transportation. Some of this oil was not spilled in well-publicized tanker accidents but was released during loading, discharging, and flushing of tanker ships. Between 150,000 and 450,000 marine birds are believed to be killed each year by oil released from tankers.

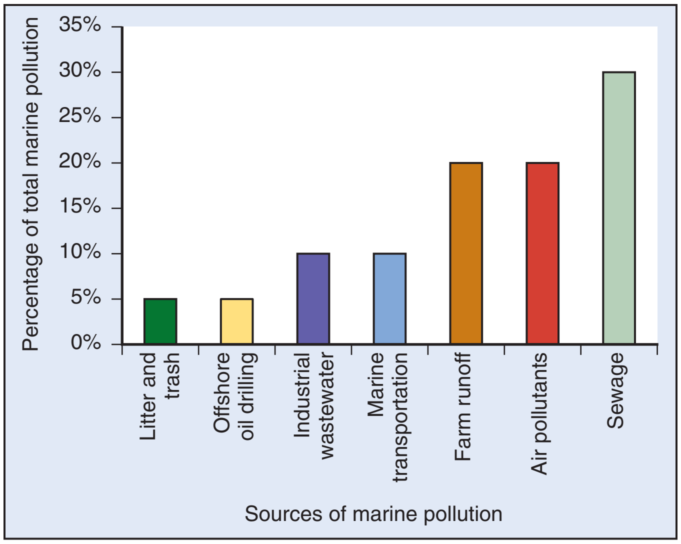


FIGURE 123-10 Sources of marine pollution. *(Data from Garrison, T: Oceanography: An invitation to marine science, 8th ed, Independence, Ky, 2013, Cengage Learning.)*

Much more oil reaches the ocean in runoff from city streets, or as waste oil dumped down drains, poured into dirt, or hidden in trash destined for a landfill. Each year, more than 900 million liters (about 240 million gallons) of used motor oil—about eight times the volume of the 2010 Gulf of Mexico spill—finds its way to the sea. This used oil is much more toxic than crude oil because it has developed carcinogenic and metallic components from the heat and pressure within internal combustion engines. Not all hydrocarbon pollution is "wet." Aromatic compounds released when crude oil evaporates eventually find their way back into the ocean.

Spills of *crude* oil are generally larger in volume and more frequent than spills of refined oil. Most components of crude oil do not dissolve easily in water, but those that do can harm the delicate juvenile forms of marine organisms, even in minute concentrations. The remaining insoluble components form sticky layers on the surface that prevent free diffusion of gases, clog feeding structures in adult organisms, kill larvae, and decrease sunlight available for photosynthesis. Crude oil is ultimately biodegradable. Although crude oil spills look terrible and generate great media attention, most forms of marine life in an area recover from the effects of a moderate spill within about 5 years.

Spills of *refined* oil, especially near shore where marine life is abundant, can be more disruptive for longer periods. The refining process removes and breaks up the heavier components of crude oil and concentrates the remaining lighter, more biologically active ones. Components added to oil during the refining process also make it more toxic. Spills of refined oil are of growing concern because the amount of refined oil transported to the United States rose dramatically through the 1980s and 1990s.

The volatile components of any oil spill eventually evaporate into the air, leaving the heavier tars behind. Wave action causes the tar to form into balls of varying sizes. Some of the tar balls fall to the bottom, where they may be assimilated by bottom organisms or incorporated into sediments. Bacteria eventually decompose these spheres, but the process may take years to complete, especially in cold polar waters. This oil residue, especially if derived from refined oil, can have long-lasting effects on seafloor communities.

The methods used to contain and clean up an oil spill sometimes cause more damage than the oil itself. Detergents used to disperse oil are especially harmful to living things. An accident on the *Deepwater Horizon* oil platform off the Louisiana coast on 20 April 2010 resulted in the largest accidental release of oil in the history of the U.S. petroleum industry. Eleven workers were killed, 17 injured, and the rig itself failed and sank. Oil under high geologic pressure shot from the stump of the broken drill pipe. The best estimate of the total release of oil during the 85 days required to plug the well is 4.9 million barrels (206 million gallons) of crude oil, of which 800,000 barrels (33 million gallons, about 16%) were captured by direct recovery from the drill head. Skimming and burning at the surface were also employed, recovering an additional 8%. The remainder of the oil dispersed into the surrounding ocean and atmosphere. About 26% of it formed tar balls, washed ashore, was buried in sediments, or remained as surface sheen. Lighter components of the oil, about 25%, evaporated or dissolved in seawater.

About 8% of the oil was dispersed by chemicals injected at the wellhead or sprayed on the ocean surface. The dispersants were toxic to the very organisms that would metabolize the oil naturally. A vast amount of dispersant, mostly the chemical Corexit, was deployed; estimates suggest that 2 million gallons were sprayed over the water or injected directly into the gushing wellhead on the seafloor. When mixed with the dispersant, oil that would normally float was able to linger far below the surface and affect fishes and bottom-dwelling organisms. Research continues, but it has been suggested the dispersed oil was more likely to be toxic than the crude oil by itself.

The best way to deal with oil pollution is to prevent it from happening. Tanker design is being modified to limit the amount of oil intentionally released in transport. Oil companies limit new tanker construction to stronger, double-hull designs, and platforms to contain redundant fail-safe components. During the past decade, improved production technology and safety training of personnel have significantly reduced both blowouts and daily operational spills. Currently, accidental spills from platforms represent about 1% of petroleum discharged in North American waters and about 3% worldwide.

PLASTIC WASTE

Plastic waste is another serious and growing problem. Approximately 134 million metric tons of plastic are produced each year, of which 10% ends up in the ocean. Americans use more plastic per person than any other nationality. Americans generate about 31 million metric tons of plastic waste, about 120 kg (264 lb) per person, each year. They consume an average of 167 plastic bottles of water per person per year—about 25 million per hour! Slightly more than 4% of world oil production goes into the manufacture of plastics.

The attributes that make plastic items useful to consumers—durability and stability—make them a problem in marine environments. Scientists estimate that certain forms of synthetic materials, such as plastic six-pack holders, will not completely decompose for 400 years. Although oil spills receive more attention as a potential environmental threat, plastic is a much more serious danger. Oil is harmful, but plastic does not biodegrade.

The problem is not confined to the coasts. The North Pacific subtropical gyre covers a large area of the Pacific where the water slowly circulates in a clockwise direction. Winds are light. The currents tend to move any floating material into the low-energy center of the gyre. There are few islands on which this floating material can beach, so it remains in the gyre. This area, about the size of Texas, has been dubbed the “Eastern Pacific Garbage Patch.” A smaller, western Pacific equivalent has formed midway between San Francisco and Hawaii; another lies off the U.S. East Coast. One researcher estimates the weight of the debris trapped in gyres to be about 3 million metric tons, comparable to 1 year’s deposition at Los Angeles’s largest landfill.

Hundreds of marine mammals and thousands of seabirds die each year after ingesting or being caught in plastic debris. Sea turtles mistake plastic bags for jellyfish prey and die from intestinal blockages. Seals and sea lions starve after becoming entangled in nets or muzzled by six-pack rings. The same rings strangle fish and seabirds. About one-quarter million Laysan albatross chicks die each year when their parents feed them bits of plastic instead of food (see Figure 123-10).

Plankton productivity is adversely affected by plastics. Sunlight, wave action, and mechanical abrasion break plastic into ever smaller particles. This microfine plastic debris tends to attract oily toxic residues such as polychlorinated biphenyls, dioxin, brominated flame retardants, and other noxious organic chemicals. In the middle of the Pacific Ocean, 1 million times the amount of toxins are concentrated on the plastic debris and plastic particles (e.g., microbeads used in mildly abrasive skin cleaners) than are estimated to reside in ambient sea water. The microscopic plastic particles outweighed zooplankton by six times in water taken from the North Pacific subtropical gyre. Filter-feeding zooplankton mistake plastic particles for food, and the attendant synthetic chemicals may be interfering with aspects of phytoplankton physiology.

Not all plastic floats. About 70% of discarded plastic sinks to the ocean bottom. In the North Sea, Dutch scientists have counted 110 pieces of litter for every square kilometer of seabed, or about 600,000 metric tons in the North Sea alone. These plastics can smother benthic life-forms.

CHAPTER 124

Brief Introduction to Forestry

DONALD L. GREBNER AND PETE BETTINGER

Although forests have provided material, safety, and welfare for thousands of years, only in the last 100 years have humans begun to manage forests carefully. Early humans used forests as a source of shelter and products to sustain primitive communities. More recently, humans have used forests to develop modern economies. For most of human history, humans have only minimally impacted forested resources, largely because of low and widely dispersed population densities. More recently, human populations have expanded dramatically, leading to increased interactions between people and forested landscapes.

When the first European colonists arrived at the eastern seaboard of North America, settlers could only marvel at the immensity of the forested landscape.[6] For centuries, it seemed as if this resource had no limit and could never be depleted. Humans entered the forest and extracted material to build houses, hunted game to provide food, and used wood to support smelting of iron ore.[8] Settlers converted forested lands to develop agricultural fields to produce crops, which were used to feed themselves and expand communities. A culture of sustainability was not inherent.

Forested landscapes may pose risks to people as they travel in and around them. These landscapes can be flat or steep, wet or dry, and densely vegetative or sparsely covered. Types of birds, mammals, insects, reptiles and amphibians, and fish found in forests vary depending on latitude, elevation, and physiographic features. Size and shape of water bodies found within forested landscapes vary greatly, from small ponds to large lakes, and from seasonal streams that only flow at certain times of the year to swift rivers carrying melted-snow water from mountainous peaks. The landscapes vary depending on location, which can be extremely important in case of medical emergencies and search and rescue operations. Landowners and land managers who care for these landscapes can have important impacts on risks assumed by visitors. Small, private forest landowners, as well as larger industrial landowners, typically have easy access to their properties. However, in certain areas, particularly in the eastern part of North America, access to these lands is often restricted. State, province, and county governments that control public forests typically allow unimpeded access, as do national forests and parks. Many privately owned forested areas in Scandinavia allow free public access. Larger forests in less densely populated and remote areas often have poorer access characteristics (e.g., roads may not be well maintained), which can pose critical challenges to search and rescue teams as well as to medical evacuation.

DEFINITIONS

FORESTRY

Forestry involves management of forests to meet economic, ecologic, or social goals. Early references to the term *Forstwesen*, which is similar to “forestry,” can be traced to Europe in the early to mid-19th century, in association with publications such as the *Swiss Forestry Journal* (now called *Schweizerische Zeitschrift fur Forstwesen*). Forestry as a profession began in Europe during the Middle Ages, when royalty was interested in controlling which persons hunted wildlife on crown lands.[2] In Germany, exclusive hunting rights in forests were provided to nobility from the ninth through the 18th centuries.[5] Given humankind’s long-term use of forests to locate food, it was an important task for foresters to conserve a sufficient supply of animals for their countries’ inhabitants (in particular their nobles) to hunt. At present, perceptions of forests are much different. In some countries (e.g., Scotland), forestry practice outcomes are now among the lowest environmental concerns of local societies (i.e., when compared to treatment of human waste or nuclear technology), but forest practices remain important social concerns with respect to climate change.[4]

In Europe, the transition from managing forests as wildlife habitat to timber production preceded that within the United States. This transition did not begin in the United States until the middle to later 1800s. Early European settlers viewed the North American forested landscape as endless and placed little emphasis on resource management. Attention focused on resource extraction, as was pursued with nonrenewable resources (e.g., coal, petroleum). Early American foresters trained in Europe, such as Bernhard Fernow (Prussia) and Gifford Pinchot (France), led the way in developing the profession of forestry in the United States. Pinchot was instrumental in founding the Society of American Foresters (SAF).

The Dictionary of Forestry (http://dictionary of forestry.org) of the SAF defines *forestry* as “the profession embracing the science, art, and practice of creating, managing, using, and conserving forests and associated resources for human benefit and in a sustainable manner to meet desired goals, needs, and values.”[3] This definition has changed over time, becoming more complex with increased recognition of the benefits, not previously considered, that forests provide to individuals as well as to human society.

During the late 1800s and early 1900s, there were growing concerns about excessive harvesting of trees from forested landscapes. This contributed to development of scientific forestry as well as efforts to establish a national park system that focused on protection of “wild” forested landscapes. These events led to debate over whether public forests should be “conserved” and managed for human needs, or whether they should be “preserved” and managed to minimize human interaction with their innate resources. Advocates of the “conservation” perspective included Pinchot, and advocates of the “preservation” perspective included John Muir. Debate over management of forests is ongoing. Whether forest are managed or preserved can have significant implications for structure and health of a forested landscape, as well as for how people interact with the forest.

For much of the 1900s, forestry schools taught students such skills as how to manage a forest, focusing on extraction of wood, how to reforest cutover areas, and how to implement fire protection strategies. During the early phases of the forestry profession’s development, American society was largely agrarian in nature. Since that time, more than 85% of the U.S. population now resides in an urban environment. This dramatic change in U.S. demographics has important implications for how people view forests and forested landscapes. These changes have also impacted how foresters and other natural resource professionals manage forested landscapes and ecosystems, for what are often competing uses.

FOREST

What defines a forest can vary depending on standards set by societies, as well as the perspective of individuals or organizations tasked with managing them. For some, any area that contains trees can be viewed as a forest. For others, a stand of trees at the north end of their cornfield is simply considered a woodlot, not a forest. The size of a forest can play a critical role in shaping one’s perspective on what is or is not a forested landscape. Some forests can be 100 acres, about 40 hectares (ha), whereas others can be 100,000 acres (~40,470 ha) or even as large as a 1 million acres (~404,700 ha). The SAF dictionary defines a forest as “an ecosystem characterized by a more or less dense and extensive tree cover, often consisting of stands varying in characteristics such as species composition, structure, age class, and associated process, and commonly including meadows, streams, fish and wildlife.”[3]

The U.S. Forest Service[7] uses a different definition, as found in its Forest Inventory and Analysis (FIA) program:

> Land that is at least 10 percent stocked with trees of any size, or that formerly had such tree cover and is not currently developed for a nonforest use. The minimum area for classification of forest land is one acre. The components that make up forest land and all noncommercial forest land.

This definition is consistent with that used by the Food and Agriculture Organization (FAO) of the United Nations[1]:

> Land with tree crown cover (or equivalent stocking level) of more than 10 percent and area of more than 0.5 hectares (ha). The trees should be able to reach a minimum height of 5 meters (m) at maturity *in situ*. May consist either of closed forest formations where trees of various stories and undergrowth cover a high proportion of the ground; or open forest formations with a continuous vegetation cover in which tree crown cover exceeds 10 percent. Young natural stands and all plantations established for forestry purposes which have yet to reach a crown density of 10 percent or tree height of 5 m are included under forest, as are areas normally forming part of the forest area which are temporarily unstocked as a result of human intervention or natural causes but which are expected to revert to forest.

The FAO definition is similar to others but is more specific. Forests are basically areas with a number of trees that vary in size and shape. They can include other features (e.g., bodies of water, meadows, open areas) and can accommodate wildlife, insects, and fish. In many cases, humans live there.

Wilderness

A forest's classification as *wilderness* depends on remoteness of a forested area, size of forested area, and presence (or lack of) human structures as well as the presence (or lack of) human activity. Typically, wilderness areas are those where forests are not actively managed. A wilderness can be situated in a forested landscape, but not all wilderness areas are forested areas. Wilderness areas are managed from a "preservationist" perspective, which allows virtually no human interference in ecologic processes occurring on those properties. The SAF forestry dictionary refers to the Wilderness Act of 1964 in its definition of a wilderness area[3]: "A wilderness, in contrast with those areas where man and his works dominate the landscape, is hereby recognized as an area where the earth and its community of life are untrammeled by man, where man himself is a visitor who does not remain."

In the United States, a wilderness area is an area that the U.S. Congress has designated as such, regardless of the land's past use. The Sleeping Bear Dunes is a great example of a recently designated wilderness area. Located along Lake Michigan, the dunes are places where people have often recreated. Others, such as the Sipsey Wilderness Area located within the Bankhead National Forest in northwestern Alabama, are largely devoid of human structures except around their boundaries. The Sipsey Wilderness Area contains one remnant road that is being slowly reclaimed by nature as time passes, and the area hosts a wide collection of diverse plant and animal species, beautiful rock formations, waterfalls, and pristine river ways. It is also a popular hiking and camping destination.

Key differences between wilderness and nonwilderness areas center on the amount of human activity and whether human presence is transitory. Regardless of a forest's official designation, wilderness and nonwilderness forests both have the same attributes. All forests have a distinctive vertical structure of vegetation dominated by trees. In mature forests, trees dominate the *overstory,* the uppermost portion of the forest. This canopy section can be home to different types of wildlife, including birds and insects, as well as mammals such as flying squirrels. Beneath the canopy layer are three other layers of forest. The next level down from the canopy, the *midstory,* comprises trees that are either more shade tolerant or suppressed by the overstory canopy. These trees are shade tolerant and live and grow in lower levels of sunlight. Trees that are shade intolerant require more sunlight and are more likely to be found in the canopy layer. The next layer is called the *understory,* or *shrub layer,* and includes smaller trees as well as shrubs. Not all forests have this layer. Forests with extensive shrub layers can be difficult to navigate. In some parts of the United States, landowners may burn their properties because it makes it easier for them to travel through the forest. Below the shrub layer are found the *herbaceous layer* and *grass layer,* which are located just above the soil. These lower layers provide food and cover to numerous wildlife species. Their presence can depend on sufficient levels of sunlight reaching the forest floor. Not all forest structures are alive. Dead trees, known as "snags," may be standing in the midstory or overstory, and downed logs may be lying on the ground.

Trees and Rainfall

Types of tree species are an important aspect of forests. Which tree species exist in a forest depends not only on region of the world, but also on latitude and other physiographic and climatic factors (e.g., elevation, average local rainfall, average seasonal temperatures). A forest in Maine may be composed of species such as balsam fir (*Abies balsamea*), white spruce (*Picea glauca*), red maple (*Acer rubrum*), sugar maple (*Acer saccharum*), quaking aspen (*Populus tremuloides*), white (paper) birch (*Betula papyrifera*), and American beech (*Fagus grandifolia*). As one moves further south, changes in average annual temperatures prohibit many of these species from being found at lower latitudes or elevations. In Mississippi, a common tree species is loblolly pine (*Pinus taeda*). It can be found throughout the southern United States, but as one moves further north, loblolly pine becomes less common because it is susceptible to damage caused by ice storms.

Rainfall is a crucial factor in determining structure and species composition of a forest. In general, the eastern United States receives more rainfall than does the western part, except along the Pacific Coast. Areas west of the Rocky Mountains (e.g., northern California to Washington) are typically wetter than the eastern sides of the Cascade or Sierra Nevada Ranges. Redwood (*Sequoia sempervirens*) trees in northern California need sea fog for optimal growth, whereas ponderosa pines (*Pinus ponderosa*) are generally found much farther inland and often east of the Cascade or Sierra Nevada Ranges and can thrive in much drier climates.

Factors affecting forest type and structure are important for a variety of reasons. First, abundance of different types of species can impact what can be used for human consumption. Different tree species have different useful properties to humans, animals, and insects. The forest structure can affect who enters the forest and why they would do so. Drier forests tend to possess fewer trees that are more widely spaced. This landscape can be more prone to fire. In addition, drier forests provide fewer opportunities for finding water, which would be important to hikers and campers.

Where substrate (soil) conditions are favorable, and where seeds or seedlings (if planted) are provided, forests will become established. The establishment phase can occur after a natural disaster or following human activity. Many conifers are shade-intolerant trees. Full sunlight (e.g., large openings) may be required to successfully reestablish coniferous forests. Shade-tolerant trees may not require full sunlight for successful reestablishment. If left unattended, forests progress through stages of "succession." In the last stage of forest succession, the *climax stage,* the forest ecosystem perpetuates itself indefinitely through minor changes in forest structure, including individual tree mortality and regrowth of trees in the resulting gaps created. Currently, the climax stage of forests is often not attained because of human intervention.

TYPES OF FORESTS

Forests naturally occur in areas that are conducive to tree growth and are influenced by factors such as latitude, elevation, soil condition, and precipitation. Different environments lead to very different types of forests. The three major forest biomes found around the world are the boreal, temperate, and tropical. Each contains ecosystems with similar climatic characteristics. Boreal and temperate biomes contain forests with trees that become dormant in the winter. Tropical biome contains forests that have a year-long growing cycle.

Boreal forests are found in higher latitudes (e.g., Canada and Russia in northern hemisphere) and mainly contain conifers such as spruces and firs (Figure 124-1). Growing season is generally short, and forests can be located on relatively dry sites (e.g., western Alberta) or wet sites (peat bogs). *Montane* (i.e., high-elevation forests) can be found in the boreal biome or in temperate and tropical biomes. These forests also typically contain conifers; a good example is the dry coniferous forests of the Rocky Mountain range in the intermountain western United States (Figure 124-2). For these forests, elevation influences the amount of precipitation available for trees to grow.

Temperate deciduous forests contain oaks, maples, hickories, and much broader variety of broad-leaved trees that are typically found in the Appalachian Mountains and northeastern part of the United States (Figure 124-3), as well as vast areas of northern Europe. Temperate coniferous forests contain pines or firs; a good example of these are the pine forests along the southeastern U.S. Coastal Plain (Figure 124-4). Between these two forest types and across the eastern U.S. and northern Europe are mixed temperate coniferous and deciduous forests. In the U.S. Pacific Northwest, temperate rain forests contain fir and hemlock trees (Figure 124-5). Precipitation patterns influenced by the Pacific Ocean result in high annual rainfall amounts, and tree growth in this type of forest can be substantial. Throughout California and southern Europe, forests grow in climates that typically have wet winters and long, dry summers. These are often characterized as woodlands, chaparral, or Mediterranean forests (Figure 124-6).

FIGURE 124-1 Boreal forest, northern Minnesota. *(Courtesy Steven Katovich, USDA Forest Service. www.Bugwood.org.)*

FIGURE 124-2. Montane forest. *(Photo from Wikimedia Commons; courtesy Walter Siegmund. https://commons.wikimedia.org/wiki/File:Clover_Lake_8556.JPG; licensing agreement: https://creativecommons.org/licenses/by-sa/3.0/deed.en.)*

FIGURE 124-3 Temperate deciduous forest, northern United States. *(Courtesy Joseph O'Brien, USDA Forest Service. www.Bugwood.org.)*

FIGURE 124-4 Temperate pine forest, south Georgia. *(Courtesy Chuck Bargeron, University of Georgia. www.Bugwood.org.)*

Tropical deciduous forests can be found along the east coast of South America and the east coast of India, among other locations. Tropical rain forests are located throughout the Amazon Basin and Central America (Figure 124-7). Precipitation occurs nearly year-round in these forests. Monsoon tropical forests are somewhat different and involve a prolonged dry season and short rainy season. Cloud forests have a consistent low-level cloud cover at the tree canopy level (Figure 124-8). Examples of these can be found in Hawaii and Central America.

The gradient between forests and prairies or deserts can contain savannah forests (Figure 124-9). These areas are often grassy and contain sparse amounts of trees. Examples are the xeric scrublands of the intermountain western United States and subtropical deserts.

FIGURE 124-5 Temperate rain forest, Alaska. *(Courtesy Dave Powell, USDA Forest Service, retired. www.Bugwood.org.)*

FIGURE 124-6 Mediterranean forests. *(Photo from Wikimedia Commons; courtesy Ramessos. https://commons.wikimedia.org/wiki/File:KleineSchweizIsrael.jpg; licensing agreement: https://creativecommons.org/licenses/by-sa/3.0/deed.en.)*

FIGURE 124-7 Tropical rain forest, Honduras. *(Courtesy Howard F. Schwartz, Colorado State University. www.Bugwood.org.)*

FIGURE 124-8 Cloud forest, Ecuador. *(Courtesy Paul Bolstad, University of Minnesota, www.Bugwood.org.)*

FIGURE 124-9 Savannah forests. *(Photo from Wikimedia Commons; courtesy Gentry George, US Fish and Wildlife Service. https://commons.wikimedia.org/wiki/File:Bald_top_oak_savannah.jpg; licensing agreement: public domain.)*

IMPORTANCE OF FORESTS TO PEOPLE

Forests provide services to people and their communities. What people desire from forests can change over time depending on the changes in technology. During the early development of human society, forests provided wood used to cook food as well as to generate warmth. Wood was used to create tools to create weapons for hunting and self-defense. Early human societies depended heavily on hunting and gathering foodstuffs from the forest. This can still be seen in modern times in remote areas of the world where indigenous peoples are found. Forests in both early and modern human societies have played an important role in spiritual and religious practices. Many people prefer a particular forest area that, when visited, provides a sense of calm and peace.

Many types of products or commodities come from forests.[2] For example, "wood products" can be considered to be solid wood, pulp and paper, composites and engineered woods, chemicals, and non–timber-based products. Solid-wood products include items such as boards, beams, posts, fuel wood, charcoal, piles, posts, furniture, barrels, kegs, casks, musical instruments, sailing masts, tool handles, weapons, ice hockey sticks, gun stocks, flooring, canoes, bowls, utensils, pencils, chests, and knickknacks.

Paper production began in China almost 2000 years ago. The process has evolved over time so that now a wide variety of pulp and paper products are produced. Common products include writing and copying paper, printing paper, books, magazines, newspapers, post-it (or sticky) notes, calendars, envelopes, maps, business cards, holiday cards, photographic paper, and wrapping paper. Wood pulp is often used for creating napkins, newspapers, magazines, books, maps, toilet tissue, diapers, and paper towels. Other paper-based products include the corrugated cardboard containers widely used as packaging.

Composites and engineered products are created from small pieces of wood that are arranged and glued together to create a resource with a specific, desired property. A well-known composite and engineered product is plywood, which is made from peeled layers of wood that are overlaid with each other and glued together. Particleboard is another composite-based product made from chips, sawdust, and shavings. Oriented strand board is made from wood wafers, and newer products (e.g., decking, seating, and in some cases sidewalks) are currently made from combining wood fibers and plastic polymers to create wood-plastic composites.

Forests provide chemicals and residues. Some common chemicals that have been created include creosote (a wood preservative), acetone (a solvent), and acetic acid (used to make wood glue). Other chemicals produced include formic acid, butyric acid, propionic acid, methanol, turpentine, and various other oils and acids. Naval stores were typically extracted from specific tree species, such as pines, to provide resin, pitch, and tar for shipyards. In addition, lye is created by leaching potash from deciduous tree species and can be used in making soap.

Although wood-based products derived from forest resources are extremely important, many non–timber-based products provide spiritual, medicinal, food-based, or recreational value to humans and society. In certain U.S. areas, people enjoy extracting sap from (tapping) maple and birch trees to create syrups. In Europe, mushroom picking is a common forest-based activity. Some people collect medicinal plants, such as yarrow (*Achillea millefolium*), wild ginger (*Asarum canadense*), mayapple (*Podophyllum peltatum*), ginseng (*Panax quinquefolius*), witch hazel (*Hamamelis virginiana*), and sarsaparilla (*Smilax regelli*).

Wildlife habitat is an important nontimber forest product. Wildlife habitats are environments where vertebrates and invertebrates procreate and obtain shelter, food, and water. Many landowners view wildlife habitat as important because it can provide conditions favorable for supporting preferred species. Encouraging wildlife habit may promote better hunting, animal- and bird-watching opportunities, and protection of endangered species. Recreation is an important nontimber benefit. In forested landscapes, this usually refers to activities such as camping, hiking, backpacking, mountain biking, canoeing, fishing, hunting, bird watching, rock climbing, zip lining, and spelunking. Another important nontimber product is provision of water from forested ecosystems.

HUMAN INTERACTION WITH FORESTS

Foresters or natural-resource professionals interact with the forest in different ways, depending on the management objectives for a forested landscape. One of the most common objectives for managing a forested landscape is to maximize a financial return. This typically entails extraction of biologic material from the landscape. This objective is achieved through extraction or harvesting of trees to be sold to various types of forest products manufacturing facilities. The harvesting system employed depends on type of topography, forest access, and type(s) of tree species. Areas that are flat typically employ harvesting systems that use feller-bunchers to harvest trees (Figure 124-10) and then skidders to drag the trees to a landing area close to an access road. In steeper terrain, it is common to see loggers using chainsaws for felling trees and in the western United States, employing cable logging systems (Figure 124-11). These systems use a tower and cables to drag logs up to the roadside for placement on trucks.

People visit forests for recreational purposes. Hunting is a common activity. Although considered a recreational sport in many areas, foresters consider hunting an important tool for managing wildlife populations. The decline of natural predators, such as wolves (*Canis lupus*), cougars (*Puma concolor*), and bears in many areas, as well as changes in vegetative communities helps promote expansion of white-tailed deer (*Odocoileus virginianus*) and elk (*Cervus canadensis*) populations. On smaller forested properties, one may find hunting camps in remote locations. In the United States, large landowners, such as real estate investment trusts (REITS), forest products companies, and timber investment management organizations (TIMOS), often lease large portions of their lands to hunt clubs to generate additional revenues, as well as recruit people to monitor activities on forested properties. Game wardens, who typically are state employees, are responsible for monitoring hunting activities on state lands.

FIGURE 124-10 Feller-buncher. *(Photo courtesy USDA Forest Service. www.Bugwood.org.)*

FIGURE 124-11 High-lead cable logging system with loader in the foreground, tower and carriage in background. *(Courtesy Donald L. Grebner, Mississippi State University.)*

As we noted, recreation activities situated in forests include day hiking, backpacking, camping, rock climbing, mountain biking, skiing, spelunking, bird watching, and mushroom and berry picking. Topography can dictate the types of potential activities. For example, mountainous areas are popular for rock climbers, spelunkers, skiers, backpackers, and campers. Time of year can be equally important. Mushroom and berry pickers forage in the woods during growing seasons (summer months). Bird watchers depend on location and time of year and on species linked to those landscapes. Especially in state or national forests, natural-resource professionals actively promote these recreational opportunities.

Foresters measure various natural resources throughout the year. Successful forest management requires considerable data. Foresters identify living plant species, location, and size and density (heights and diameters). They measure amount of dead vegetation, both standing as snags and as the volume of dead trees lying on the forest floor. Foresters measure these conditions over time to evaluate forest growth and health. Foresters conduct field surveys to measure density and populations of wildlife. These animals may include common white-tailed deer or endangered and threatened species, such as the American bald eagle (*Haliaeetus leucocephalus*) or red-cockaded woodpecker (*Leuconotopicus borealis*). Natural-resource professionals and other scientists often monitor water quality in remote rivers and streams and in lakes and ponds.

Forest fires, whether wildfires started by lightning strikes or arson or planned prescribed fires, are of special interest to foresters. When wildfires are found in remote locations, firefighters known as “smoke jumpers” will parachute into “hot spots” in an attempt to put out a fire quickly before it spreads. If they are unsuccessful, large numbers of people may be employed to build firebreaks around burning areas to control a fire’s spread. Aircraft equipped to carry water or fire-retardant chemicals may be used to extinguish fire or suppress a fire. Wildfires can be greatly influenced by the amount of dry vegetation on the forest floor, topography, and general weather conditions. Prescribed fires employ many of the same personnel and equipment but are typically planned events for which firebreaks and other safeguards are put in place before the fire is started.

HAZARDS TO PEOPLE IN FORESTS

People may face extensive hazards in forested landscapes. Both natural-resource professionals and laypersons can become disoriented and lost, especially in forests with dense understory and overstory levels. Misdirection can lead to panic and feelings of being lost, which can lead to poor decision making and injuries. Foresters are trained to use a compass, and many carry Global Positioning System (GPS) instrumentation and cellular phones with GPS capabilities. However, these tools may not always be available or may fail. Nonprofessionals venturing into the forest may not have these tools but may possess a cell phone with GPS tracking capability.

Forests pose risk for injuries incurred by falling or by being hit by falling objects. Risks are increased when traveling off trail or on uneven terrain where it is difficult to see hidden holes covered by vegetation. Walking on top of downed trees is particularly hazardous. Falls can result in both blunt trauma (i.e., against the ground or tree) and penetrating trauma (i.e., impalements on broken stems or branches) that can be fatal. Travelers in the forest may be exposed to falling rocks and debris upslope and falling branches and trees under canopies. Falling branches that are loosened from a tree's crown through wind events are called "widow makers." Foresters can be injured or killed by falling dead trees or snags, especially while evaluating their suitability for wildlife habitat. Additional hazards include hypothermia, hyperthermia, and exposure to flash floods.

Traveling through uneven topography can be physically stressful and pose hazards. Difficulty of traveling through a forest is influenced by the number of live trees, composition of the understory, and density of fallen tree branches. Managers of recreation facilities typically develop trails to reduce the impact of these impediments and to provide higher-quality experiences for people wanting to walk through forests. In undeveloped forests, considerable exertion may be required to navigate through a forest, particularly in areas with steep or uneven topography. Field foresters are therefore typically required to be in good physical condition.

Forests also pose risk of exposure to infectious diseases. In North America, these include giardiasis, West Nile virus disease, Rocky Mountain spotted fever, rabies, and Lyme disease. Insects, mammals, and reptiles may pose hazards. Bee stings can be life threatening. Animal bites can produce lacerations and infections. Large animals, such as deer, moose (*Alcesalces*), or mountain goats (*Oreamnos americanus*), may attack a person if threatened. Their size, horns, and hooves may lead to fatal blunt and penetrating trauma. It is prudent to use care when stepping over logs or other woody debris in uneven terrain to avoid bites and envenomation. Toxic plants pose hazards. In North America, the most common toxic plant is poison ivy (*Toxicodendron radicans*).

Forest management operations (e.g., tree harvesting, road building, invasive insect and plant control) pose unique hazards. Heavy machinery can crush a person's limb or body; chainsaws and axes can cause puncture wounds and severe lacerations. Vehicles can roll over on steep slopes. Someone may be struck because of standing in the path of a falling tree. Where harvesting operations use overhead cable systems for dragging logs up steep terrain, cables can snap under high tension, and logs can also slip from their harnesses (chokers) and crush persons. Heavy equipment is a hazard during road construction. Also invasive insect and plant control operations also pose hazards because of the extensive use of chemicals.

Forest fires are extremely hazardous to foresters and others working, living, or recreating in forested areas (see Chapter 14). Firefighters use heavy equipment and chainsaws to create firebreaks, which are critical to slowing progression of or containing a fire. Hilly terrain and changes in wind patterns pose threats to firefighters, who may become trapped by rapidly advancing fires. Firefighters can employ "back fires" initiated upwind of a human-made or natural firebreak that burn available fuel back to the wildfire and thus widen the firebreak.

REFERENCES

Complete references used in this text are available online at expertconsult.inkling.com.

CHAPTER 125

Brief Introduction to Earth Sciences

WAYNE D. RANNEY

The planet Earth is a rocky sphere covered in a rich veneer of life. Even without its biologic life, however, Earth could still be considered alive because of the way that it creates heat in its interior. Although sometimes overlooked or even taken for granted, this heat engine creates a multitude of tangible effects on Earth's surface, including earthquakes and volcanoes felt and witnessed by all living beings. Escape of heat from the earth's interior drives the motions of the continents (which move at about the rate our fingernails grow) to create the mountains and ocean basins that cover the globe. Formation of these uplifted and down-warped features on Earth's surface determines circulation patterns in the world's oceans, which in turn redistributes warm water heated in the tropics toward the more temperate and polar regions. These ocean circulation patterns work in conjunction with feedback from various orbital phenomena, including the sun and the inclination of Earth's spin axis, to drive our planet's climate system. Tropical forests, arid and sandy deserts, temperate growing regions, and polar ice caps are all the result of this system. All life on Earth is affected by this multitude of inputs, which originate from the simple starting point that the earth generates its own heat and can be considered "alive" on the inside. The moon, as well as some other planetary bodies, is not alive in this sense. This chapter explores some of the products of existence on a living planet. Civilization and all life-forms respond to, and are affected by, the sometimes dramatic effects that are felt and experienced on planet Earth.

EARTH'S ORIGIN

Recent advances in astronomical measurements have allowed for consensus on the age, origin, and life cycle of the universe and its component parts. These observations into deep space reveal

REFERENCES

1. Food and Agriculture Organization of the United Nations. FRA 2000 terms and definitions, FRA Working Paper 1, FAO Forestry Department. <http://www.fao.org/docrep/006/ad665e/ad665e06.htm>; 1998.
2. Grebner DL, Bettinger P, Siry JP. Introduction to forestry and natural resources. New York: Academic Press; 2013.
3. Helms JA. The dictionary of forestry. Bethesda, Md: Society of American Foresters; 1998.
4. Hinds K, Carmichael K, Snowling H. Public attitudes to the environment in Scotland 2002. Research Findings No 24. Edinburgh: Environmental and Rural Affairs Research Programme, Scottish Executive Social Research; 2002.
5. Hocker VR. Beginn, höhepunkt und ende der kurfürstlichenjagd im Rheinland und Westfalen. Zeitschr Jagdwissensch 1997;43(2):105–15.
6. Perlin J. A forest journey: The story of wood and civilization. Woodstock, Vt: The Countryman Press; 2005.
7. US Forest Service. Common definitions used by the Forest Inventory and Analysis, <http://www.fs.fed.us/ne/fia/methodology/def_ah.htm>; 2013.
8. Williams M. Americans and their forests: A historical geography. New York: Cambridge University Press; 1989.

that the Big Bang occurred about 13.8 billion years ago (it is possible at this time to observe outward to about 380,000 years after the Big Bang). Considering the antiquity of the age of the universe, it could be said that our relatively "average" solar system is *only* about one-third of this age, or about 4.566 billion years old.

Our sun and its solar system formed as did many other common stars, from a *nebula* (Latin for "cloud"), which is composed of the scattered remains of previously existing stars (Figure 125-1). Stars have a life cycle from birth to death that begins in nebulae (Figure 125-2). As the far-flung material begins to collapse inward and coalesce, it swirls into a *protostar* that is surrounded by a disk of gas and dust. The gas and dust particles begin to adhere to one another and eventually grow into *protoplanets.* It is believed that only 1 million years elapsed from the initial collapse of a nebula to the birth of our protostar and the protoplanets. Then, for about 50 to 100 million years, these protoplanets collided with various neighbors to create the *planetesimals.* During this time, a large planetesimal hit Earth with a glancing blow to create the moon. Planetesimals in turn collided to form the planets we know today. The solar system is about halfway through its life cycle. About 5 billion years from now, the sun will have used up its inheritance of hydrogen fuel and begin to die, first becoming a red giant, and then exploding into the surrounding region to become the interstellar dust and gas that will eventually collapse into future stars. The cycle will begin anew.

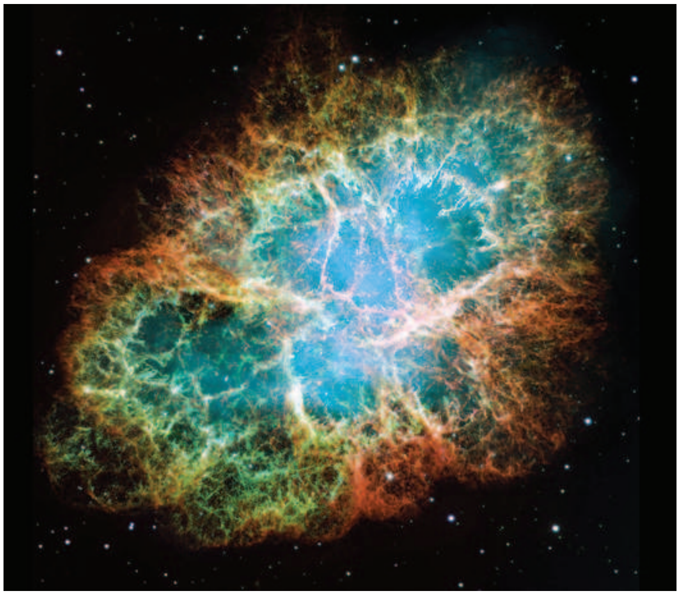


FIGURE 125-1 The Crab Nebula is an example of the material that once made up a solar system. The light from this nebula reached Earth in the year 1054, but it is 6500 light-years from Earth, meaning that it exploded about 7500 years ago.

EARTH'S INTERIOR STRUCTURE

To the geologist, Earth can be considered "alive" because of the way in which it generates its own heat, causing its interior to roil and convulse and its surface to be subjected to constant motion and change. Physical expressions of this living nature are the many earthquakes and volcanoes that shake our existence on this restless planet. The moon is an example of a "dead" planet in that it does not generate its own heat and thus does not presently have active volcanoes or tectonic earthquakes. Our planet extends from its core to the outer reaches of the atmosphere (the atmosphere can be considered a part of the layered Earth). The planet is technically an oblong spheroid, meaning that it bulges slightly at the equator and is squashed minimally at the poles. It has a radius of nearly 6400 km (4000 miles) and is layered into concentric shells with the heaviest, densest layers at the center and the least dense on the outside. The contrast in density between these layers is the result of differences in the composition of each layer and increasing pressure toward the center. In composition from the center outward, our rocky Earth contains

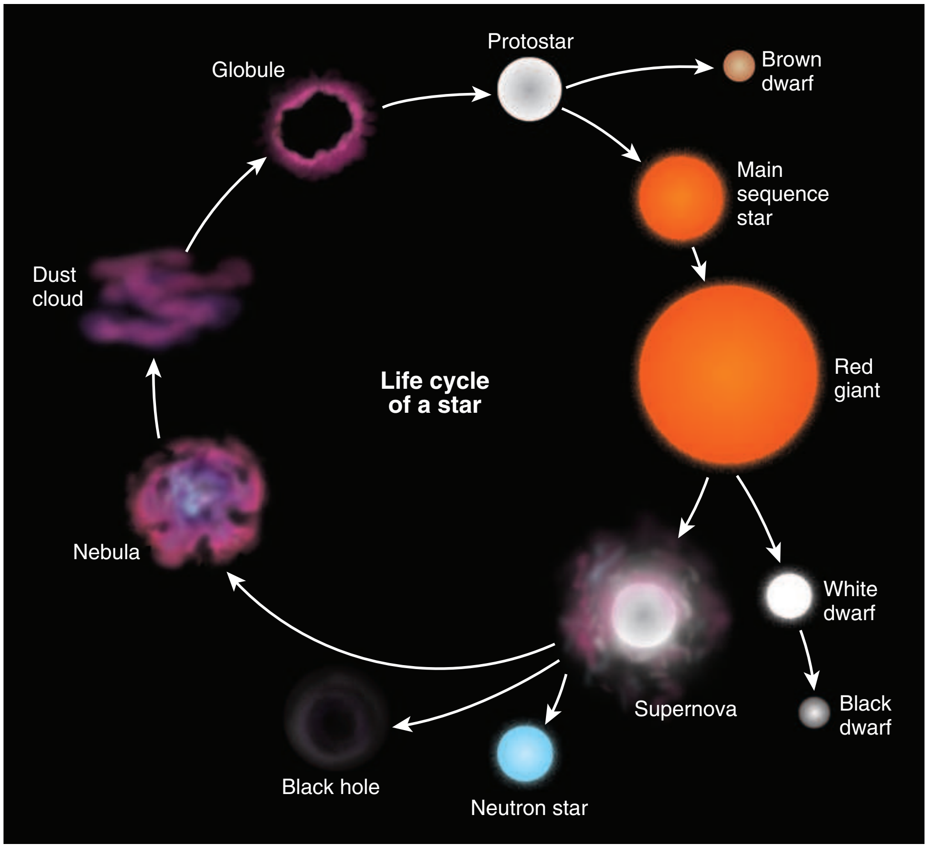


FIGURE 125-2 Life cycle of a typical star, such as our sun.

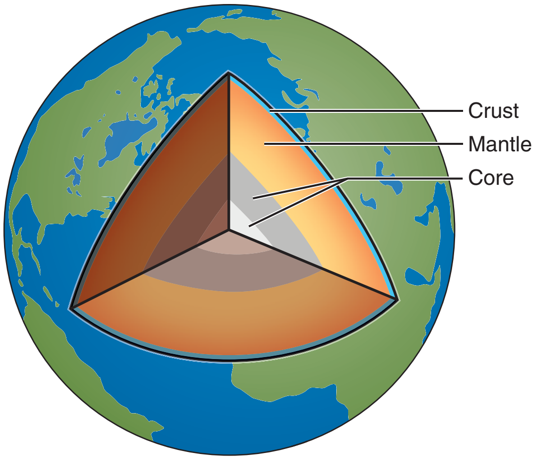


FIGURE 125-3 Earth's concentric layers consist of the core (both inner solid and outer liquid), mantle, and crust.

the core, mantle, and crust. A second scheme differentiates the earth's interior based on the variable properties that the rocks exhibit pertaining to their depth and pressure. These layers do not correspond precisely to the compositional layers just mentioned and are known as the asthenosphere and lithosphere, which together correspond only to the upper few hundred miles of the earth's interior (Figure 125-3).

EARTH'S COMPOSITION

The *core* is the densest part of the earth and was created when the heaviest components of the planet "sank" inward during its formation. It consists of two parts, a solid inner core and liquid outer core, both composed of iron, nickel, and sulfur. No one has ever sampled the core; its density and temperature are too great to access, among many other physical limitations. Its composition is determined in two ways: from the makeup of metallic meteorites whose chemistry mimics that of the earth's interior, and from the way earthquake waves propagate through the earth's interior. Even though both parts of the core exist at essentially the same temperature (4000° to 7000° C [7232° to 12,632° F]), the inner core is solid because of the extreme pressure on it. The area near the core-mantle boundary contains pockets of radioactive elements, which undergo fission and create the earth's internal heat. Uneven distribution of these radioactive pockets drives convection and movement of the mantle and ultimately, the earth's living mobility, and gives rise to its magnetic field.

Surrounding the core is the *mantle,* a layer not quite as thick as the core but comprising almost 80% of Earth's volume. It is composed of iron- and magnesium-silicate minerals, which often combine with other elements, such as aluminum, iron, calcium, sodium, magnesium, and potassium. This chemistry makes it only about one-third as dense as the core. No one has sampled the mantle, but in some instances, fragments have been carried to the earth's surface by volcanic eruptions or, rarely, when mantle material is scraped up onto the earth's surface by plate tectonic processes (see next section). Even though it is solid, the mantle "flows" slowly on the currents of heat that are generated at the core-mantle boundary. All solid material, even rock, can flow when subjected to enough heat and pressure.

The outermost, rocky layer of the earth is the *crust.* We live on the crust and interact with aspects of it every day, from copper in our cell phones, to aluminum and steel in our cars, to salt on the dinner table. This relatively thin layer is proportionally equal to that of the skin of a peach or an apple. In fact, a peach may be a perfectly proportional analog to Earth, with its central pit mimicking the core, its juicy flesh representing the mantle, and its thin skin as the crust (Figure 125-4). Two types of crust exist on Earth. Ocean crust underlies most of the ocean basins and is relatively dense (3.3 times as dense as water) but relatively thin, at only 5 to 10 km (3 to 6 miles) thick. These properties cause ocean crust to have a low average elevation of about 5 km (3 miles) below sea level and therefore be covered with seawater. There are a few places where strands of the ocean crust have been forcibly shoved up onto the continents, leaving samples that are easily obtained. Modern technology now allows us to sample the ocean floor. The average composition of ocean crust is very close to that of basalt rock.

Continental crust underlies all the continents and is less dense (2.7 times as dense as water) than ocean crust but on average is much thicker, at about 30 to 80 km (20 to 50 miles) thick. It therefore stands higher on Earth's surface than the ocean crust. The average composition of continental crust is very close to that of granite rock. Because continental crust is less dense than ocean crust, the continents essentially "float" on the mantle, which is why they rise above sea level to form land. The outermost edges of most continents are sometimes inundated with seawater. These fringes of the continents are called *continental shelves* and may extend out to sea up to 225 km (140 miles). The geologic record provides abundant evidence that these shelves have often been more extensive in the past than at present, forming what are known as *epicontinental seas* ("upon the continent"). This partially explains why areas far from the sea today sometimes expose marine rocks on land (Figure 125-5).

Examination of how rocks behave in the upper part of the earth's interior reveals two layers: the *asthenosphere* (meaning "without strength" or "weak layer") and the *lithosphere* ("rocky" or "rigid layer"). The asthenosphere is slightly more than 320 km (200 miles) thick and is found up to 100 km (60 miles) below the lithosphere. It exhibits temperatures close to the melting temperature of rock (about 1600° C [2900° F]) and thus consists of mushy, plastic-like rock containing pockets of molten material. Therefore, the asthenosphere behaves as a ductile (bends without breaking) material and is structurally weak. Its weak, ductile properties impair the velocity of earthquake waves; in fact, this is how the asthenosphere was first recognized. It is sandwiched between stronger, more rigid layers above and below and provides the medium on which Earth's tectonic plates "float." The asthenosphere can be viewed as a yielding cushion upon which the continents drift (Figure 125-6).

The lithosphere has the properties of a rigid, brittle medium. It is formed from the entire crust and very uppermost part of the mantle, which is cool enough at this level to be brittle. The lithosphere is 1 to 60 miles thick (1.6 to 100 km), is relatively strong, and floats on the ductile asthenosphere. The brittle nature of the lithosphere causes it to behave like a broken eggshell, and thus it is broken into a series of tectonic plates (from the Greek *tectos*, "to build"). Heat generated near the core-mantle boundary is what drives slow convection of the mantle, which ultimately allows the brittle lithosphere to move across the face of the earth and fracture into plates. The lithosphere can be viewed as pieces of Styrofoam drifting in a swimming pool, with water currents generated beneath them by activation of the filter system. The

FIGURE 125-4 Various layers inside a peach may serve as a convenient analog to Earth's interior, with the peach pit representing the core, the flesh representing the mantle, and the thin skin of the peach representing Earth's crust. *(Photo by Wayne Ranney.)*

process of plate tectonics drives most landscape-forming processes and is the overriding theory that has shaped much geologic thought in the past 50 years.

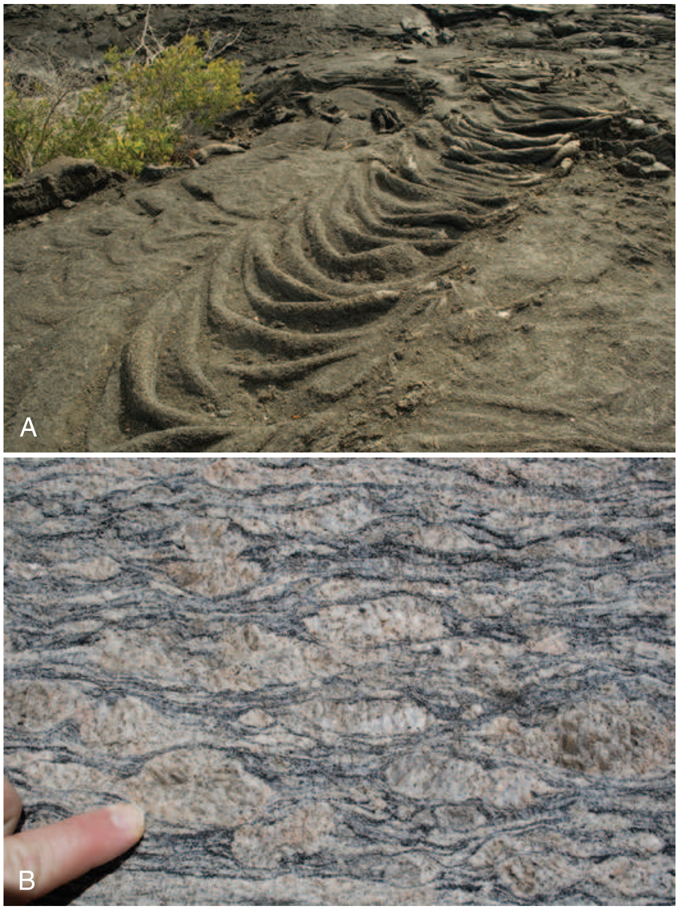


FIGURE 125-5 The two types of earth crust are exemplified by two rock types: **A,** basalt lava flow on the Galápagos Islands, and **B,** granite outcrop in Rio de Janeiro. *(Photos by Wayne Ranney.)*

BOX 125-1 Concept of Theory in Plate Tectonic Theory

A theory is an observation supported by evidence. For a theory to stand up to scientific inquiry, no piece of evidence can falsify the theory. In the case of plate tectonic theory, no scientific evidence of any kind disproves it (although a few remaining diehards see evidence that does not fit the theory). This does not mean that we know everything about plate tectonics, or that plate tectonics can answer every question, but rather that the theory at the fundamental level is not contradicted by scientific evidence.

THE BRITTLE, RESTLESS CRUST: PLATE TECTONICS

The branch of geology that studies movements of the earth's lithosphere is called *plate tectonics.* Based on an enormous amount of evidence, geologists developed this theory in the 1960s. It holds that the lithosphere, the outermost rigid layer of the earth, is divided into separate sections called *plates,* which move in response to convection of the mantle, itself a byproduct of radioactive decay in the earth's interior. Plate tectonic processes build up mountains and down-warp basins. This facilitates erosion of mountains, which delivers sediment to the basins. Most geologic concepts are related to plate tectonics (Figure 125-7 and Box 125-1).

A considerable amount of tectonic activity occurs along the plate boundaries, or margins. Earthquakes and volcanoes are concentrated along these narrow boundaries, which run in an arcuate pattern across the earth's surface, like the stitches of a baseball (Figure 125-8). Three major types of plate boundaries are recognized on Earth: (1) *divergent margins,* where plates separate and new ocean crust is formed; (2) *convergent margins,* where one plate crashes into another and is sometimes consumed or destroyed in subduction; and (3) *transform* or *strike-slip margins,* where plates slide past one another and where crust is generally not created or destroyed.

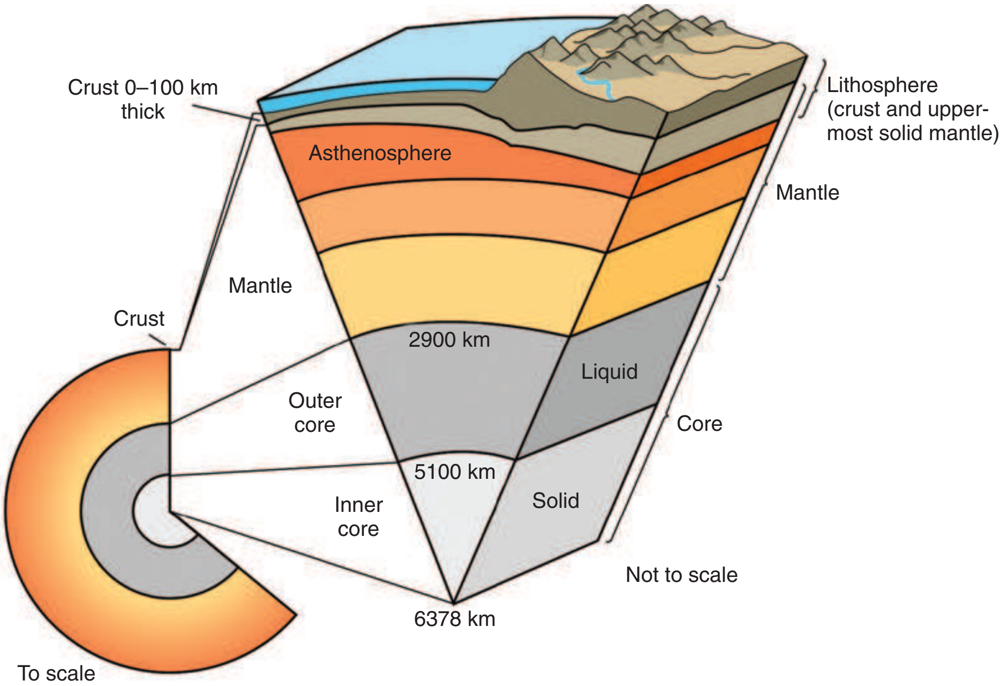


FIGURE 125-6 The lithosphere is composed of the crust and rigid part of the upper mantle. The aesthenosphere lies below it and is the ductile portion of the mantle. *(From* http://upload.wikimedia.org/wikipedia/commons/thumb/8/8a/Earth-cutaway-schematic-english.svg/2000px-Earth-cutaway-schematic-english.svg.png.*)*

DIVERGENT PLATE MARGINS

Also known as *spreading centers,* divergent plate margins are places where plates move away from one another. They are sinuous but lengthy traces that reveal where heat is escaping from the earth's interior. Portions of the upper mantle melt here and force their way into the overlying crust, causing it to separate. As the older crust is pushed away, lava takes its place and solidifies into a rock type called basalt. This process produces long, sinuous chains of submarine mountains called *midocean ridges.* Evidence for this slow, ongoing process comes from the increasing age of ocean crust away from the midocean ridges, alternating bands of magnetism that run parallel to these spreading centers, and mirror images of these phenomena produced on either side of the spreading centers. In fact, the observation of alternating bands of magnetism in ocean crust first led to the concept of seafloor spreading. Basalt contains appreciable amounts of iron, and when it is extruded on the ocean floor, preserves a record of the earth's changing magnetic field. Because the earth's magnetic poles occasionally switch from north to south, and vice versa, an alternating magnetic signature is progressively recorded in ocean crust at spreading centers. This pattern of alternating magnetism is identical on either side of the ridge, showing how the crust has spread apart. Midocean ridges occur in all the world's major oceans (Figure 125-9).

CONVERGENT PLATE MARGINS

When new crust is formed at midocean ridges, something must be destroyed on the opposing edge of the plate (unless the earth is growing larger, an idea for which there is no evidence). Zones where crustal plates move together and are often destroyed are called *convergent plate margins.* They occur variably between two oceanic plates (as in the Japan, Philippines, or Aleutian island chains), two continental plates (as in the Himalayan chain, where the Indian plate is colliding with the Eurasian plate), or between an ocean and a continent, such as the western edge of South America. The South American example provides a typical series of events. As lithospheric plates converge, the oceanic crust that is being pushed from behind at the spreading center (and tending to be denser and cooler) sinks beneath the lighter, continental crust in a process called *subduction.* As the ocean plate is subducted, usually at an angle of 35 to 60 degrees, it is pushed into the earth, where portions of it melt after achieving a depth of about 100 km (60 miles), although parts of the descending slab can reach depths of 640 km (400 miles) before melting entirely. Pockets of molten material then rise buoyantly through the continental crust, creating slightly curved chains of volcanoes called *volcanic arcs* (Figure 125-10). These arcs typically erupt through overthickened crust, formed by the convergence of two plates. Oceanic trenches form offshore, where the ocean crust slides beneath the continental edge and can be quite deep. Plate boundaries where two continents collide are also the site of convergence but generally do not produce volcanism. This is because when two pieces of continental crust meet, they are equally buoyant to preclude subduction. Thus, no melting occurs in continent-to-continent convergence. Great mountain ranges such as the Himalayas are thrust upward in these settings, resulting in continental rocks becoming greatly deformed and folded.

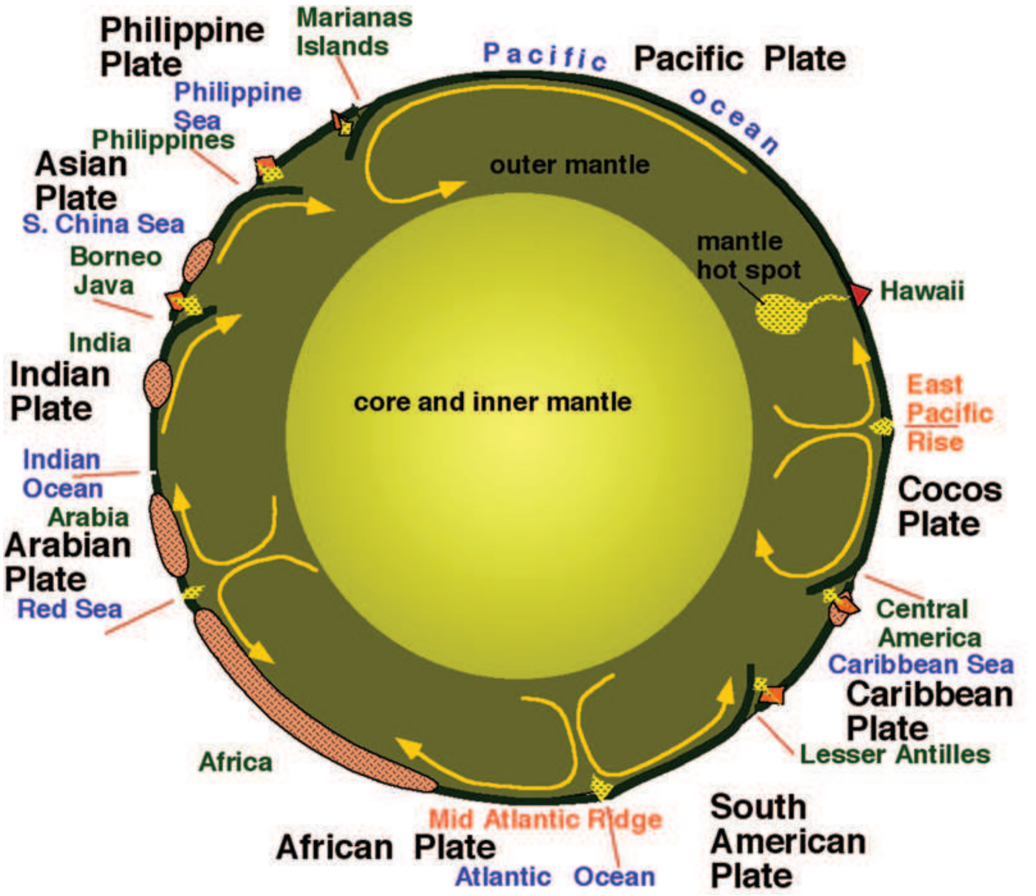


FIGURE 125-7 Cross section through Earth at about 10 degrees north latitude showing the components of plate tectonic theory. The *yellow arrows* show convection cells in the mantle that drive the surface motion of the earth's plates. Continental crust is depicted in *pinkish color* with ocean crust as a *solid dark line.* The plate boundaries are shown with *thin red lines* pointing to the crust. Major tectonic features are named and labeled as plates in *black,* oceans in *purple,* arcs and continental crust in *dark green,* and midocean ridges in *orange.* All thicknesses depicted are exaggerated and not to scale. *(Courtesy Ron Blakey.)*

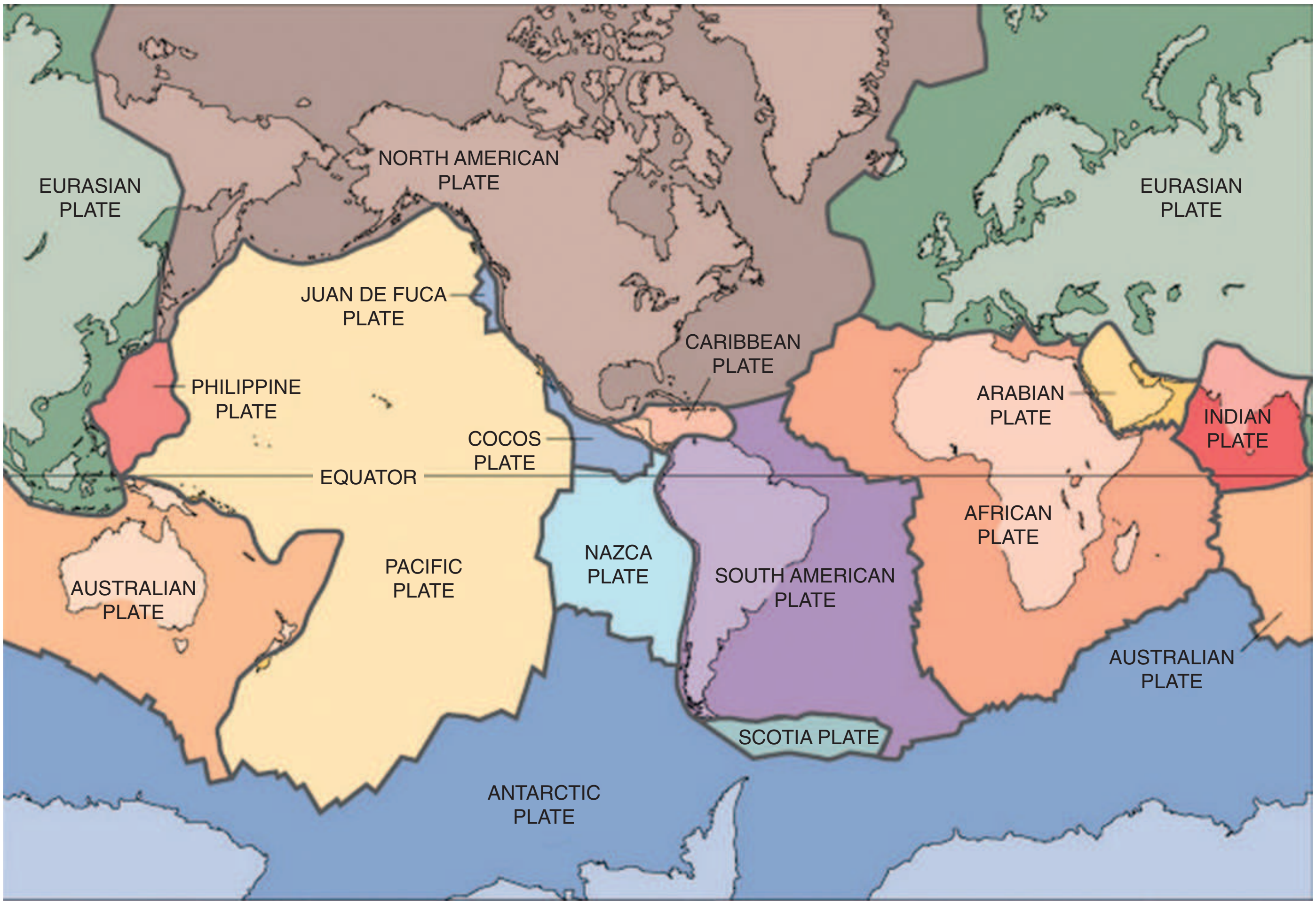


FIGURE 125-8 Map of the earth showing its major plates. *Lighter* shade of any color shows the continental portion of the plate, while *darker* color shows the oceanic component of the same plate. Fifteen of the largest plates are labeled here, with dozens of smaller plates not shown. *(From the United States Geological Survey.* http://pubs.usgs.gov/gip/dynamic/slabs.html. *Accessed February 9, 2016.)*

STRIKE-SLIP OR TRANSFORM PLATE MARGINS

Some plates slide horizontally past one another in such a way that little or no creation or destruction of crust occurs. Such boundaries are called strike-slip or transform margins because the plate motions transform the offset laterally. Few volcanoes occur in these tectonic settings because there is no subduction and consequent melting of the crust. Earthquakes, however, are quite abundant and tend to be rather destructive as the two rigid plates grind past one another. The San Andreas Fault in California and the Alpine Fault in New Zealand are the most famous examples of such plate boundaries. Earthquakes, therefore, are the result of breaking of brittle lithosphere. When this solid rock is subjected to stress, it sends off shock waves that generate earthquakes. Most earthquakes occur along the earth's plate boundaries, but it is also possible to have intraplate earthquakes, such as the famous New Madrid, Missouri, quakes of 1811-1812 or the Virginia earthquake of August 23, 2011. Plotting major earthquake epicenters over time clearly delineates the location of the plate boundaries. Shallow earthquakes tend to occur along all plate boundaries, whereas intermediate and deep earthquakes occur only in subduction zones along convergent margins. Shallow earthquakes are defined as those occurring down to 60 km (40 miles); intermediate, between 60 and 320 km (40 and 200 miles); and deep, 320 to 640 km (200 to 400 miles). Below this depth, the mantle is ductile and rocks do not break.

Mountains, the most obvious result of the uplift generated by Earth's tectonic system, are formed in many ways. Almost all mountain ranges are located near present or ancient plate boundaries. The Appalachian chain formed near an ancient plate boundary, and the Andes formed on a modern margin. Mountains at convergent margins are the result of plates pushing rock together, which thickens the crust either by squeezing it or by forcing some sections to override other sections (*thrusting*). Volcanoes can form mountains as molten material accumulates on the surface and cools into rock. Mountains are also created at divergent plate boundaries because upwelling heat initially swells the earth's crust before it rifts apart; the flow of hot material expands the crust upward, and piles of volcanic material are added to its top.

A less obvious but no less important result of tectonism is formation of tectonic or sedimentary basins. These form in conjunction with mountain uplift but are often overlooked by nongeologists because they do not form the rugged, spectacular scenery of a mountain range. Sedimentary basins, however, are corollaries of the same tectonic story. As plates converge, they unevenly warp the crust into both mountains and basins. Consider the way a throw rug is variably warped as it is pushed along a hardwood floor into a wall. You can see both humps (mountains) and depressions (basins) that form on the rug's surface. In tectonic settings, rocks are eroded off the mountains and transported by rivers and wind into the basins. These basins preserve most of the earth's sedimentary deposits, which reveal details about Earth's prior environments, climate, and life. It is the workings of plate tectonics that allows this clear view into the ancient past.

TECTONIC ORIGIN OF ROCKS

Pervasiveness of plate tectonics on Earth creates the conditions necessary to form the many different types of rocks found on our planet. These rocks constitute the record of Earth's history.

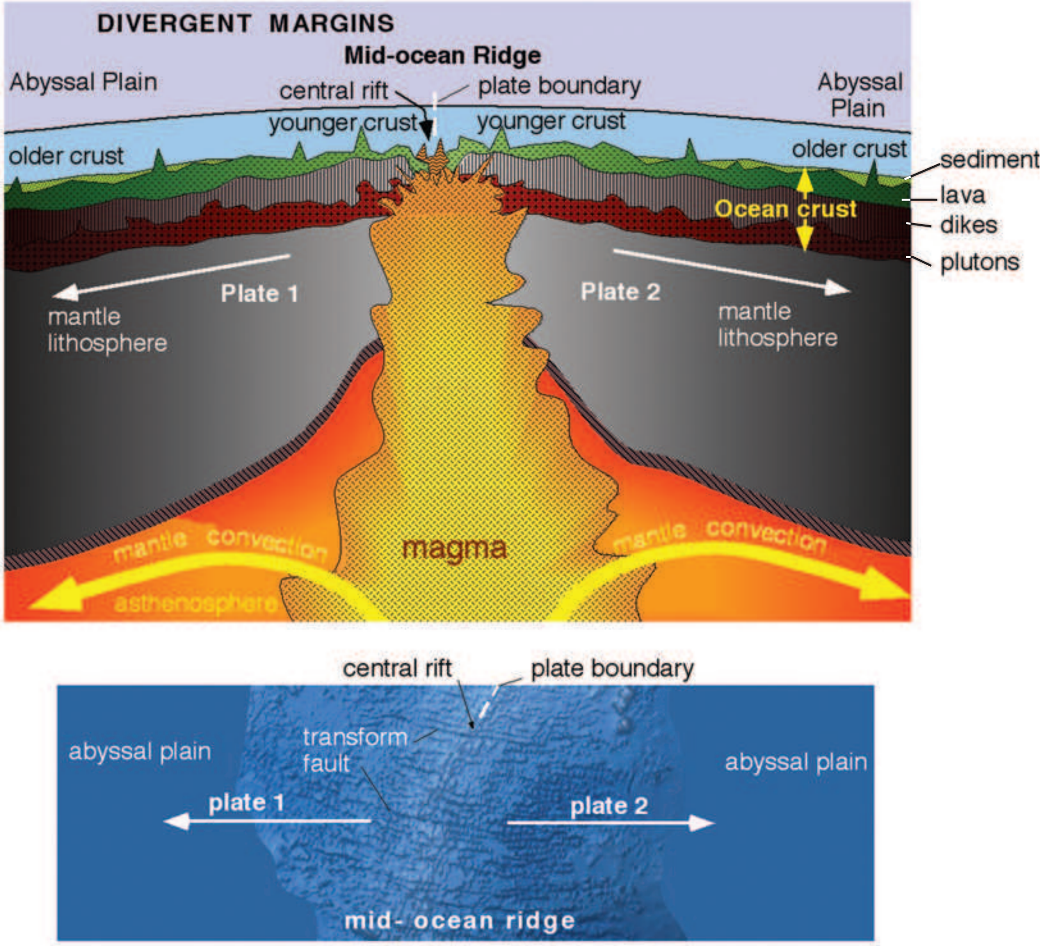


FIGURE 125-9 Cross section through a typical midocean ridge or divergent plate margin. Note how magma rises through the lithosphere and is erupted onto the ocean floor, creating new ocean crust. *(Courtesy Ron Blakey.)*

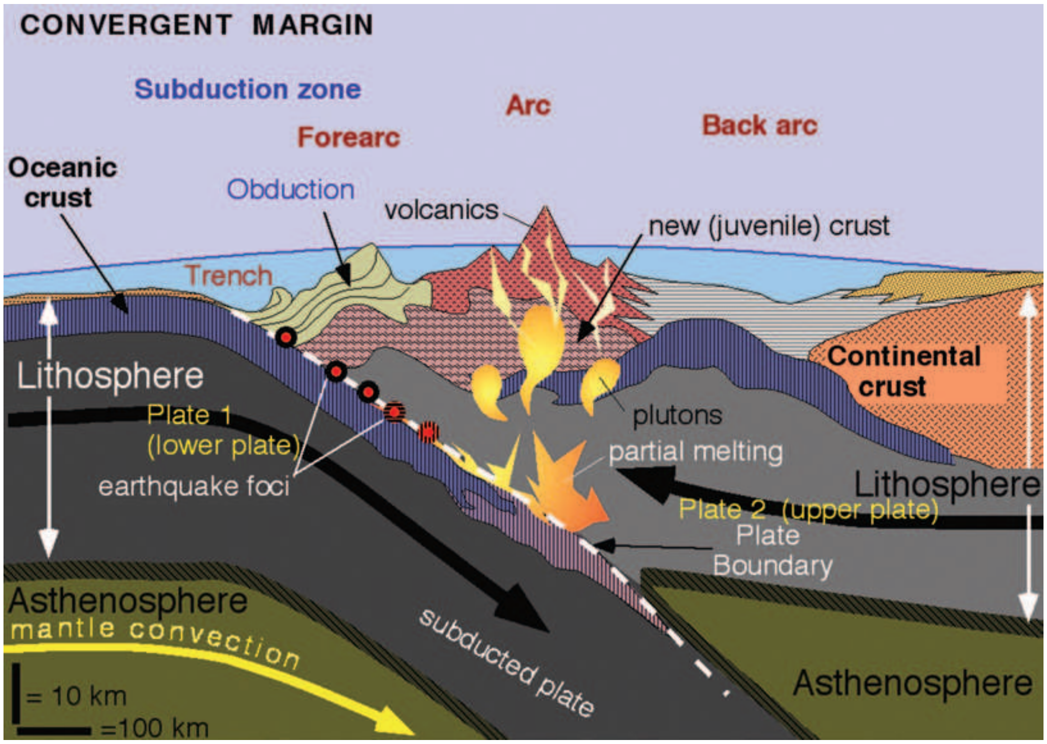


FIGURE 125-10 Convergent margins form where oceanic crust is subducted beneath the edge of a continent (or between two continents or the edge of another ocean plate). Although some material may be scraped off near the top (*obduction*), most descends beneath the continent because ocean crust is cooler and denser than continental crust. Melting of the subducted plate at depth of about 100 km (60 miles) creates magma that erupts into the crumpled edge of the continent to create a volcanic arc. The Andes Mountains in South America are a good example of this type of plate margin. *(Courtesy Ron Blakey.)*

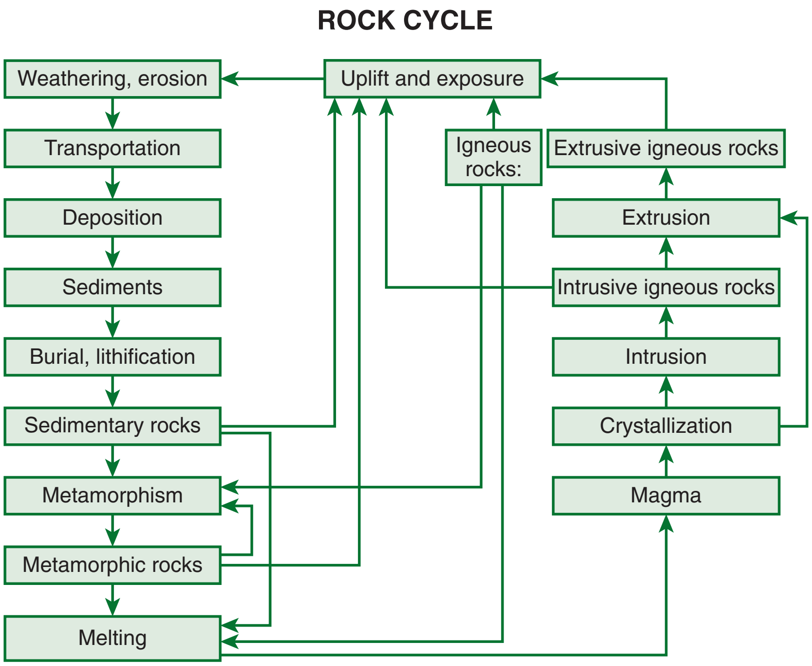


FIGURE 125-11 Typical rock cycle showing the relationship among the three classes of rocks. Follow the "Uplift and exposure" box (top, center) counterclockwise for the simplest cycle, noting only that rocks can take "shortcuts" at any stage within the circle.

Most people are familiar with the three main types of rocks: igneous, metamorphic, and sedimentary, but many may not be aware of the close relationship these rocks have with tectonic processes (Figure 125-11). Igneous and metamorphic rocks, which contain groupings of one or more mineral crystals, are the so-called crystalline rocks. They form from heat and except for volcanic rocks, within the earth. Sedimentary rocks contain a record of Earth's surface and form from broken bits of other rocks (including other sedimentary rocks), precipitation of solutions, or compaction of shells.

Igneous Rocks

Igneous comes from the Latin word meaning "to ignite" ("burning" or "fiery"). Igneous rocks are born of heat. When the earth's crust or upper mantle is subjected to extreme heat, melting occurs and molten magma is formed. Therefore, all igneous rocks originate initially as magma. Some of this magma may remain deep within the crust, where it cools slowly to form solid rock (e.g., granite). Such rocks take a long time to form, which allows for growth of large crystals as the magma cools and solidifies. Such igneous rocks are called *intrusive* or *plutonic* (after Pluto, the Roman god of the underworld). They contain coarse-grained crystals that are easily seen without magnification. Large bodies of plutonic rock are called *plutons* or *batholiths*; small bodies less than 259 km^2 (100 miles^2) are called *stocks* (Figure 125-12).

Some magma makes its way to the surface in liquid form and flows out rather gently, or it is blasted out violently onto Earth's surface. This igneous material cools relatively quickly, and crystals often do not have time to form. If crystals form, they are very fine-grained and are not usually visible without magnification. Basalt lava flows or rhyolite ash beds are examples of *extrusive* or *volcanic* rocks (after Vulcan, the Roman god of the forge).

Volcanoes are simply mountains composed of piles of volcanic material—cinders, lava, and ash—erupted from a central vent. Volcanoes are classified according to their shape or the processes that form them. *Shield volcanoes* form from piles of extremely fluid basalt lava that flows and cools far from the vent; they tend to be broad at the base with a very low profile, reflecting the low viscosity of the lava. *Cinder cones* (or *scoria cones*) are formed from small droplets of molten rock that erupt into the air and cool as pea-sized particles in a steep pile around the vent. Cinder cones tend to be the smallest volcanoes, at about 300 m (1000 feet) in height. *Composite volcanoes* or *stratovolcanoes* are formed from alternating layers of andesite lava flows and ash. These form some of the largest, most symmetric, and dangerous mountain peaks on Earth: Mt St Helens, Mt Fuji, and Mt Pinatubo are examples. These types of volcanoes yield huge amounts of ash. When this material travels to distant basins, it can become preserved within sedimentary layers. These ash beds are readily dated and allow geologists to date strata, which otherwise could

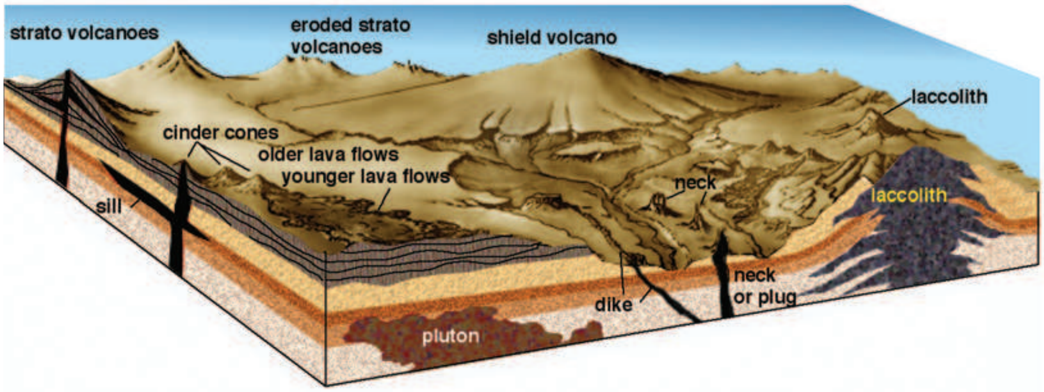


FIGURE 125-12 Hypothetical landscape, showing many of the igneous landforms that can be found on Earth. *(Courtesy Ron Blakey.)*

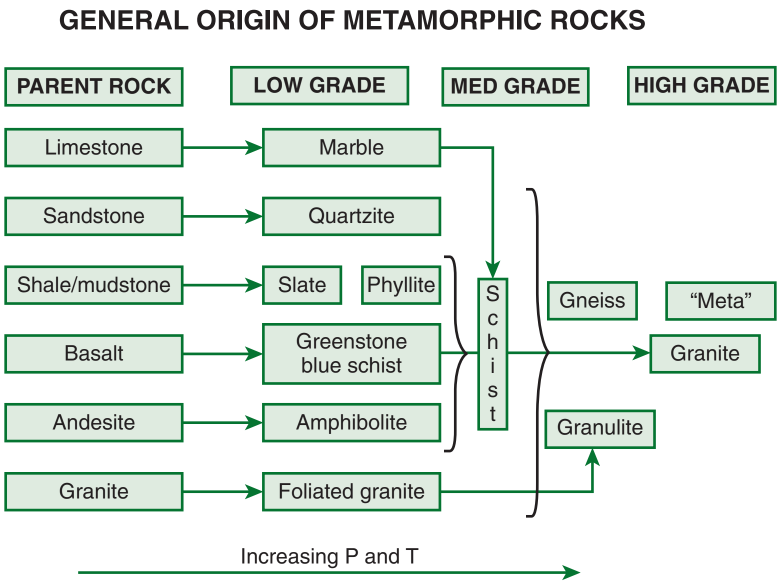


FIGURE 125-13 Protolith or parent rocks and their metamorphic equivalents, with increasing grade of metamorphism shown *left* to *right*.

not be dated directly. *Dome volcanoes,* the last type of volcanic edifice, form from very viscous magma that resists flow away from the vent.

Volcanoes have the potential to cause much destruction. Lava flows pose an obvious hazard, but more serious are volcanic explosions and their associated gases, hot dust (ash), and mudflows triggered by melting ice and snow. The more destruction and death associated with volcanic eruptions, the more they are reported. However, many significant eruptions in unpopulated areas go virtually unreported and unnoticed.

Metamorphic Rocks

Rocks formed by heat and pressure, regardless of their original makeup, are called metamorphic rocks. The word *metamorphic* means "changed form." These rocks have been altered considerably since their original incarnation. Heat and pressure applied by tectonic processes (with perhaps associated magmatism) can cause a significant change in the appearance of a rock, such that an entirely new rock is formed. These changes include formation of entirely new minerals, recrystallization of old minerals (usually with an increase in crystal size), or alignment, banding, or segregation of differently colored minerals. Some common metamorphic rocks are slate, marble, quartzite, schist, and gneiss. Each of these has an original rock type, or *protolith,* from which it formed. The protoliths can often be determined by the overall chemistry and texture of a metamorphic rock and may help in determining the ancient setting of the prealtered rocks (Figure 125-13).

There is a relative increase in metamorphic grade as rocks become subjected to increasing temperatures and pressures. *Low-grade* metamorphic rocks begin to form between 200° and 300° C (392° and 572° F). Examples of these are phyllite and slate (both different grades of altered shale) and quartzite (altered sandstone). *Medium-grade* metamorphism begins at 300° C (572° F) and ends at about 500° C (932° F). These rocks include schist (increasingly altered shale), marble (from limestone), and amphibolite (from basalt). *High-grade* metamorphism begins at about 500° C [932° F]) and includes gneiss (extremely altered shale or granite) and migmatite (tectonically deformed and cooked granite). At about 950° C (1742° F), rocks begin to melt, and the igneous environment begins.

Through time, metamorphic rocks find their way back to the surface in tectonic uplifts, where erosion exposes them to view. Awareness of all these rock-forming processes is useful in deciphering the seemingly unconnected parts of Earth's history. The metamorphic grade in a rock allows us to know the specific conditions that existed during its creation. Certain minerals can grow only within a narrow range of temperatures and pressures, and geologists use the known depths where those conditions are present today to infer the ancient depth of the rocks when they were created. The chemistry of the rock can suggest a protolith, which can divulge the specific sedimentary environment of the rock before its burial and metamorphism. This may provide a clue to its tectonic setting just before the mountain-building event that changed it. Taken together, all this information might reveal a sensible sequence of tectonic events in which a rock originated in a sedimentary basin, only to become involved in a convergent mountain-building event that folded, buried, and altered the rock deep within the earth.

The beauty of the plate tectonic concept and of modern geologic thought is that discreet bits of evidence scattered across the globe allow us to know the sequential evolution of our planet, if only we can recognize the evidence. Before the birth of this concept, metamorphic rocks could only be described in stone-faced prose and dry classification schemes that were simply organized around their texture. Attempts were made to interpret how the rocks might have originated, but there was no single, fundamental concept that could show a link between the seemingly unrelated aspects of a rock's metamorphic grade, its sedimentary protolith, or how it was uplifted and exposed. A plate tectonic view has allowed us to see how the transitions in Earth's history have been actualized. Many of the details are still being determined, but scientists continue to unravel Earth's history at an astounding rate.

DYNAMICS OF SEDIMENTATION AND SEDIMENTARY ROCKS

Sedimentary rocks are the most important surface rocks in reconstructing the details of past events. If sedimentary rocks are found in a particular area, they preserve something of the surface history that occurred there. We usually can recognize parts of that past, because all modern environments also existed in the past. Geologists are like detectives who arrive at a crime scene long after the fact. In much the same way that 20-year-old fingerprints can pinpoint the person who left the scene of a crime, ordinary sandstone, limestone, and mudstone record the specific environments that once covered an area, even if that area is now completely changed. Sediments accumulate in layers called *strata* to form sedimentary rocks. Because strata form under surface conditions, they reflect the extent and nature of the environments in which they formed. This allows geologists to reconstruct the ancient landscapes that once existed on Earth. *Clasts* (broken grains or particles) in sedimentary rocks tell us about the parent rocks from which they were derived. Clasts can also tell us about the conditions of weathering, erosion, and the transportation history of the grains. Fossils in sedimentary rocks reflect specific ecologic and

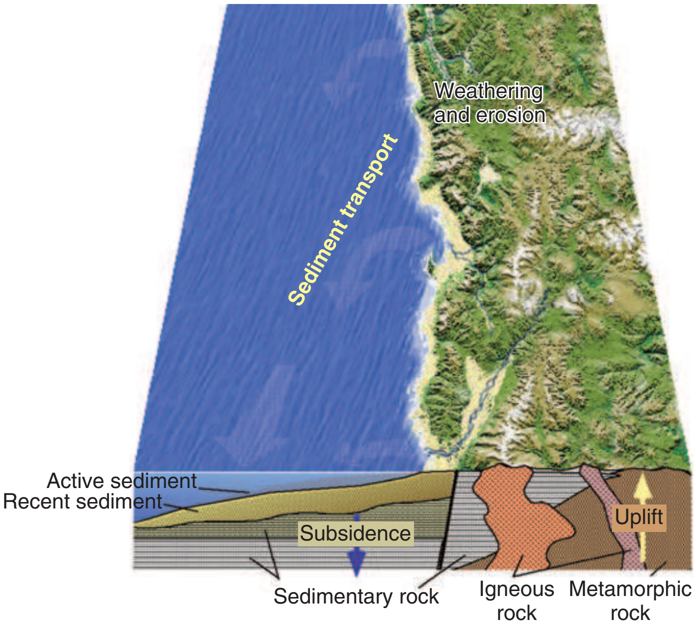


FIGURE 125-14 Sedimentary cycle, in which rocks of any type are uplifted, weathered, eroded, and transported to a depositional basin. Subsidence in the basin creates the space where the sediment becomes buried and lithified. *(Courtesy Ron Blakey.)*

environmental conditions at the site of deposition. *Evaporite rocks* such as halite and gypsum tell us about the chemistry of waters, as well as the climate during deposition. Organic material also reflects the extent of the fecundity of the biosphere.

Ironically, sedimentation begins with weathering and erosion. Any type of rock exposed on Earth's surface eventually breaks down as it is subjected to vagaries of the atmosphere or hydrosphere. Eventually, these clasts are transported downwind or downstream and come to rest in a basin. Formation of sedimentary rocks involves initial weathering (physical or chemical) and erosion of preexisting rocks, transportation of the broken pieces to other areas as sediment, and its ultimate deposition as a layer of new sediment. These steps can be presented graphically in a *sedimentary cycle*. Factors such as the tectonic setting, climate, rock types in the source area, and distance to the sea affect the sedimentary cycle (Figure 125-14).

Weathering is the breakdown and change of rocks at or near the earth's surface. Many rocks are formed under heat and pressure within the earth, where there are relatively low amounts of water. Conversely, the surface of the earth has lower temperatures and pressures but large amounts of water. These opposed conditions tend to cause minerals to change either by physical breakdown of the rock or by chemical alteration. Physical weathering is decrease in the size of clasts from larger pieces to smaller ones. Freeze-and-thaw processes in rock cavities split off chunks of rock and are a good example of physical weathering. Chemical alteration in composition of the parent material is an illustration of chemical weathering. For example, feldspar, a common mineral in granite, interacts chemically with rain and groundwater to form clay minerals. This process results in a new substance with totally different properties than the parent mineral. Chemical weathering results in release of dissolved ions, such as calcium, iron, and silica, into the surface water and/or groundwater. These ions are carried by surface water to the sea and become the source of dissolved sea salts. They are carried by groundwater through rocks and often become the cementing agents for loose sediment (e.g., turning sand into sandstone).

Erosion is the process that removes weathered bits of rock and carries them to a new site. It is accomplished by moving water, blowing wind, soil creep, or flowing glacial ice. Erosion by running water forms the canyons and gullies that are so prevalent on our planet. The force of running water removes loose particles and is the method by which canyons are lengthened, deepened, and widened. Each new flood carries away more material to increase the size of the canyon. Meanwhile, weathering loosens more material, and the process continues. Erosion by running water is by far the most important aspect of this process. Wind erosion is much less important with respect to the amount of material moved but can affect weakly cemented rocks when they are buffeted by other particles traveling with the wind. The most efficient erosion is that accomplished by glacial ice. Glaciers form when more snow is deposited in winter than melts in spring and summer. This must occur for a number of successive years for major glaciers or ice fields to form. As ice accumulates and thickens, it flows downhill and has a tremendous capacity to erode soil and even solid rock. Ice is a major factor in sculpting some landscapes, especially in mountainous areas.

Numerous factors affect the rate and type of erosion that occurs across the landscape. Climate determines the distribution and amount of precipitation and thus the pace of erosion. Arid landscapes tend to be more angular in appearance than humid landscapes because erosion is concentrated along river courses, even if those channels are usually dry. Intervening areas (mesas and plateaus) between the major rivers have much slower rates of erosion in this dry environment and thus stand tall relative to the deep canyons. Humid areas, such as Brazil or the eastern seaboard of North America, experience considerable chemical weathering, and the intervening areas between rivers tend to "round off" at about the same rate as the rivers dissect. Elevation and *relief* (the elevation difference between the highest and lowest points in a given area) are other important factors and determine the potential energy of running water and how far the water drops to attain *base level*, defined as the final destination of a stream, usually a lake or the ocean, which is the ultimate base level. Steep-gradient rivers tend to both deepen and lengthen their channels. If a river has a steep gradient to base level, erosion will be facilitated.

Bedrock type can determine the ultimate shapes we see on landscapes. Harder rocks, such as granite or limestone, tend to produce steep-walled canyons, whereas softer rocks such as shale or mudstone typically yield broader valleys. The geologic age of a landscape (how long the area has been subjected to the current conditions of weathering and erosion) determines the overall appearance of a landscape. Relatively young landscapes tend to be more rugged and angular in appearance and to have the greatest relief. Older landscapes tend to have more rounded slopes and hills and to be mostly low-lying with broad, open river valleys.

Many areas undergo erosion as *mass movement*. Mass movement involves pulses or relatively short spurts of activity where large amounts of material move downslope. The type of mass movement is related to the type of material being moved by erosion and amount of water associated with the process. Mudflows involve large amounts of water and relatively fine material, such as soil and mud. Rockfalls involve little water and large amounts of loose rock. *Slumps* are sudden slippage, usually of water-saturated soil or rock, and *creep* involves movement of water-saturated soil and rock by means of numerous small pulses of short duration. Mass movement causes billions of dollars of damage worldwide on an annual basis. Roads, buildings, and other cultural features can be destroyed rapidly, often without warning.

The products of erosion are usually transported by water, wind, or ice and then deposited (Figure 125-15). These deposits include gravel, sand, and mud (technically a mixture of clay and silt particles), precipitates (material that changes from solution to solid in bodies of water), and organic deposits (material formed from living organisms). Common precipitates include halite (common salt), gypsum, potash, and some limestones. Common organic deposits include most limestones, coals, and phosphates.

Sedimentary rocks can be dated by several different methods, including *fossils,* the remains of organisms that lived during deposition. Because life has changed constantly throughout geologic time, fossils can be used to date sedimentary rocks in a relative way. They do not tell us "how many millions of years ago," but rather "what came first" and "what came later." Fossils that are widespread laterally but that lived for only a short time make excellent *index fossils*, used to correlate widely separated rock units. Presence of datable grains and volcanic ash beds allows absolute dating that can tell us "how many millions of years ago." Dating of clasts, such as zircon, must be used with

SEDIMENT AND SEDIMENTARY ROCK

Sediment	Common sedimentary environments	Rock name
Gravel: rock frags, quartz (grains >2 mm dia)	River channels, alluvial fans, some beaches, glaciers	Conglomerate
Sand: mostly quartz, some feldspar (grains 1/16–2 mm dia)	River channels, eolian dunes, beaches, shorelines	Sandstone
Mud: any mix of silt and clay (grains <1/16 mm dia)	River floodplains, lakes, swamps, low-energy marine and coastal plain	Mudstone (Shale)
Lime: $CaCO_3$	Clear, warm, shallow marine shelves, some lakes	Limestone, Dolomite (with Mg in crystal structure)
Evaporite: salts	Restricted marine, sabkha, arid lakes	Halite, Gypsum, some Limestone, Potash
Peat	Humid coastal plains and swamps	Coal

FIGURE 125-15 Sediment types with their equivalent rock names and the environments in which they form.

care, because the grains provide only the age of the parent rock, not the age of the sedimentary deposit. Using this method, geologists may only be able to tell a strata's maximum age. Volcanic ash beds are an important dating method for sedimentary strata. Ash deposits originate from eruptions far away and are carried by the wind to become trapped in the deposits of a sedimentary basin. Their presence in strata can be cryptic and difficult for an untrained eye to detect, but once found, provide useful dating tools and yield increasingly reliable dates. Taken together, use of fossils and datable materials has allowed global correlation of strata.

PRESENT-DAY GEOLOGY AS A KEY TO UNDERSTANDING THE PAST

The concept of using the present to understand the past is an invaluable tool to the geologist. This concept is called *uniformitarianism* and widely applied in the field of historical geology. For example, one can study the processes and subsequent deposits associated with the flooding of a major modern-day river. As floodwaters recede, they leave behind newly deposited sediment, and geologists may trench through the sand, mud, and debris to observe what is preserved. They will look to see how the grains of sand are arranged, what sizes they are, and how they all "stack up" when viewed comprehensively. Geologists are interested in a sediment's *texture* and the *sedimentary structures* it contains. The texture and structure of these modern flood deposits (which have first-person accounts that verify how they originated) can be used to help recognize ancient flood deposits whose origins were not directly observed. In this way, flood events that happened many millions of years ago can be documented, showing how a present reality becomes a key to understanding the past. Because the same rigorous observations are performed on other types of deposits, such as sand dunes in the Sahara, deltas in India, and beaches in eastern North America, a body of knowledge exists that helps to interpret and differentiate the various types of deposits, regardless of their antiquity.

This ability to "read the rocks" has only evolved in the last 250 years. Before that, people saw oddly located fossils and lacked the geologic framework to make sense of how seashells came to rest in the mountains. Before development of geologic thought, they could not know that rocks contained a sequential record of Earth history. Modern geologic concepts began with the Age of Enlightenment at the end of 18th century and have progressed such that it is now possible to reconstruct major portions of Earth's past, even though there were no humans present to witness it. Several assumptions need to be accepted, however, most importantly that the processes acting on the earth today are similar to those that have acted in the past. The evidence, critically reviewed by thousands of scientists, is overwhelming that the earth behaves today much as it has in the past.

HOW ROCKS ARE DATED

Many nonscientists express skepticism when presented with ideas regarding the staggering antiquity of Earth. They wonder, "How can it be known?" that events happened hundreds of millions or billions of years ago. Certain physical properties of some common elements allow scientists to discover how long ago rocks formed or events occurred. It involves radioactive decay from a parent product into a daughter product. These naturally occurring decay series happen at a known and constant rate. One example is radiocarbon decay. Organisms such as animals and plants absorb carbon-14 (^{14}C) from the atmosphere. When their living functions cease, they stop absorbing ^{14}C, which is radioactive and begins to decay to ^{12}C. The half-life of ^{14}C is 5730 years (±40 years), meaning that after that amount of time, half the ^{14}C has decayed to ^{12}C. This radiocarbon method is good for dating organic matter less than about 62,000 years old. After that, there is insufficient residual ^{14}C to obtain a ratio.

Therefore, other decay series must be used to date older materials. Some common decay series used to date rocks are uranium-235 to lead-207 (half-life, 700 million years); uranium-238 to lead-206 (half-life, 4.5 billion years); potassium-40 to argon-40 (half-life, 1.3 billion years); and rubidium-87 to strontium-87 (half-life, 50 billion years).

GEOLOGIC TIME

The concept of geologic time, or *deep time*, exposes humans, who are used to much smaller time frames, to vast numbers of years. One million years is difficult to comprehend even to the geologist, yet represents only a fraction of Earth's history. An analogy may help. Medium-sized sand grains, the common building block of many sedimentary rocks and widespread in modern sand dunes and beaches, are approximately 10 mm (1⁄25 inch) in diameter. Although 25 sand grains laid side by side in a line is only 2.5 cm (1 inch), 1 billion sand grains would stretch over more than 1015 km (630 miles). Astonishingly slow geologic rates, such as the movement of the earth's plates, can accomplish amazing feats in a relatively short geologic time span. For example, with an average rate of plate motion at about 2.5 cm (1 inch) per year, the Atlantic Ocean has widened another 14 m (44 feet) since Christopher Columbus sailed from Spain to the New World in 1492. In just 100 million years, widening of the Atlantic has proceeded more than 1900 km (1200 miles) (Figure 125-16).

THE GRAND CANYON: AN EXAMPLE OF EARTH SCIENCE AT WORK

INTRODUCTION AND PHYSICAL SETTING

The Grand Canyon is one of Earth's most iconic landscapes (Figure 125-17). It provides an exceptional window into the workings of planet Earth and serves to highlight the basic concepts presented in this chapter. Grand Canyon National Park offers a host of colorful viewpoints from the rim that present visitors a platform from which to view a spectacular display of Earth history. Numerous trails leave the rim to provide access to the Colorado River, offering an exciting white-water ride through the length of the canyon. Almost everyone who visits the canyon is immediately impressed with its immense size, rugged and colorful topography, and stunning skies and changeable weather.

The Colorado River and its tributaries have likely carved the Grand Canyon in only the last 5 to 6 Ma (*mega-annum,* or million years ago), but neither the exact age nor the specific processes that acted to create it are resolved. The river flows through the canyon for 450 km (277 miles), but nowhere can the canyon be viewed in its entirety from the ground. On average, the canyon measures about 16 km (10 miles) wide, with an extreme width of 29 km (18 miles). It is more than 1.6 km (1 mile) deep in most places, with about 4170 km^3 (1000 miles^3) of rock having been removed by erosion. Much of this material now resides in the area of the Gulf of California, where the Colorado River ends its 2333-km (1450-mile) journey to the sea.

The canyon is located entirely within the state of Arizona and on the southwestern edge of the Colorado Plateau, one of 26 geographic provinces described within the boundaries of the United States. Because of its extreme relief and longitudinal profile, the canyon is home to 1750 species of plants (more than any other National Park in the country), 373 species of birds, 47 reptile species, and 34 species of mammals. Its archaeological record extends back at least 4500 years and is based on radiocarbon-dated willow-stick figurines found in caves. The record may extend back 12,000 years or more, to the time when people first arrived in the Americas, but based only on a single projectile point, found on the canyon's South Rim.

Persons of European descent first saw the canyon in 1540, when native guides led members of the Coronado Expedition to the canyon's edge. These explorers were not impressed, and the canyon was not truly appreciated by humans until the first geologist visited in 1858. From that time onward, people have come to the Grand Canyon to experience its sublime grandeur and spectacular vistas. Today, it is visited by almost 5.5 million people a year, with more than 40% coming from outside the United States.

CREATING THE ROCKS: 2 BILLION YEARS OF EARTH HISTORY

Basement Rocks

The geologic story at Grand Canyon National Park begins about 1840 Ma, when other *terranes* collided with North America and became attached to it (a *terrane* is a discreet portion of the crust containing related rocks with a shared geologic history) (Figure 125-18). This collision compressed the rocks and both raised mountains on the surface and folded them to great depths within the crust, altering them to medium- to high-grade metamorphic rocks. Garnet minerals within rocks exposed today reveal burial depths of up to about 25 km (15 miles) with temperatures of 750° C (1382° F). They are now formally described as the Grand Canyon Metamorphic Suite, but historically are known as Vishnu Schist.

At even greater depths, the same rocks were melted and then rose buoyantly into the still-deforming metamorphic assemblage. These rocks were intruded as light-colored granitic dikes and plutons between about 1710 and 1660 Ma. They have been formally classified as the Zoroaster Plutonic Complex, but are historically known as the Zoroaster Granite. The resulting assemblage of igneous and metamorphic rocks records the dynamic changes that added crust to the North American continent over a 180-million-year period.

The entire assemblage of metamorphic and igneous rocks is informally referred to as the "Vishnu basement" (Figure 125-19). Mica minerals within the schist record how the overlying 21 km (13 miles) of rock was eroded away between 1350 and 1254 Ma. As erosion continued through time, the confining pressures were gradually lessened, and the rocks below rose isostatically. This is how rocks once found at 21 km (13 miles) down were brought back to the earth's surface. A generally flat-lying erosion surface was ultimately worn down to near sea level.

Grand Canyon Supergroup

Following this period of erosion, during which the Vishnu basement rocks were planed to near sea level, sediments began to be deposited on top of them around 1250 Ma. These rocks belong to the Grand Canyon Supergroup, a package of mostly sedimentary rocks containing nine formations that is more than 3800 m (12,500 feet) thick. The lower part of the Supergroup records deposition in offshore (limestone), nearshore (shale), and continental (sandstone) environments. A volcanic period with eruptive lava flows and forceful intrusions ended deposition about 1100 Ma.

The upper part of the Supergroup was preceded by an interval of erosion lasting up to 400 million years, and the rocks in this package were laid down in shallow marine and nearshore settings. Few rocks of this age (between 780 and 740 Ma) are known on Earth. The Grand Canyon Supergroup contains an important record of the diversification of single-celled life and the appearance of heterotrophic life—organisms that gain their nutrition from other organisms rather than exclusively through photosynthesis.

The Grand Canyon Supergroup is found in only about 10% of the canyon and always as isolated, fault-bounded, and tilted blocks. The preserved blocks represent areas that were downfaulted low into the ancient crust, allowing them to escape subsequent erosion. Blocks that were faulted higher were eroded away completely between 650 and 540 Ma. Some lithologies in the Supergroup were particularly resistant to erosion and stood as cliffs on the ancient landscape. These cliffs were preserved beneath the next package of rocks (Figure 125-20).

Paleozoic Rocks

A 1200-m (4000-foot) section of flat-lying Paleozoic strata makes up the upper four-fifths of the walls of Grand Canyon, composing the easily recognizable, stratified profile of the canyon (Figure 125-21). These rocks record deposition over a 255-million-year period spanning the length of the Paleozoic Era. There are 14 separate formations that make up this stack of rocks, each with a distinctive story to tell about the environments that once existed there (Figure 125-22).

Cambrian-age rocks include the Tapeats Sandstone, Bright Angel Shale, and Muav Limestone. This three-part assemblage exposes a continental-to-marine sequence that records the gradual onlap (about 30 million years) of the sea onto the continental margin (Figure 125-23). A hiatus of no less than 135 million years separates the Muav Limestone from the Devonian-age Temple Butte Limestone. This rock unit is only exposed in

GEOLOGIC TIME SCALE

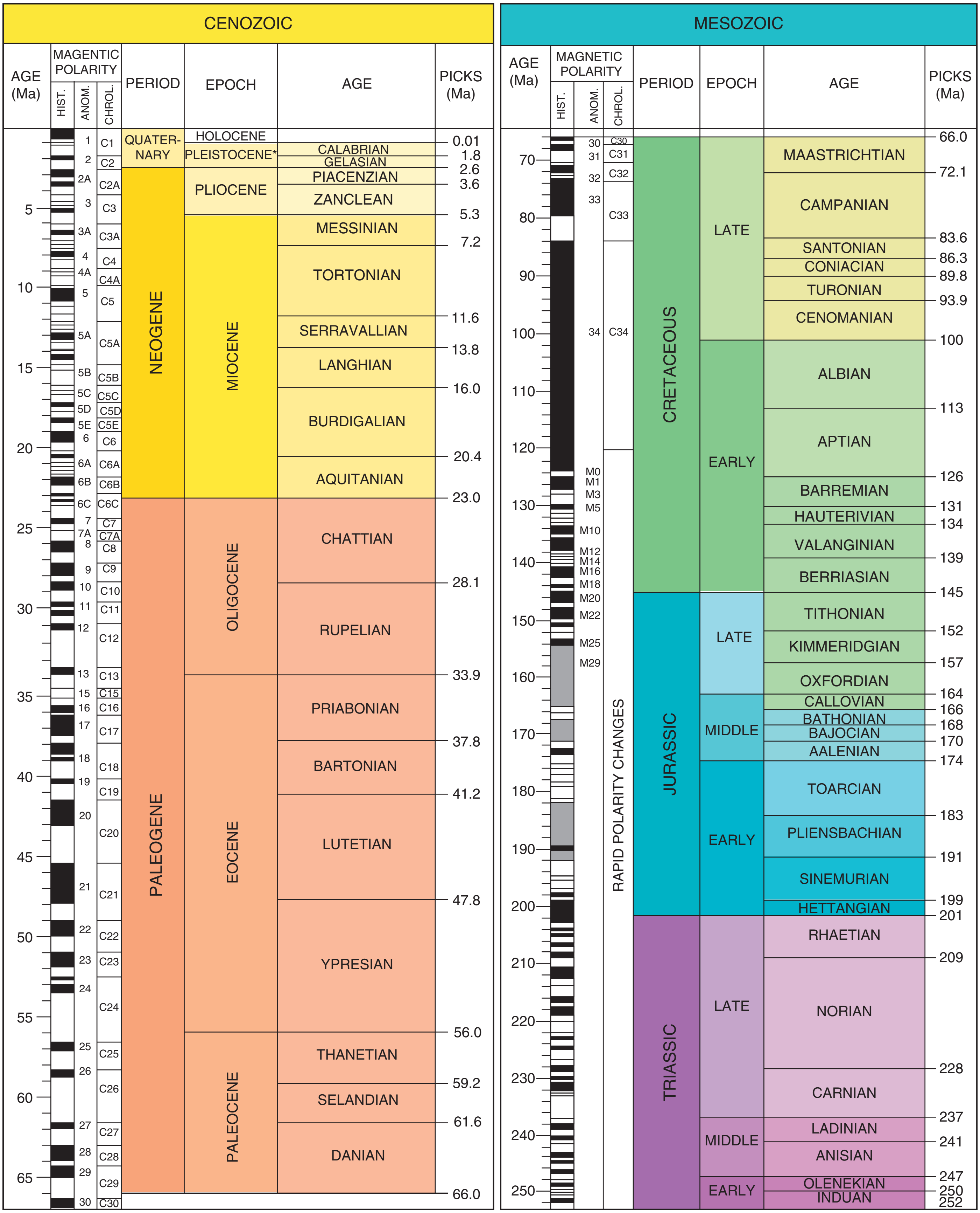
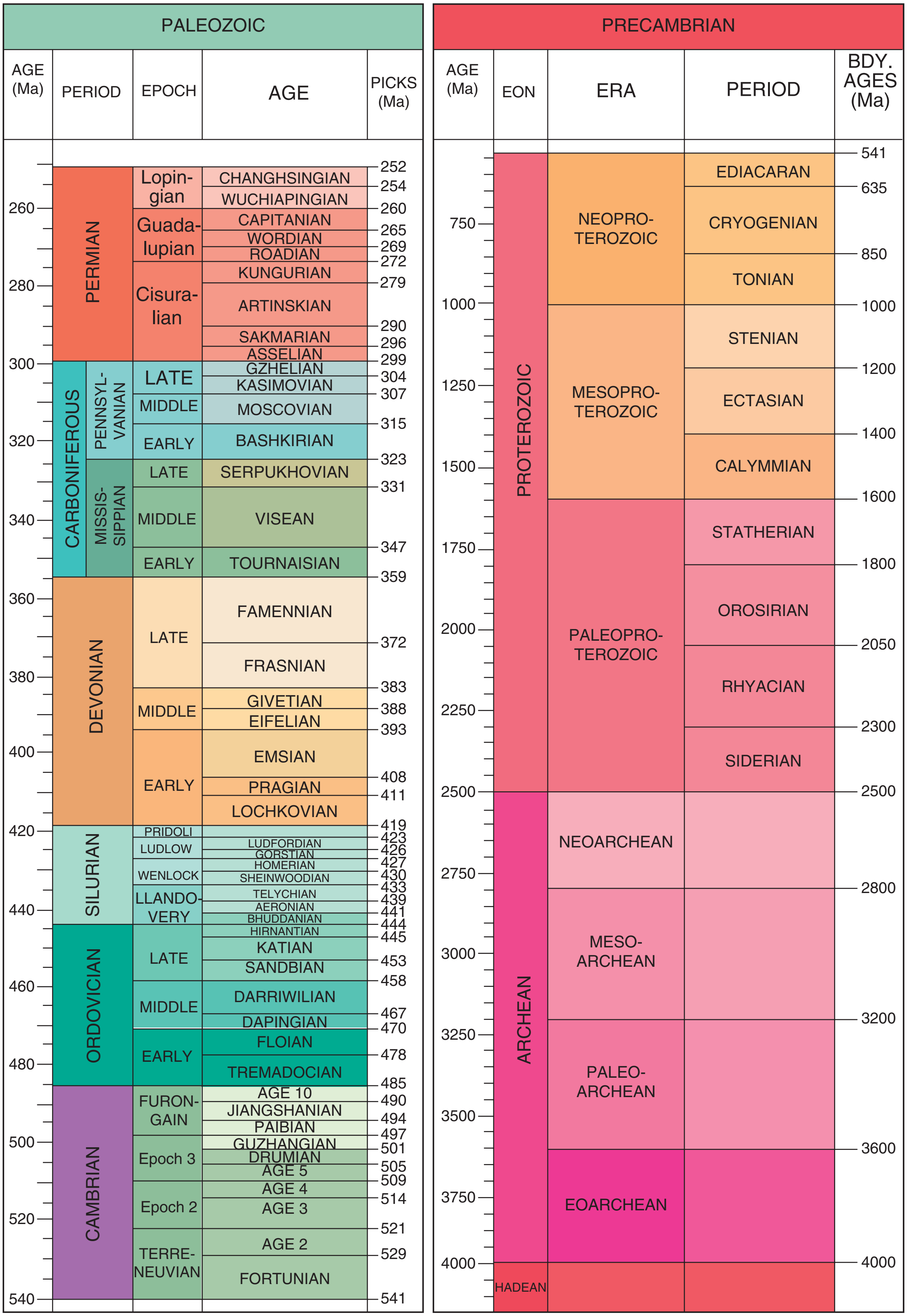


FIGURE 125-16 The geologic time scale.

GEOLOGIC TIME SCALE continued



PALEOZOIC			
PERIOD	EPOCH	AGE	PICKS (Ma)
			252
PERMIAN	Lopingian	CHANGHSINGIAN	254
		WUCHIAPINGIAN	260
	Guadalupian	CAPITANIAN	265
		WORDIAN	269
		ROADIAN	272
	Cisuralian	KUNGURIAN	279
		ARTINSKIAN	290
		SAKMARIAN	296
		ASSELIAN	299
CARBONIFEROUS – PENNSYLVANIAN	LATE	GZHELIAN	304
		KASIMOVIAN	307
	MIDDLE	MOSCOVIAN	315
	EARLY	BASHKIRIAN	323
CARBONIFEROUS – MISSISSIPPIAN	LATE	SERPUKHOVIAN	331
	MIDDLE	VISEAN	347
	EARLY	TOURNAISIAN	359
DEVONIAN	LATE	FAMENNIAN	372
		FRASNIAN	383
	MIDDLE	GIVETIAN	388
		EIFELIAN	393
	EARLY	EMSIAN	408
		PRAGIAN	411
		LOCHKOVIAN	419
SILURIAN	PRIDOLI		423
	LUDLOW	LUDFORDIAN	426
		GORSTIAN	427
	WENLOCK	HOMERIAN	430
		SHEINWOODIAN	433
	LLANDOVERY	TELYCHIAN	439
		AERONIAN	441
		RHUDDANIAN	444
ORDOVICIAN	LATE	HIRNANTIAN	445
		KATIAN	453
		SANDBIAN	458
	MIDDLE	DARRIWILIAN	467
		DAPINGIAN	470
	EARLY	FLOIAN	478
		TREMADOCIAN	485
CAMBRIAN	FURONGIAN	AGE 10	490
		JIANGSHANIAN	494
		PAIBIAN	497
	Epoch 3	GUZHANGIAN	501
		DRUMIAN	505
		AGE 5	509
	Epoch 2	AGE 4	514
		AGE 3	521
	TERRENEUVIAN	AGE 2	529
		FORTUNIAN	541

PRECAMBRIAN			
EON	ERA	PERIOD	BDY. AGES (Ma)
			541
PROTEROZOIC	NEOPROTEROZOIC	EDIACARAN	635
		CRYOGENIAN	850
		TONIAN	1000
	MESOPROTEROZOIC	STENIAN	1200
		ECTASIAN	1400
		CALYMMIAN	1600
	PALEOPROTEROZOIC	STATHERIAN	1800
		OROSIRIAN	2050
		RHYACIAN	2300
		SIDERIAN	2500
ARCHEAN	NEOARCHEAN		2800
	MESOARCHEAN		3200
	PALEOARCHEAN		3600
	EOARCHEAN		4000
HADEAN			

FIGURE 125-16, cont'd

FIGURE 125-17 The Grand Canyon of the Colorado River in Arizona is perhaps our planet's best monument to geologic time, sedimentation, and erosion. *(Photo by Wayne Ranney.)*

discontinuous channels in eastern Grand Canyon that thicken and converge into a 120-m-thick (400-foot-thick) continuous deposit in the western canyon.

Overlying the Temple Butte Limestone is the Redwall Limestone, a Carboniferous-age massive cliff-former located midway up the canyon walls. The deposit is 150 m (500 ft) thick and formed in an open marine setting that contains abundant crinoid, bryozoan, brachiopod, and coral fossils (Figure 125-24). The Tapeats to Redwall section of rocks reflects the passive margin (continental shelf) conditions that existed in western North America during the early Paleozoic, after the opening of the proto–Pacific Ocean.

Upper Carboniferous to Permian rocks known as the Supai Group document the gradual replacement of marine environments to more continental conditions during the late Paleozoic. The Supai rocks are interpreted in ascending order as nearshore, coastal floodplain, and eolian (wind-derived) deposits. Some vertebrate trackways have been found. The overlying brick-red Hermit Formation consists of sandstone, mudstone, and pebble conglomerate. It formed on a broad coastal plain in mostly fluvial (river) and eolian settings.

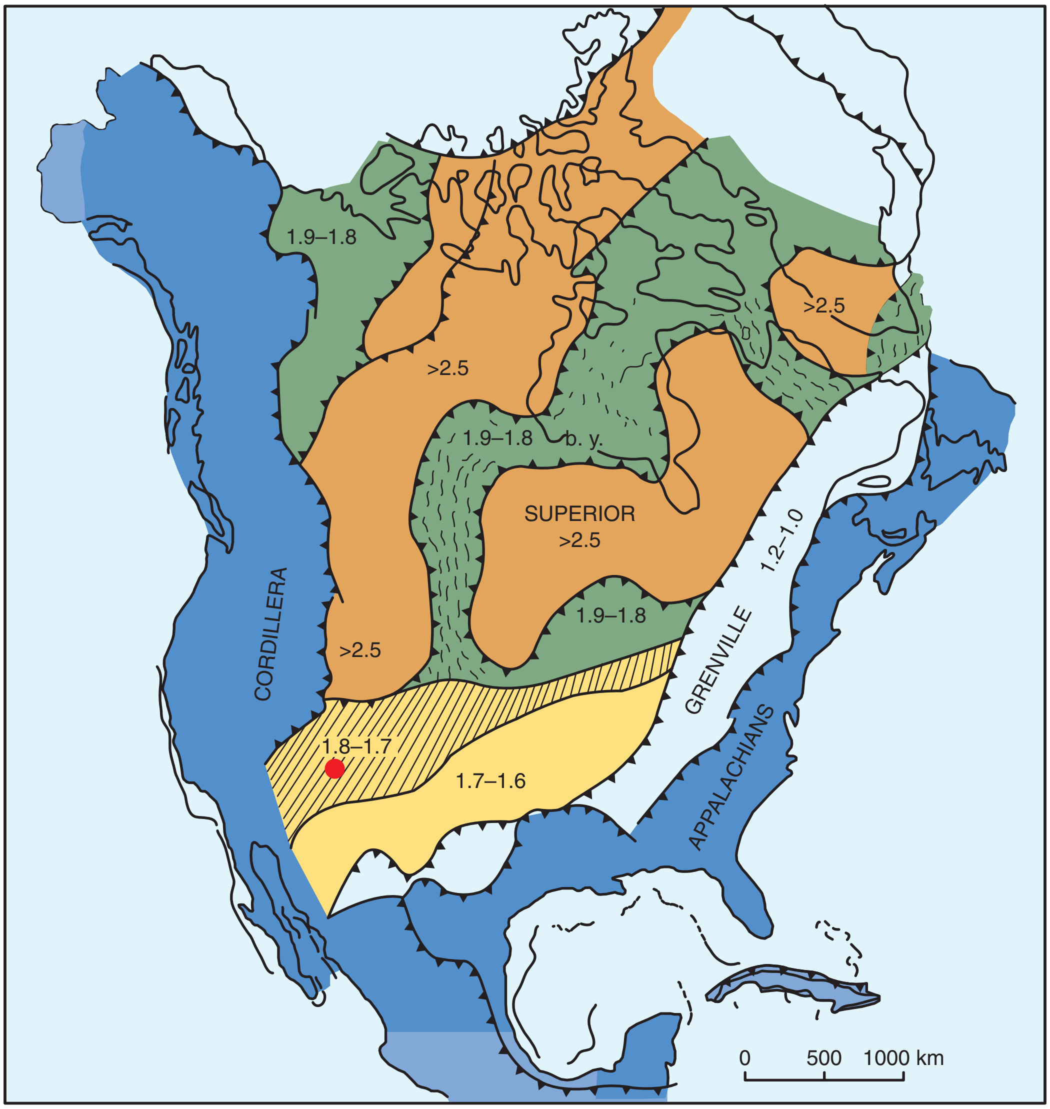


FIGURE 125-18 Map showing incremental growth of the North American continent. The oldest parts of the crust, formed more than 2500 million years ago (Ma), are labeled in *orange*; crust between 1900 and 1800 Ma, in *green*; 1800 to 1600 Ma, in *yellow*; and less than 300 Ma, in *blue*. The Vishnu rocks in the Grand Canyon are part of a terrane that arrived about 1750 to 1680 Ma. Red dot indicates approximate location of the Grand Canyon.

FIGURE 125-19 Spectacular view of the Vishnu basement rocks in the Inner Gorge of the Grand Canyon. *(Photo by Wayne Ranney.)*

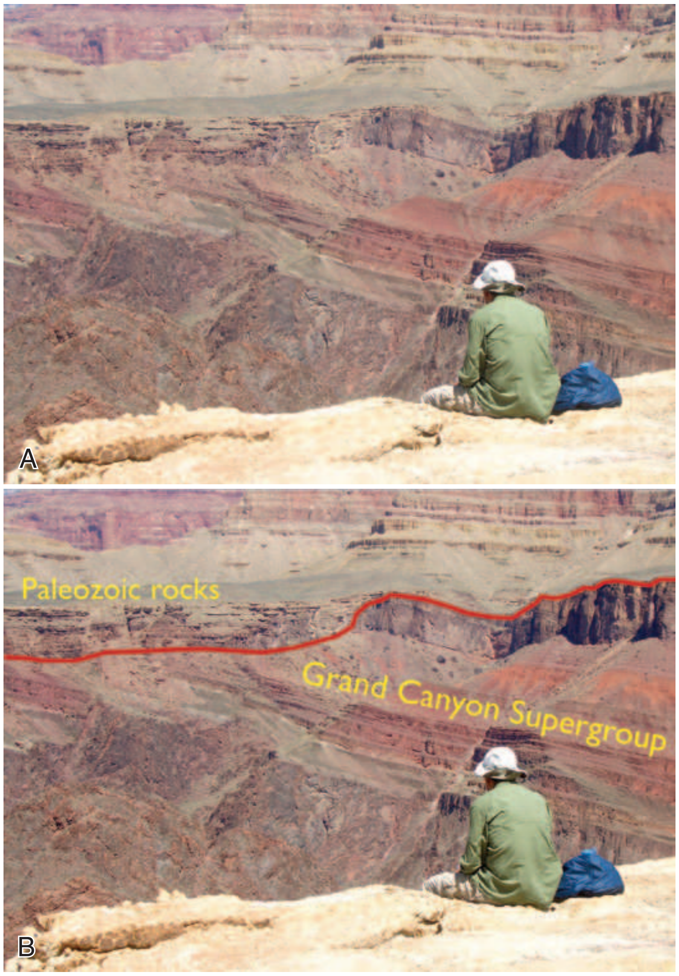


FIGURE 125-20 **A,** The Grand Canyon Supergroup, originally more than 3658 m (12,000 feet) thick, is found in only about 10% of the Grand Canyon and is tilted everywhere when seen. **B,** The *red line* shows the eroded top surface of the Supergroup, with a resistant cliff standing in the center of the photo. The younger, flat-lying Paleozoic layers (*top*) ultimately buried the Supergroup rocks. *(Photo by Wayne Ranney.)*

The next deposit is the Coconino Sandstone, a pale-yellow unit that is 100 m (330 feet) thick and forms sheer cliffs everywhere within the canyon. It originated in an arid, inland dune environment similar to the modern desert in the Sahara (Figure 125-25). Some layers contain numerous and well-preserved reptile trackways. The overlying Toroweap Formation is often overlooked in Grand Canyon because it forms slopes of easily eroded siltstone and gypsum that are covered in trees. It was deposited along the shore of a sea that encroached from the west. Capping the Grand Canyon and completing the entire Paleozoic section is the Kaibab Limestone. It represents a final transgression of the late Paleozoic sea into the area. Numerous chert horizons (a microcrystalline form of silica that often precipitates from seawater) help to solidify the Kaibab and make it the durable rock that "holds up" the strata in the canyon.

Mesozoic Rocks

A voluminous stack of Mesozoic-age rocks once covered the Grand Canyon area, but erosion has removed most of them. These rocks were once on the order of 1500 to 3000 m (5000 to 10,000 feet) thick and can be found in Zion and Bryce Canyon National Parks to the north, on the Navajo Indian Reservation to the east, and near Las Vegas, Nevada, to the west. Because the Grand Canyon lies between the three areas, it is logical to assume the rocks were once here as well, before erosion stripped them away.

Cenozoic Rocks

Rocks of Cenozoic age are relatively scarce at the Grand Canyon, because this was a time of regional uplift and erosion. Some river gravels are preserved that may indicate a time when the Grand Canyon began to form, no earlier than about 70 Ma. More recent volcanic rocks are located in the western Grand Canyon, where basalt lava, ranging in age from 830,000 to 1000 years, erupted along a 16-km (10-mile) stretch of the Colorado River. Numerous flows and cones are found perched above and within the canyon walls, with one remnant having traveled 135 km (84 miles) down the river channel, while others are perched up to 330 m (1100 feet) high. Up to 17 lava dams blocked the Colorado River, and

FIGURE 125-21 Paleozoic rocks were deposited in the Grand Canyon from 525 to 270 Ma and form the upper four-fifths of the canyon walls. *(Photo by Wayne Ranney.)*

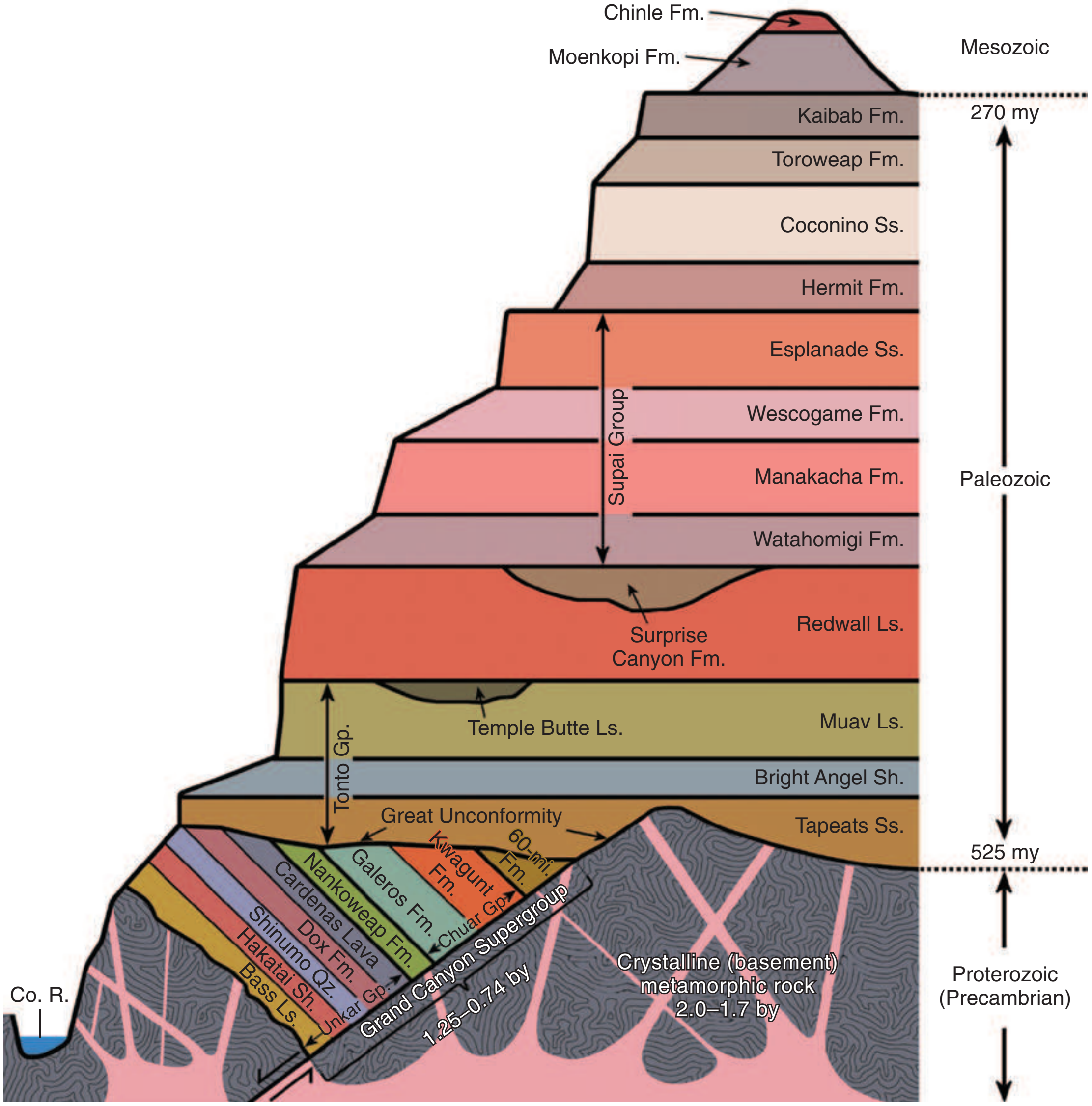


FIGURE 125-22 Rock column for the Grand Canyon.

on at least five occasions, huge outburst floods were the result of catastrophic failure of these dams. Deposits from these outburst floods are found 45 to 200 m (150 to 650 feet) above the modern channel, with clast sizes up to 30 m (100 feet) in diameter.

CARVING GRAND CANYON

Many theories have been proposed for how the Colorado River (or some ancestor to it) actually carved the Grand Canyon. Despite extensive research, some important details about its formation have escaped full detection. However, a broad outline is known, and future research will likely elucidate more of the story. This much is known: the Colorado River and its tributaries excavated this great space; it could not have happened before about 70 Ma; and it is likely that much of what we see today has formed in just the last 5 or 6 million years.

For approximately 455 million years, the Grand Canyon region was situated near, and many times below, sea level, thus precluding the presence of a deep canyon. A fantastic 4.8-km-thick (3-mile-thick) section of stratified rocks reveal this long-lived nearshore setting, a time when the rocks in Grand Canyon were formed. Beginning about 70 million years ago, the region was uplifted as the Farallon Plate was subducted beneath western North America. This event, known as the *Laramide orogeny* (an orogeny is a mountain-building event), uplifted a broad section of the continent from southern Arizona to the Rocky Mountains. A range of mountains southwest of the Grand Canyon, called the Mogollon Highlands, caused the first rivers to run from the southern mountains to the northeast, exactly opposite the direction of the modern Colorado River. This drainage lasted until at least 25 million years ago (Figure 125-26).

It is currently under debate if portions of the Grand Canyon could have been cut by this early northeast drainage system. Modern laboratory techniques are being used to tease more information out of Grand Canyon's stubborn rocks, but to date have yielded conflicting data. One group of researchers proposes that by about 70 Ma, some portions of the canyon were cut to within 100 m (330 feet) of its current depth, whereas another group says that only one of five subsections of the canyon was carved at this early date.

Geologists "lose sight" of the river and canyon for about 20 million years, because a period of erosion or perhaps even nondeposition ensued. This certainly was the time of drainage reversal, whereby the old northeast-directed system gave way to the current southwest-flowing river. Deposits at the mouth of the Grand Canyon that are between 16 and 6 million years old show that the modern river might not have been flowing here before about 5 or 6 million years ago. This is why the majority of geologists think that the river (and thus the canyon) could be no older than this age. However, some alternative scenarios have been proposed that can explain why no Colorado River sediment

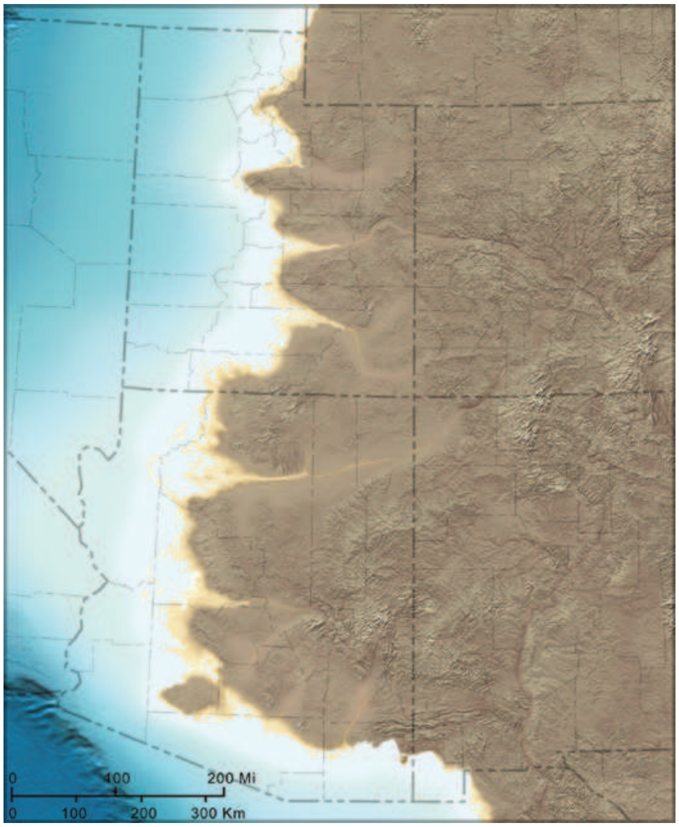


FIGURE 125-23 Paleogeographic view of the Four Corner states during deposition of the Tapeats Sandstone 525 Ma. The map reflects a time before land plants evolved; thus the land area (*right*) is shown as *brown. (From Blakey R, Ranney W. Ancient landscapes of the Colorado Plateau, 2008, Grand Canyon Association.)*

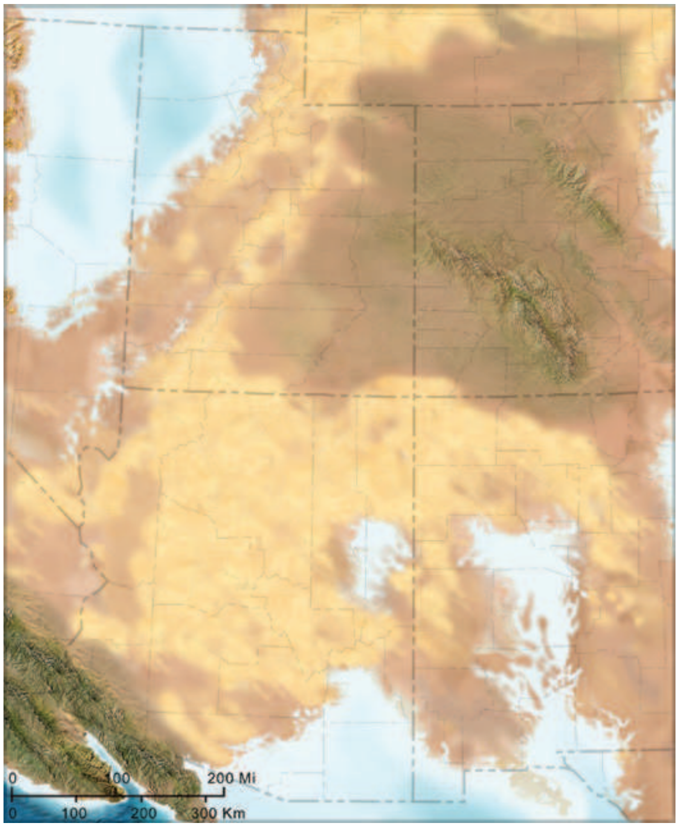


FIGURE 125-25 The Coconino Sandstone reveals the presence of a Sahara-like desert in the Grand Canyon area about 275 Ma.

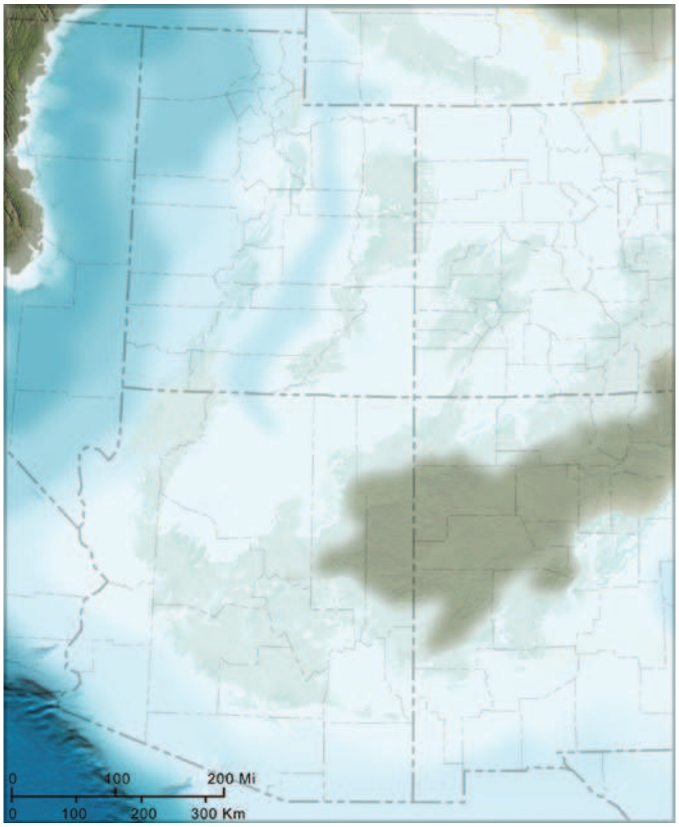


FIGURE 125-24 The Redwall Sea covered much of the American Southwest about 340 Ma.

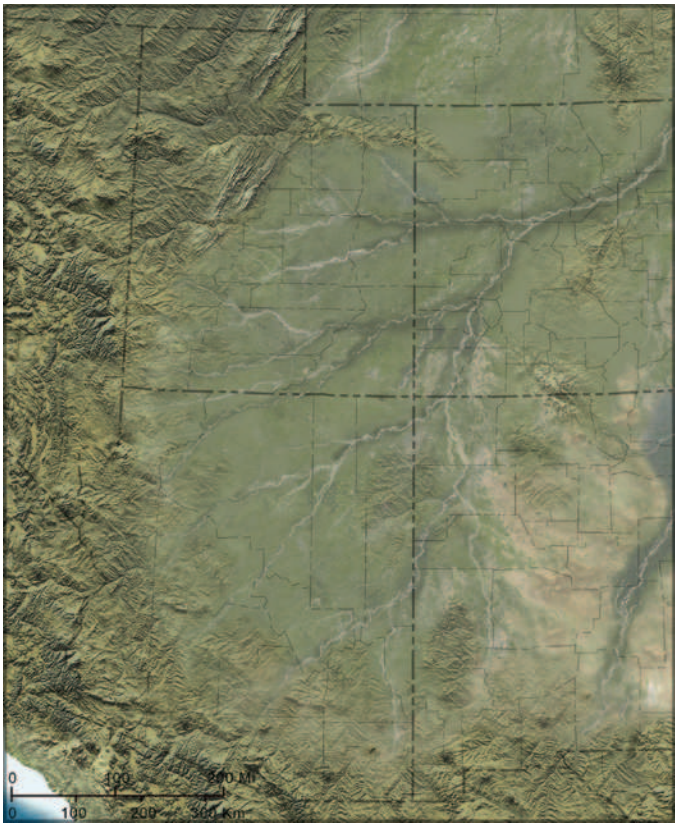


FIGURE 125-26 When the sea finally left the Grand Canyon region for the last time, an initial river system developed from the Mogollon Highlands (*bottom*) to the northeast. This is opposite to the direction of the modern river system and lasted for at least 45 million years.

resides in these deposits (16 to 6 million years old), and so the "old canyon/young canyon" debate continues.

What is known is that a modern Colorado River arrived at the mouth of the Grand Canyon and began to fill a series of lake basins along the present-day course of the lower Colorado River on its way to the Gulf of California. Certain deposits reveal that the river sequentially filled closed basins, first with water and then sediment. Eventually, the lake water overtopped a divide and began to fill another downstream closed basin. This occurred four times on its way to the sea and thus created a course of the lower Colorado River. All this occurred over a 1- to 1.5-million-year period from about 5.6 to 4.1 million years ago.

The previous process is known as *basin spillover* and is one of three processes invoked for how the river in Grand Canyon may have been integrated from separate ancestors. The other two are *headward erosion,* whereby one river lengthens its channel upstream to intersect and capture another river, and *karst collapse*, whereby groundwater establishes a subsurface connection through caverns that ultimately collapses to form a surface connection. Any one, two, or all three of these processes could explain the Colorado River in the Grand Canyon.

SUMMARY OF GRAND CANYON GEOLOGY

A picture that is emerging is that headward erosion, basin spillover, and/or karst collapse helped integrate the Colorado River and create the Grand Canyon. Most geologists agree that the canyon we see today is the result of deep incision within only the last 5 to 6 million years. However, parts of this "modern" canyon may overprint or incorporate some sections of older canyons. A broad outline of the major events forming the Grand Canyon is now possible: (1) the Laramide orogeny produced an initial river system with drainage to the northeast, with possible early segments of Grand Canyon carved; (2) drainage in the region became disrupted between about 25 and 6 Ma, resulting in few deposits; and (3) integration of the lower Colorado River by basin spillover created an outlet to the sea from the Rocky Mountains.

The Grand Canyon continues to inspire as a unique and remarkable landform. It remains one of the most impressive outdoor laboratories for the study of Earth history. In the almost 160 years that it has been studied scientifically, much has been learned about fluvial processes acting on an uplifted, arid landscape. Over time, and with improved dating methods, the canyon will continued to reveal more of its secrets.

SELECTED RESOURCES

Selected resources used in this text are available online at expertconsult.inkling.com.

CHAPTER 126

Space Medicine: The Next Frontier

THOMAS H. MARSHBURN, RICHARD W. COLE, AND REBECCA S. BLUE

Humans are at the threshold of a phase of exploration equal in significance to any yet undertaken by the species. Having developed technology that enables escape from the gravitational forces that hold us to the earth's surface, human space programs have demonstrated that we can survive, function, and perform complex tasks in continuous microgravity. The past century has witnessed humans piloting spacecraft through launches, landings, rendezvous, and dockings in Earth orbit; constructing complex vehicles and habitats that allow long-term occupancy; and expanding human presence in the Solar System beyond the limits of Earth's atmosphere.

An international coalition of space programs from the United States, Russia, Europe, Canada, and Japan continuously inhabit low Earth orbit. We have collectively experienced more than 15 years of human presence aboard the International Space Station (ISS), the largest space vehicle ever assembled, with a module length of 50.9 m (167 feet) and a habitable volume of about 906.14 m^3 (32,000 $feet^3$) (Figure 126-1). More recently, China has joined the roster of spacefaring nations, launching five-crewed missions since 2003 and maintaining the Tiangong-1 space station in low Earth orbit for visiting crews since 2012.

With increasing human experience in space, the field of space medicine continues to mature and advance, improving understanding of the unique threats spaceflight poses to human health and developing a foundation of medical diagnosis and treatment techniques specific to spaceflight-related concerns. While drawing heavily from many medical and surgical specialties, aerospace medicine is a unique field of study and medical practice. It addresses the challenges of maintaining long-term human health and performance in the face of the following four fundamental threats:

1. *Microgravity.* Astronauts travel in craft with sufficient energy to attain a free-fall state in perpetuity, no longer encumbered by Earth's gravity gradient. The resultant unloading of body fluids and tissues has profound effects on human physiology and health and can have an impact on timing and mechanism of recovery from injury and illness. Further, the microgravity environment alters the ability to deliver medical care.
2. *Radiation.* The farther humans travel into space, away from the protective shroud of Earth's geomagnetosphere, the more they are exposed to greater intensities of highly ionizing radiation, pervasive throughout the universe. This type of radiation has the potential to inflict immediate injury and long-term harm to human health.
3. *Psychological effects of isolation.* Living in space entails living in close quarters with few crewmates in a dangerous environment, with visual and tactile surroundings quite dissimilar to those found in daily life on Earth, and an unprecedented disconnect from the rest of humankind. The psychological impact of this unique combination of stressors can be a significant challenge to human performance during space exploration.
4. *The spacecraft environment.* To survive in space, astronauts must rely entirely on small, enclosed, Earth-like ecosystems aboard their vehicles. These systems maintain a breathable atmosphere, water and food supply, hygiene methods, and other critical elements that, when disrupted, can negatively affect human health and are vulnerable to any insults that could endanger their structural integrity.

Despite these challenges, access to space may soon increase exponentially. Spacefaring nations are expanding their reach toward the goals of habitation on the lunar surface, interplanetary

SELECTED RESOURCES

Blakey R, Middleton L. Geologic history and paleogeography of Paleozoic and early Mesozoic sedimentary rocks, eastern Grand Canyon, Arizona. In: Timmons J, Karlstrom K, editors. Grand Canyon geology: Two billion years of earth history. Geological Society of America Special Paper 489. 2012. p. 81–92.

Davis S, Dickinson W, Gehrels G, et al. The Paleogene California River: Evidence of Mojave-Uinta paleodrainage from U-Pb ages of detrital zircons. Geology 2010;38:931–4.

Flowers R, Wernicke B, Farley K. Unroofing, incision and uplift history of the southwestern Colorado Plateau from apatite (U-Th)/He thermochronometry. Geol Soc Am Bull 2008;120:571–87.

Hill C, Ranney W. A proposed Laramide proto-Grand Canyon. Geomorphology 2008;102:482–95.

Hill C, Eberz N, Buecher R. A Karst Connection model for Grand Canyon, Arizona. Geomorphology 2008;95:316–34.

House K, Pearthree P, Perkins M. Stratigraphic evidence for the role of lake spillover in the inception of the lower Colorado River in southern Nevada and western Arizona. In: Reheis M, Hershler R, Miller D, editors. Late Cenozoic drainage history of the southwestern Great Basin and lower Colorado River region: Geologic and Biotic perspectives. Geological Society of America Special Paper 439. 2008. p. 335–53.

Karlstrom K, Beard S, House K, et al. CRevolution 2: Origin of the Colorado River system II, Geological Society of America. Geosphere 2012; 8:1–7.

Karlstrom K, Ilg B, Hawkins D, et al. Vishnu basement rocks of the Upper Granite Gorge: Continent formation 1.84 to 1.66 billion years ago. In: Timmons M, Karlstrom K, editors. Grand Canyon geology: Two billion years of earth history. Geological Society of America Special Paper 489. 2012. p. 7–24.

McKee E, Wilson R, Breed W, Breed C, editors. Evolution of the Colorado River in Arizona. Museum of Northern Arizona Bulletin 44. 1967.

Ranney W. Carving Grand Canyon: Evidence, theories and mystery. Grand Canyon Association; 2012.

Wernicke B. The Colorado River and its role in carving Grand Canyon. Geol Soc Am Bull 2011;123:1288–316.

Young R, Spamer E, editors. The Colorado River: origin and evolution. Grand Canyon Association Monograph No 12. 2004.

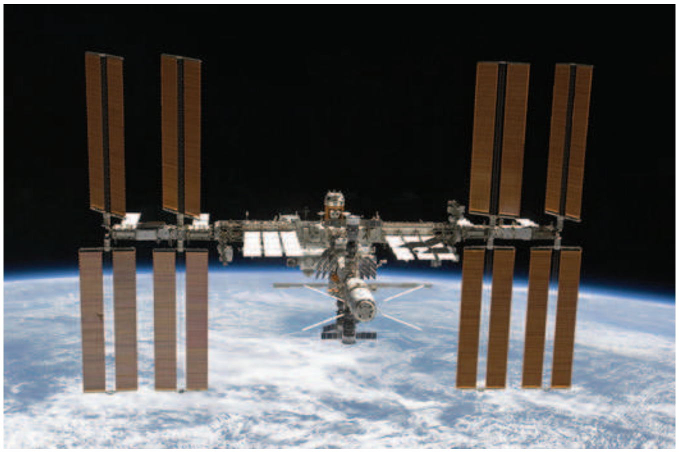


FIGURE 126-1 The International Space Station (ISS), a National Research Laboratory currently orbiting the earth approximately 250 nautical miles above Earth's surface. Assembled by five partner agencies, the ISS has been continuously manned since November 2000.

travel to Mars, and commercial spaceflight offered to the general population. Acknowledging this rapid expansion of spaceflight and the unique role of space as the ultimate wilderness environment, this chapter summarizes the distinctive environmental factors experienced by astronauts; describes the clinically relevant physiologic and psychological effects of space travel, particularly as related to human function and adaptation in space; and provides an understanding of the medical practice of modern flight surgeons and astronaut physicians. We further explore gaps in current clinical understanding with respect to the knowledge we must gain to reduce barriers to longer and more expansive space travel in the future.

THE SPACEFLIGHT ENVIRONMENT

The spaceflight environment is complex and variable. In general, space is uninhabitable to humans and fraught with dangers that would rapidly prove fatal without the protection afforded by a vehicle with intrinsic life support systems. Major challenges of the spaceflight environment include pressure (specifically, lack of pressure), thermal concerns, gaseous factors (e.g., availability of oxygen, concentrations of carbon monoxide and carbon dioxide), presence of onboard hazardous materials (e.g., coolants, propellants, other contaminants), and risk of fire.

PRESSURE

The human body requires a certain amount of ambient pressure to sustain life. Humans exposed to vacuum conditions, as exist outside of Earth's atmosphere, would experience a number of physiologic insults, including decompression illness, hypoxia and impeded gas exchange, and ebullism (the evolution of body fluids to gas); each of these insults can rapidly lead to physiologic dysfunction and death. Integrity of life support systems, particularly in providing the livable pressure atmosphere, is a major determinant of human survival in spaceflight.

Spaceflight Decompression Risks

Decompression is a serious concern during spaceflight and can occur either intentionally or from a contingency event, such as collision, structural failure, or human error. Intentional decompressions occur in normal flight operations, usually to support spacewalks, or extravehicular activities (EVAs). Astronauts performing EVAs do so in spacesuits that operate at pressures lower than the common sea level cabin pressure of 14.7 pounds per square inch absolute (psia), typically between 4 and 6 psia depending on the suit, using 100% oxygen (O_2) as a breathing gas. To prevent evolution of tissue nitrogen bubbles during decompression to these lower pressures, ISS astronauts perform an O_2 prebreathe session to afford nitrogen washout, followed by a staged decompression within the airlock to final suit pressure. The usual EVA sortie lasts for about 6 hours, after which crews can fairly rapidly recompress back to station sea level pressure. Although such intentional depressurizations can cause injury, including decompression sickness (DCS), such events are rare. In fact, the only recorded DCS events in spaceflight occurred with planned cabin depressurization during launch of a Gemini spacecraft from sea level to an in-flight operating pressure of 5 psia, when a crewmember experienced sharp pain in one knee. The same astronaut experienced similar pain when he later launched in an Apollo spacecraft.[89]

Contingency depressurization is of greater concern than are planned decompression protocols with respect to likelihood of crew injury. The most likely events leading to contingency depressurization are collision with other vehicles, as occurred between the Russian Mir station and an uncrewed cargo vessel in 1997, and hypervelocity orbital debris impact. Whereas vehicle collision can be avoided with engineering and human factor practices, orbital debris is a much less controlled hazard. As human presence in space has increased, so has the risk of orbital debris impacts. Any component of derelict spacecraft, including abandoned stages, decomposing satellites, or even minute debris such as paint particles, can pose a risk to other space vehicles. While orbital debris pieces larger than 10 cm (4 inches) in diameter can be carefully tracked and avoided by means of ground-based radar, smaller pieces are difficult to track and can be highly damaging because of the speeds at which they travel. The highest-risk objects are 1 to 10 cm in (0.4 to 4 inches) diameter, large enough to be seen by ground radar but not tracked and avoided, and still able to transmit highly damaging force to spacecraft. The force imparted by collision with even a small debris particle is surprisingly large because of the velocity of the object and kinetic energy being proportional to the square of this velocity. For example, a 1 cm^2 (0.15 $inch^2$) piece of aluminum weighing 3 g (0.0066 lb) and traveling at orbital velocity of approximately 8 km/sec (17.89 miles/hr) in low Earth orbit would deliver 96,000 J of force on impact. The effects of orbital debris impacts depend on velocity of the debris, angle of impact, and the object mass, as well as sensitivity of the surface that is hit. A small, glancing blow may not cause significant damage to a heavily reinforced portion of a space vehicle, but it could cause catastrophic damage to a solar panel. Unfortunately, impacts are now a feature of space travel, with numerous reports of impact craters found on space vehicles, one of the largest being 3.8 mm (0.15 inch), on the window of Soyuz T-9 in 1983.[170]

The longer a vehicle stays in orbit, the greater its exposure to some manner of debris, and the more important it is to ensure that shielding is effective. The ISS has the most robust shielding ever flown and routinely performs maneuvers to adjust the orbit of the entire station to avoid collision with orbital debris; even so, small debris impacts do occur. For this reason, protocols are in place to prepare astronauts for a contingency decompression after debris strike or vehicle collision. In a decompression event, sensors would alarm at a pressure differential of greater than 1.0 psi/hr or a total pressure decrease of 0.4 psi. Crews are trained to respond by first ensuring individual access to supplemental O_2 and a clear route to an escape vehicle, then detecting and isolating the leak or, in worst-case scenarios, abandoning the station to return to Earth.

Decompression-Related Injuries

Decompression-related injuries in spaceflight are expected to be similar to those seen in other dysbaric events, such as rapid pressure changes during scuba diving or rapid decompression of an aircraft at altitude. Sequelae can include hypoxia from reduced partial pressure of oxygen (P_{O_2}), barotrauma events (e.g., pneumothorax, gastrointestinal barotrauma), musculoskeletal decompression symptoms (e.g., joint pain), neurologic sequelae of DCS, and arterial gas embolism (clinical manifestations are discussed in Chapters 71 and 72). DCS is avoided during nominal decompression procedures for EVA through a combination of prebreathing 100% O_2 and staged crew decompression to facilitate nitrogen washout (Figure 126-2). One recently adopted adjunct

FIGURE 126-2 Astronauts Rick Mastracchio (*right*) and Mike Hopkins in preparation for a spacewalk in the ISS airlock, breathing 100% oxygen while donning their spacesuits as part of a protocol to reduce the risk of decompression sickness. Note the white cooling and ventilation undergarments that crews wear inside their spacesuits to regulate body temperature.

is addition of a light-exercise protocol; extensive research has evaluated exercise as a means of accelerating nitrogen washout during reduced-pressure O_2 prebreathe sessions.[36] Another option is to lengthen the O_2 prebreathe, then decompress directly to suit pressure, although this may imply logistical overhead of more in-suit time before decompression.

Accidental DCS events during a controlled decompression could occur; most would be expected to be treatable by returning to sea level pressure and continuing to breathe 100% O_2.[126] Decompression injury could also be treated on orbit by using an EVA suit as a pressure chamber. At standard sea level atmospheric pressure inside the ISS, the suit with the affected crewmember inside can be further pressurized with 100% O_2 to attain 22 psia. Although this is not as effective as a hyperbaric chamber and a standard hyperbaric protocol, the increased pressure may offer some benefit for acute DCS injuries, although isolation of the affected crewmember in a suit would limit other examination and interventional opportunities.[73] Further, limitations of on-orbit medical capability renders treatment of most of the more severe decompression-related conditions listed practically untenable, and evacuation and return to Earth may be necessary.

With the most catastrophic events, exposure to true vacuum conditions is also possible. If crewmembers are exposed to ambient pressure of less than 1 psia, ebullism occurs in addition to the classic DCS, hypoxia, and barotrauma insults. *Ebullism* is spontaneous evolution of water in tissues from a liquid to gaseous state at ambient pressures of less than 47 mm Hg (~0.9 psia), where the boiling point of water is less than or equal to the homeostatic temperature of the human body.[165] In *ebullism syndrome,* multiple insults occur, including anoxia, trapped-gas expansion and formation of nitrogen and water vapor bubbles, and rapid body fluid loss.[165] With an explosive decompression to vacuum conditions, crewmembers would be rapidly rendered unconscious, generally in less than 10 seconds. Even if astronauts were wearing supplemental O_2, the ambient pressure would be too low to allow for effective pulmonary gas exchange. Within 30 seconds of exposure, severe neurologic hypoxic sequelae would occur, leading ultimately to brain death as a result of hypoxia or ischemic insult from bubble formation within the vasculature. Bubbles can also form within thoracic organs, including the heart, and impede intrathoracic circulation, compounding ischemic insult. Circulatory arrest would rapidly follow.

However, very short exposures to complete vacuum are survivable. In 1966, a spacesuit technician was exposed during a test to near-vacuum conditions for 30 seconds in an altitude chamber after his suit umbilical accidentally disconnected from its pressurized O_2 supply. The technician lost consciousness in 12 to 15 seconds,[197] although was later able to report noting a fizzing sensation on his tongue from sublimation of saliva. He regained consciousness seconds after rapid repressurization and surprisingly, suffered no sequelae (LeBlanc J: Personal communication, 2015). An individual exposed to 22,555 m (74,000 feet) equivalent altitude for approximately 3 minutes in another vacuum chamber accident suffered pulmonary barotrauma and massive pulmonary and neurologic sequelae of DCS, with eventual clinical resolution after hyperbaric and prolonged medical treatment.[120] Protocols have been developed and published for treatment of ebullism syndrome based on successful rehabilitation of these individuals and continuous advancement of supportive medical care.[165] Aside from rapid recompression to cabin atmosphere, few options exist for ebullism treatment on orbit. Unless exposure is isolated to a single crewmember and ambient pressure could be rapidly restored, or exposure occurs during a rapid descent to normal atmosphere during reentry, ebullism injury in a spacecraft vehicle would likely prove rapidly fatal.

In-Flight Decompression Events

Despite careful debris tracking and avoidance and extensive crew training on the risks and mitigations of loss of cabin pressure, significant decompressions have occurred during human spaceflight. In 1997, the Russian Spektr module was damaged and depressurized by a collision with an uncrewed Progress resupply ship during an attempted docking with the Mir space station. In this case, the crew quickly sealed the hatch to the leaking Spektr, preventing further depressurization in the remaining modules of the space station and avoiding crew injury.

A more catastrophic loss of pressure occurred in 1971 when a pressure-equalization valve failed during reentry in the Soyuz 11 spacecraft at an altitude of 168 km (105 miles). The three cosmonauts aboard the Soyuz (Georgi Dobrovolskiy, Vladislav Volkov, and Victor Patsayev) rapidly lost consciousness during the decompression, and it is believed they had suffered fatal injuries within 1 minute of exposure. Preliminary autopsy results reportedly revealed intracerebral and pulmonary hemorrhage as a result of the approximately 10-minute exposure to near-vacuum conditions.[205]

Given the potential for catastrophic outcomes of decompression-related events in spaceflight, prevention of any loss of cabin pressure will remain of the utmost importance in vehicle design, trajectory and orbital maneuvering, and crewmember training.

OXYGEN

Directly related to the atmospheric pressure of a space vehicle cabin is the amount of O_2 available. Decreases in ambient pressure lead to resultant decreases in the partial pressure of O_2 available for gas exchange, which can lead to hypoxic symptoms despite compensatory pulmonary and cardiac responses. All current crewed vehicles, including the ISS and all its resupply craft, are nominally pressurized to 14.7 psia with a dual-gas system at near-terrestrial composition (21% O_2 and 79% nitrogen). However, astronauts are occasionally exposed intentionally to lower ambient pressures, most often during EVA. Lower spacesuit pressures are useful for many reasons, particularly for manipulating suit components such as gloves, improving mobility, and reducing the physical strain of activity within the suits. However, in addition to the risk of DCS, the lower pressure of a spacesuit does expose astronauts to the potential for hypoxia from reduced PO_2. Given these lower pressures and the simplicity of a single-gas system, suits are operated at 100% O_2.

Given the risk of vehicular depressurization or disruption of the internal gas environment from equipment malfunction, spacecraft are rigorously monitored for O_2 levels, and an insidious decrease in PO_2 would likely be detected by automatic sensing systems before symptomatically affecting the crew. Similarly, the EVA suit environment is carefully monitored for pressure and levels of O_2 and carbon dioxide. Any slight disruption of the internal suit environment would prompt rapid termination of an

EVA and return to the safety of the vehicle. Disruption of the cabin environment would prompt the crew to seek individual supplemental O_2 equipment to mitigate hypoxia until the problem is addressed, or in the worst-case scenario, the vehicle is abandoned.

The aerospace medicine practitioner must remain hypervigilant for hypoxic signs and symptoms because of the potential for exposure of entire astronaut crews to hypoxic conditions. At an ambient pressure of 10 psi (equivalent-pressure altitude, 3408 m [10,000 feet]), PO_2 drops to 109 mm Hg (~2 psi), and intra-alveolar concentrations of O_2 approach 60 mm Hg (~1.16 psi).[79] Decompression of the cabin atmosphere from 14.7 to 10 psia or lower would prompt physiologic compensatory responses to hypoxia in most individuals. Immediate responses are identical to those in patients at high altitude: hyperventilation and tachycardia to increase tissue O_2 delivery. Crewmembers may report more obvious evidence, such as ear popping, if decompression is rapid. However, hypoxic conditions secondary to dysfunction of gas-mixing systems may not be accompanied by a pressure differential, and the only signs or symptoms of a hypoxic environment may be vague and poorly defined feelings of fatigue, malaise, apprehension, paresthesias, visual changes, and nausea. Because none of the immediate symptoms of acute hypoxia can be considered pathognomonic, clinicians must remain vigilant for any crewmember complaints that may suggest an insidious hypoxic event and thus the need for O_2 replenishment or supplementation.

CARBON DIOXIDE

On Earth, production of carbon dioxide (CO_2) as a normal byproduct of ventilation is physiologically insignificant, because the gas is rapidly dispersed into the atmosphere and CO_2 concentration remains quite low, at an average of 0.23 mm Hg (.004 psi, or 0.03%). However, in a closed spacecraft environment, CO_2 would rapidly build up without effective atmospheric scrubbers. Therefore, cabins must be designed not only to provide the O_2 that a crew requires, but also to remove the CO_2 that the crew produces. Cabin scrubbers, typically adsorbents such as lithium hydroxide or zeolite, accomplish this task by chemically removing CO_2 from the cabin atmosphere. However, components of CO_2 removal systems have a limited lifetime, making it impractical to maintain CO_2 continuously at terrestrial levels. Furthermore, without effective airflow, localized concentrations of CO_2 can exist in microgravity, particularly around a stationary crewmember's face. For this reason, forced convection is established by using cabin fans in an attempt to ensure effective cabin airflow and avoid stagnant gas concentrations. Because structural elements may disturb airflow, astronauts are often exposed to elevated levels of CO_2, particularly in less ventilated areas or more populated spaces within the cabin, where individuals and their CO_2 byproducts tend to congregate and accumulate.

It was previously believed that physiologic effects of elevated CO_2 would be recognized only at levels of greater than 12 mm Hg (0.2 psi, or 1.5% CO_2). However, more recent evidence suggests that alterations in mood can occur at levels as low 0.5% CO_2 (4 mm Hg, 0.07 psi).[130] Headache is the most common initial complaint of CO_2 toxicity in space and has been reported at levels of less than 5 mm Hg of CO_2 in the cabin atmosphere. Because CO_2 is known to be a potent vasodilator that can increase cerebral blood flow and therefore possibly affect intracranial pressure (ICP), elevated CO_2 is also a possible contributor to suspected ICP increases on long-duration missions. This phenomenon is discussed later.

In a standard developed from terrestrial data and in concert with U.S. Occupational Safety and Health Administration (OSHA) and National Institute for Occupational Safety and Health (NIOSH) recommendations published in the 1980s, the average 24-hour CO_2 levels aboard the ISS were initially allowed to climb up to 7.6 mm Hg (0.14 psi).[101] This limit was later lowered to 5.3 mm Hg (0.1 psi) after several individuals onboard reported CO_2-related symptoms at the 7.6 mm Hg level; the decrease was further supported by new ground-based information from the early 1990s that suggested elevated CO_2 could lead to secondary effects, including discomfort and decreased exercise tolerability.[242] However, astronauts continued to report symptoms even at the 5.3 mm Hg level. Crew symptomatology includes headache, fatigue, and self-reports of decreased work efficiency. Although there is significant interindividual variability, the frequency of symptoms resulted in the CO_2 threshold being lowered again in 2013, to 4.0 mm Hg (0.07 psi, or 0.5%).

New data continue to accumulate. In general, CO_2 levels below 2.3 mm Hg (0.04 psi, or 0.3%) seem to cause few crew symptoms, but levels as low as 2.3 to 2.7 mm Hg (0.04 to 0.05 psi, or 0.3 to 0.35%) can result in complaints of headache or feelings of "full-headedness" and decreases in crew work efficiency. Currently, CO_2 levels are generally maintained below 3.0 mm Hg (0.06 psi, or 0.4%) to minimize symptoms, with allowable excursions above this limit for periods less than 24 hours.

TEMPERATURE

The touch temperature of external spacecraft objects in low Earth orbit can range from approximately −100° C to +130°C (−148° F to +266° F) between shaded and sunlit sides of the vehicle (Figure 126-3). These extremes require active thermal control and insulation of crewed vehicles to ensure temperature management within habitable ranges for crewmembers. On the ISS, the crew controls the temperature to a typical range of 21° to 22° C (69.8° to 71.6° F), a "shirtsleeve" environment. However, this is a greater challenge than it appears because in the space environment, convection (transfer of heat through air or fluid currents) does not naturally occur; as a result, heated objects tend to stay warm, and cold objects tend to stay cold. Temperature variations in space are primarily determined by radiation; only objects receiving direct or reflected sunlight become warmed, and objects that fall into shadow rapidly cool to the low temperature extremes.

These temperature variations become apparent to a spacewalker during EVA. Despite a spacesuit's insulation, electric glove heaters, and manual temperature controls, the surface temperatures of a spacecraft are noticeable. Spacesuit gloves are designed to protect human skin from incidental contact with the ISS external surface temperatures, but they function best in touch temperatures ranging from −115° C to +115° C (−175° F to +239° F). Therefore, contact with surfaces receiving direct sunlight may be too hot for suit gloves to mitigate the heat delivered to an astronaut's hand. Likewise, prolonged contact with metal surfaces entirely encased in shadow can result in numbness despite the

FIGURE 126-3 Astronaut Ron Evans performing a spacewalk during the transit of the Apollo 17 spacecraft from the moon back to Earth. *(Courtesy NASA.)*

use of glove heaters or, worse, might lead to significant injury. During one ground chamber test, a shaded object in a vacuum became so cold as to cause a frostbite injury from instrument handling through suit gloves. Should such an injury occur during spaceflight, limited medical capabilities could lead to permanent damage.

Thermal conduction from structural contact is not the only thermal concern for an astronaut during EVA. Work outside the spacecraft can be very demanding. Even the simplest tasks require exquisite control of a human-suit complex with considerable mass in the microgravity environment. As a result, astronauts can and do overheat during rigorous EVA activities. The human body cools through many mechanisms, but the most rapid and effective method of body heat loss is through vasodilation and convection through blood flow to the skin surface and transfer of heat to the environment. Due to the lack of natural convection in spaceflight, an overheated astronaut may experience significant delay in body cooling, leading to heat stress, fatigue, and other sequelae of heat illness. During EVA, astronauts wear cooling undergarments that circulate cooled water over the astronaut's skin, allowing convective heat transfer from the body to the cooling garment. Even with this mitigation in place, crewmembers must take care to ensure that they are allowing themselves sufficient time to cool down after particularly demanding activity, such as EVA or even daily exercise activity in the cabin, and remain vigilant for evidence of thermal stress. Some evidence suggests that body heat dissipation after long-duration flight remains impaired after landing, such that astronauts must remain wary of heat stress even after they return to Earth.[64]

Within the cabin, the environment can reach temperature extremes that challenge the crew's health. On the Mir space station, failures of the cooling and electrical systems in 1997 led to internal station temperatures of greater than 40° C (104° F).[198] After the Apollo 13 command module lost power, temperatures as low as 3° C (37.4° F) were recorded in the cabin.[18] In another example, in 1985, a Soyuz crew was launched to reclaim and repair the Salyut-7 space station, which had experienced a power failure earlier that year. The electrical systems had been nonfunctional for months before the rescue mission, and the cosmonauts who boarded the derelict station entered one of the most extreme environments encountered in spaceflight. Wearing winter coats, hats, and gloves, the crew worked in subfreezing temperatures, estimated at −10° C (14° F), to repair the station successfully.

CARBON MONOXIDE

Carbon monoxide (CO) formed from the catabolism of hemoglobin is another metabolite that can normally accumulate in spacecraft, at a rate of 32 mg per crewmember per day.[206] Most sources agree that CO levels below 70 parts per million (ppm) would not often cause symptoms; however, in the enclosed environment of a spacecraft, levels could rise rapidly if removal systems failed, and pockets of poorly circulated air may result in areas of higher CO concentration, as occurs with CO_2 buildup. The rate of CO accumulation would depend on the size of the spacecraft and efficiency of the CO removal mechanisms. For example, if CO is not actively removed, CO levels could rise to about 20 ppm in 20 days with three crewmembers in a habitable volume of 100 m^3 (3500 $feet^3$).[101] Current spacecraft utilize catalytic oxidizers for CO removal; under normal circumstances, CO accumulation is generally not a significant concern. However, CO release as a byproduct of fire is a greater risk, as discussed later.

PROPELLANTS AND COOLANTS

Accidental release of onboard propellants and coolants could pose a significant risk to astronauts. The most common propellants used in spaceflight, nitrogen tetroxide as an oxidizer and monomethyl hydrazine as fuel, are extremely hazardous. To mitigate the risk to crewmembers, these compounds are usually contained in tanks located outside the habitable volume of the vehicle (Figure 126-4). Unfortunately, leaks occur. During the Apollo-Soyuz mission in 1975, the three-member U.S. crew was exposed to 250 ppm of nitrogen tetroxide for about 5 minutes during reentry, when a valve misconfiguration allowed the gas to enter the crew cabin. Nitrogen tetroxide decomposes into nitric acid on contact with mucous membranes and is a potent irritant, causing damage to mucosal membranes and lung tissue. The crewmembers began to suffer from a burning sensation in their eyes, throats, and lungs immediately after exposure and developed significant pulmonary edema shortly after landing, requiring hospitalization and observation. Fortunately, all three crewmembers recovered completely.[51,199]

FIGURE 126-4 Emissions visible from the Space Shuttle's orbital maneuvering system during an orbital correction. Once in orbit, nitrogen tetroxide and monomethyl hydrazine are often used to provide thrust to spacecraft.

Hydrazine can be similarly irritating to mucous membranes. Significant exposure can also lead to nervous system injury, causing seizures, coma, and eventually death. Furthermore, hydrazine can cause damage to other major organs, including liver, spleen, and thyroid. Exposure to hydrazine would be of great concern during prelaunch fueling or a postlanding fuel dump. Typically, vehicles dump unused fuel shortly before or after landing; this could pose a concern to a crewmember or ground-recovery worker who might be exposed. Halocarbons are another class of potential contaminants, particularly because they are used inside the crew cabin as cleaning agents, coolants, and fire suppressants. Concentrations above 1% are known to cause cardiac irritation that can lead to dysrhythmias; these levels have not been seen in spacecraft.[101] Ethylene glycol, used in the internal cooling system of the Apollo spacecraft and Mir space station, leaked into the cabin atmosphere during a Mir mission in 1997, causing ocular and respiratory irritation in the crew. Fortunately, inhaled ethylene glycol does not pose the high risk of systemic toxicity seen with oral ingestion,[101] and the crew experienced no systemic effects.

The ISS uses liquid ammonia, pressurized in coolant lines that run outside the station, to transfer heat to space. When leaked, ammonia forms snow-like crystals near the point of leakage. During EVA, when working on these lines, astronauts may be exposed to crystallized ammonia and carry the compound on their suits into the cabin, contaminating the internal atmosphere of the airlock. If a crewmember were to come in direct contact with ammonia, exposure as low as 20 ppm would be immediately noticeable and irritating to the eyes and mucous membranes, and inhalations of 150 ppm could cause inflammation of the upper airway. Higher levels of exposure can cause tracheobronchial burns, pulmonary edema, and even death.[241] Given the risks associated with ammonia exposure, several steps are taken to prevent contaminating the internal atmosphere of the ISS after EVA. Exposing a suit to sunlight during EVA, a technique known as a "bake-out," can accelerate sublimation of any ammonia contamination. Bake-outs are required before airlock ingress any time suit contamination is suspected. In addition, colorimetric testing kits reside in the airlock to detect the presence of residual ammonia in the air after repressurization. If contamination is discovered, multiple atmosphere purges and repressurizations of the airlock may be required to dump contaminants fully before returning the crew into the habitable volume of the ISS.

FIRE

Fires are quite possible onboard spacecraft; three occurred during the Space Shuttle program from electrical arcing and combustion of wiring insulation or electrical components. Fires within the closed volume of a spacecraft can be dangerous, not only because of the flames but also because toxic pyrolytic products, such as hydrogen chloride, hydrogen cyanide, and CO, are released into the habitable volume. The Space Shuttle events produced minor combustion products easily scrubbed from the atmosphere, but more significant fires have occurred. In 1997, a fire broke out aboard the Mir space station when a solid-fuel O_2 generator (frequently used at that time to add O_2 to the environment) malfunctioned, leading to an uncontrolled and intense flame. The flames caused damage to station panels and internal structures and released thick clouds of smoke into the cabin before the crewmembers were able to extinguish the fire. This event almost led to evacuation of the station and highlighted the dangers of an uncontrolled flame in the cabin. During a separate event aboard the Mir station, paper filters inside an overheated catalytic oxidizer ignited, releasing a significant amount of CO, with measured levels as high as 400 ppm. At least one crewmember reported headache and nausea from that exposure.[101]

To minimize the risk of fire, O_2 levels are kept at the nominal constituent level of 21%. Materials onboard the ISS are carefully selected whenever possible to be flame retardant and to minimize toxic off-gassing. An example is minimizing plastics that off-gas high levels of cyanide. Care is taken to ensure that crewmembers always have immediate and unhindered access to nearby supplemental O_2 stores and fire suppressants, with a clear path to the escape vehicles. If a fire occurs and is controllable, clearance of pyrolytic contaminants is achieved with airflow and filtering. Combustion products typically absorb quickly into water vapor; thus, condensers in air-conditioning systems can assist in removing these compounds. Lastly, the crew has access to colorimetric sensors for real-time detection of atmospheric contaminants and to ensure that the air has been adequately purified.

WATER

Historically, such as during the Mercury and Gemini missions, water for drinking and hygiene has been provided to crews from storage tanks or from water generated as a byproduct of the electrochemical combination of hydrogen and O_2 in fuel cells, as occurred on the Apollo spacecraft and the Space Shuttle. At present, the ISS Water Recovery System (WRS) recycles potable water from urine, cabin humidity condensate, and spacesuit wastewater. Water from the ISS WRS is also fed into an oxygen generation assembly that produces O_2 and hydrogen through electrolysis. The O_2 created is released into the cabin atmosphere; the remaining hydrogen is combined with CO_2 in a separate system, which can then produce more water for consumption. This highly efficient generation of water reduces the weight of stowed water and consumables that must be launched and delivered to the ISS to support the nominal six-person crew; weight savings have been reported as high as 6804 kg (15,000 lb) per year.

Water contamination can be a significant risk on spacecraft. Waterborne microbes can be transported into spacecraft water systems from terrestrial sources or from crewmembers and crew activities. Potable water is chemically disinfected before crew consumption through a variety of mechanisms. Russian water stores are purified using silver biocide, and U.S. systems typically use iodination. However, use of iodine raises a concern for thyroid dysfunction. To mitigate this risk, a resin filter is used to remove iodine just before water consumption.[160] Despite disinfection efforts, water contamination occurs. *Burkholderia cepacia* and *Staphylococcus aureus* were the most frequently identified organisms in Space Shuttle and ISS water sources.[100,188] Although *B. cepacia* has fairly low virulence, *S. aureus* contamination is a concern because food cooked in *S. aureus*–contaminated water could cause food poisoning. Water spilled aboard spacecraft, if allowed to accumulate, can support microbial life. On the Mir station, condensate accumulations found behind panels after a prolonged period of water system leaks contained stable microbial colonies of algae, bacteria, ciliates, and protozoa.[198,180] No illness has been attributed to microbial contamination in spacecraft water, but prevention of biofilm growth on water system hardware is critical because of the risk and difficulty associated with handling caustic cleaning liquids in microgravity.

Chemical contamination of water systems can occur. As mentioned, the ISS uses condensate recovery and recycling to potable water; atmospheric contaminants could accumulate in condensate and contaminate drinking water. For this reason, certain chemicals, such as isopropyl alcohol, are not used aboard the ISS. In addition, water storage and filtering systems can themselves add contaminants; for example, an early iteration of the iodine removal system led to an accidental ingestion of significant levels of trialkylamines caused by resin leaching. In this case, preflight radiation sterilization of the resin accidentally caused chemical breakdown of the resin material and production of the contaminating byproduct, which leached into the water during filtration.[17]

DUST, PARTICULATES, AND OTHER CONTAMINANTS

Dust and particulates can affect a spacecraft's habitability in the weightless environment, particularly because such particulates, in the absence of gravity, do not settle on surfaces. Flakes of skin, clothing lint, food particles, and particulates from experiments are often found in samples of analyzed spacecraft air. Of particular concern are particles larger than 100 μm, which are the particle sizes most often associated with airborne microorganisms.[100,24] Ventilation systems aboard contemporary vehicles tend to be robust, relying on high-efficiency particulate arrestance (HEPA) filters to maintain a clean environment. In more confined spaces, such as an astronaut's private quarters, crewmembers can often recognize visible levels of particulate matter, particularly skin flakes. Known as *scurf,* skin flakes are released throughout long-duration spaceflight, particularly from skin surfaces that are not used in a microgravity environment, such as calluses and thickened soles on the feet. Another common source of particulate contamination is the lithium hydroxide (LiOH) compound used in CO_2 removal systems. In spacecraft such as the Space Shuttle or Soyuz, which lack the space and power for regenerative CO_2 removal systems, LiOH cartridges are used and replaced frequently, and small LiOH particles are often released during the exchange. Inhalation of LiOH dust can cause self-limited respiratory irritation.[101]

Crews are specifically trained to remove themselves from releases of "noncontainable" particulate flurries and don surgical masks and goggles if returning to work in the area. Also, the risk of particulate contaminant in any newly arriving vehicle is increased; settled tools, screws, and dust from recent manufacturing and stowage operations on Earth can escape detection until they begin to float in microgravity. For this reason, crewmembers wear surgical masks and safety goggles during the first entry into most newly docked cargo vehicles.

Human spaceflight experience to date has revealed other contaminants that can be expected to build up in the closed ecosystem of spacecraft, most often caused by off-gassing of materials. Formaldehyde, halocarbons in hardware cleaning agents, and trace organic compounds are frequently found in air samples in spacecraft, suggesting the need for continued occupational surveillance of air pollutants during long-duration flight.[101] Space station maintenance operations require close proximity of crewmembers to other exotic compounds, including cadmium and nickel (used as anticorrosives in fluid lines) and urea, sulfuric acid, and trivalent chromium (found in waste management systems). Microgravity permits unexpected and ubiquitous sources of exposure; fluids can be free floating, adherent to walls, or hidden behind panels. Thus, accidental contact or even inhalation may be more likely than on Earth, particularly in the close quarters of a spacecraft.

Other environmental considerations will arise with future planetary missions. Close contact with alien soils and dust will be inevitable, so aerospace medicine practitioners will need to anticipate potential hazards from dust that is carried into habitats for study or as a contaminant on space suits and tools. Apollo astronauts reported significant dust contamination of the vehicle cabin after lunar excursions, even noting that the dust seemed to somewhat aerosolize in the low lunar gravity and led to unavoidable inhalation of the particles. In ground studies of animal models, pulmonary inflammation, thickening of alveolar septa, fibrosis, and granulomas have all been demonstrated after lunar dust exposure, although no Apollo astronaut is known to have had such sequelae.[127] Concerns over dust contamination of cabins have led scientists and engineers to develop alternative methods to transfer a surface-walking astronaut into a vehicle, such as "docking" a spacesuit to the vehicle wall and climbing out of the suit into the vehicle, thereby avoiding bringing the dusty suit inside. Creative methods to mitigate particulate exposures will be increasingly important as renewed lunar and then interplanetary travel is realized.

THE RADIATION ENVIRONMENT

Space radiation consists of multiple particle and electromagnetic wave types with a broad spectrum of energies. The solar wind, a continuous efflux of stellar material from the sun and made up primarily of electrons and protons, bathes Earth's magnetic field. Solar particle events, which are jets of high-energy protons and other atomic species, including helium nuclei (alpha particles), also periodically erupt into interplanetary space, occasionally striking the earth and the orbital environment. Galactic cosmic rays (GCR), likely arising from distant supernovas and traveling at relativistic speeds, also contribute to this radiation milieu.

The most common range of orbital altitudes for crewed vehicles is between 240 and 480 km (150 and 300 miles). The ISS orbits between 328 and 432 km (205 and 270 miles) above the surface of the earth. While crews in low Earth orbit (LEO) receive substantially higher radiation doses than occur with terrestrial exposures, the crews still reside well within the Earth's geomagnetosphere (GMS), which deflects much of the solar wind and GCR. The inner and outer Van Allen belts, layered torus-shaped fields of electrons and protons encircling the earth along geomagnetic field lines, represent a further radiation risk to orbiting astronauts (Figure 126-5). The lower margin of the Van Allen belts is closer to Earth above the South Atlantic, because of the offset between Earth's center of mass and its rotational axis, creating a region known as the South Atlantic Anomaly (SAA), where radiation levels are much higher even in LEO. Spaceflight crews in LEO spend only a small fraction (~10%) of travel time traversing the SAA, but incur the majority of their total mission radiation dose there.[192] In addition, captured particles travel along Earth's magnetic field lines and gather closer to Earth's surface at the magnetic poles, creating a region of denser particle concentrations at higher latitudes. Thus, the highest LEO radiation exposures are incurred in polar orbits, where a vehicle travels north to south around the earth, and in orbits that pass through the SAA.

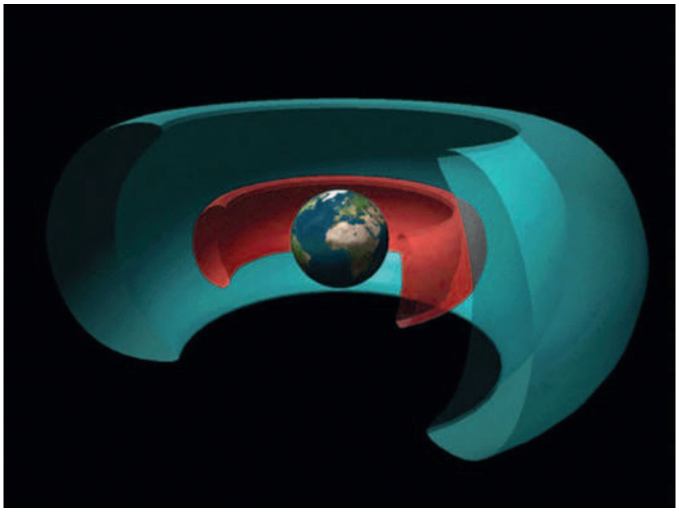


FIGURE 126-5 Artist's rendition of the Van Allen belts, consisting of solar protons and electrons that are trapped by Earth's magnetic field. *(Courtesy NASA.)*

The space radiation environment varies temporally as well as spatially; solar wind intensity and flare frequency wax and wane over an 11-year solar cycle. While increased activity near the peak of the cycle, known as *solar max,* injects more solar electrons and protons into the GMS and increases the radiation burden from these lower-energy particles, it simultaneously expands the solar magnetic field, which better deflects the more highly energetic GCR and lowers the overall radiation dose.[45] Another source of space radiation arises from collisions between GCR and the spacecraft hull, which produce showers of secondary particles. The radiation dose that a crew traveling in LEO can expect to receive therefore largely depends on orbital altitude and inclination, timing of the mission relative to the solar cycle, and the potential for secondary particle creation. Much more distant travel, such as missions to the moon and farther destinations, requires leaving the earth's GMS and exposing crews to the full brunt of solar and GCR, increasing exposure rates to GCR as much as threefold over lower orbital altitudes.[45] *Solar particle events* (SPEs) associated with solar flares, unimpeded by Earth's GMS, may represent the most potent radiation event for a crew traveling outside LEO.

Current spacecraft aluminum hulls provide some protection to crews from space radiation,[192] although the secondary generation of particles from GCR and aluminum collisions adds to the radiation environment inside of the spacecraft. Proton-dense hydrogenated materials, such as water or polyethylene, are excellent absorbers of ionizing radiation and have been considered for use as linings of "safe haven" compartments within interplanetary spacecraft, where the crew could camp for the hours to days required for an SPE to pass.[57] Kevlar can be used for micrometeoroid protection; it demonstrates shielding effectiveness against heavy ions of approximately 80% that of water.[144] However, shielding capability is limited by the weight of effective materials, which may become prohibitive for larger vehicles designed to travel greater distances from Earth. For longer and more distant missions, the most effective means of decreasing exposures would be by limiting transit time through interplanetary space. Electromagnetic or nuclear propulsion engines that provide substantially more thrust over current chemical rocket engines could significantly shorten travel duration through interplanetary space, but this capability depends on advances in rocket propulsion technology.[57]

Radiation Health Effects

Significant ionizing radiation exposure can result in biologic tissue damage and an array of adverse health events on two time scales. Acute radiation effects include burns, vomiting, hemorrhage, and even death, whereas chronic effects include increased mutagenicity, accelerated senescence, and chromosomal alterations causing carcinogenesis and genomic instability.[91] Although the short-term effects largely result from single doses of radiation, long-term issues arise with repetitive and cumulative dosages.

Correlating terrestrial exposures with doses spaceflight crews could experience is problematic because terrestrial exposures consist largely of single-particle types with limited energy spectra, whereas space radiation is a mix of multiparticle species with much higher energies, as well as electromagnetic wave radiation.[56] Because little is known about the biologic effects of many of these exotic particle types, the range of uncertainty in dose calculations can be high. It is expected, however, that acute clinical effects seen with high-dose terrestrial radiation exposure, including nausea, vomiting, fatigue, alterations to blood cell production, and even death, could also occur with high-dose acute space radiation exposures.[94,111] Similarly, long-term effects, such as mutagenesis, are expected to be similar to prolonged accumulation of radiation dosages. In addition, single-cell studies have demonstrated that damage from a single high-energy atomic particle can propagate into long-term effects, such as mutagenesis from chromosomal aberrations caused by a single impacting particle.[65] Other evidence suggests that space radiation may have tissue-specific effects on the ocular lens, bone turnover, and the nervous and cardiovascular systems, in addition to the mutagenic risk.

The health effects of human radiation exposure are commonly measured in sieverts (Sv). For context, an acquired dose of 1 Sv carries with it a 5.5% increased chance of eventually developing cancer. Terrestrial background radiation doses vary with region, but a person living in the U.S. generally receives a dose of approximately 0.0066 millisievert (mSv)/day. In LEO, the typical dose is about 75 times higher, about 0.5 mSv/day.[192] Crewmembers may receive higher exposures, closer to 1 mSv/day, during EVA, simply because of the lack of spacecraft shielding.[192] A crewmember on board the ISS for 6 months is exposed to radiation dosage still well below acute terrestrial dose exposure limits that would drive a measurable increase in cancer incidence.[142] Higher exposures would be expected, of course, with more distant or interplanetary travel. However, little is known about the biologic effects of exposure to high-energy heavy particles associated with unshielded GCR, so making health impact predictions for deep-space missions is difficult.

Cancer risk aside, the only measured pathophysiologic effect in astronauts from radiation exposure is a consistent increase in lymphocyte chromosomal aberrations after long-duration flight, relative to before flight.[72] However, an insufficient population has traveled in space, and insufficient time has passed since the beginning of human spaceflight, to detect an increase in cancer incidence in the astronaut population (Van Baalen M: Personal communication, 2015). Data on doses to astronaut crews have been accumulating since the early days of spaceflight. Crewmembers carry dosimeters on their person throughout their mission, including during EVAs. Other dosimeters are distributed throughout the interior of the ISS to develop spatial and temporal radiation maps of the living quarters. These passive dosimeters can be read and analyzed only after landing; however, a tissue equivalent proportional counter (TEPC) located on board the ISS can be read in real time for particle event monitoring. Other active dosimeters and space weather satellites are available to determine real-time changes in dose in case of unexpected SPEs. This suite of devices significantly adds to the data available to improve the collective knowledge of radiation characteristics in LEO.

Monitoring and Risk Reduction

Because no OSHA limit exists for spaceflight and terrestrial guidelines are too restrictive, the U.S. National Aeronautics and Space Administration (NASA) has sought external guidance from the National Council on Radiation Protection and Measurements (NCRP) and the National Radiation Council (NRC) in determining an acceptable radiation exposure risk level for astronauts, and to recommend means to quantify and mitigate these risks. The NCRP stated that human spaceflight should be conducted while keeping radiation exposure risk “as low as reasonably achievable” (ALARA). Currently, the limit set for astronauts in LEO is a 3% increased incidence (95% confidence interval) of fatal cancer above the background incidence in the normal terrestrial population.[46] Risk reduction currently is largely accomplished through mission planning. Orbital altitude and inclination, mission duration, and timing with the solar cycle can be incorporated into flight planning to reduce exposure. These flight-profile characteristics are largely fixed for operations aboard the ISS, although a few mission parameters can be adjusted to reduce dose. For example, timing spacewalks to occur during orbits that do not transit the SAA can reduce an astronaut’s unshielded exposure to elevated radiation levels. Flight rules have been developed to guide a rapid response to acute radiation events (e.g., SPEs) by terminating an EVA, reorienting the ISS, or directing the crew to cease working and move to a safe haven in a more heavily shielded part of the station.

Limiting an individual astronaut’s time in space is the other means of risk reduction currently practiced. Before assignment of astronauts to a mission, the total radiation burden predicted for that mission is combined with each astronaut’s radiation history from previous occupational, medical, and spaceflight exposures.[111] An individual’s gender and age at first spaceflight are major determinants of susceptibility to mutagenicity from space radiation, because women and younger individuals are more susceptible to effects from identical dosages than men or older individuals.[46] These factors are also incorporated into an astronaut’s risk calculation.

MISSION CONSIDERATIONS

Spaceflight missions vary in time, trajectory, and intent. However, a number of physiologic phenomena are usually experienced during the various phases of flight, including the preflight period, launch, time spent in microgravity, reentry, and landing. These events, as well as strategies to minimize unwanted effects, are described next, specifically with an introduction to the signs and symptoms experienced by the astronaut or witnessed by the aerospace physician.

PREFLIGHT

Space programs usually institute a period of quarantine for about 2 weeks before launch, primarily to prevent infectious illness from affecting the launching crew. A comprehensive quarantine program creates a controlled environment that raises awareness among workforce and family members to limit their contact with the crew and enables a medical screening process for visitors. Children younger than 14 years, who are less able to recognize their own infectious disease symptoms and are more likely to be carriers, constitute the greatest threat of infecting the crew.

A quarantine period has other benefits as well. The final weeks before a spaceflight can be an intensely active time for the astronaut, often requiring last-minute academic preparation and contacts with family, friends, trainers, and program managers. Sequestering the crew in a controlled environment can assist in limiting distractions and enables establishment of a well-balanced work-rest schedule. Furthermore, missions to orbiting destinations, such as the ISS, may require launches at any time of the day to synchronize the trajectories of the launching and destination spacecraft. A controlled lighting environment in a quarantine facility can assist sleep-shifting the crew to help ensure optimal wakefulness at launch. Finally, since crewed spacecraft have very small interiors with primitive waste and hygiene systems, astronauts occasionally desire liquid diets or enemas just before launch, to minimize the need for defecation in the first days of flight, which an adequate quarantine facility and staff can facilitate.

Regardless of precautions taken, an aerospace physician may see a number of physiologic alterations in a crew preparing to launch. Astronauts are often sleep-deprived on the day of launch, if not from the academic workload of preparation for the flight or an extended circadian shift, then certainly from the emotional excitement of anticipating their flight. Wary of inadvertent injury, prelaunch astronauts may limit their exercise intensity and thus may not be at their peak physical condition. They may also be slightly dehydrated, having reduced their oral intake to limit the need for urination or defecation in the first hours of their mission.

LAUNCH

To reach Earth orbit using current chemical propulsion rockets, astronauts nominally undergo acceleration forces of three to four times that of gravity (+3 to 4 G) for approximately 8 to 9 minutes. This acceleration occurs in peaks and valleys to balance human and structural tolerance and fuel and engine performance, eventually attaining a target orbital velocity of approximately 28,000 km/hr (17,500 miles/hr) for the altitudes usually occupied by orbiting spacecraft. While acceleration loads on the order of several Gs may be endured by humans, the human body is particularly vulnerable to excess acceleration in the head-to-toe

FIGURE 126-6 The facial "puffiness" common among astronauts in space is attributed to headward fluid shifts in microgravity, as demonstrated clearly in these preflight and in-flight images of astronaut Chris Cassidy during his 6-month mission aboard the ISS as a member of Expeditions 35 and 36 in 2013.

direction (+Gz). +Gz acceleration results in blood pooling in the lower extremities, which can overcome the cardiovascular system's ability to provide adequate blood flow to the brain and possibly lead to G-induced loss of consciousness. Given these necessary acceleration loads, most spacecraft position crewmembers in a recumbent posture such that the major accelerations of rocket ascent are taken in the more favorable chest-to-back direction (+Gx).

As crews assume a recumbent position in the vehicle on the pad before launch, a cephalad intravascular fluid shift occurs, as it would in any individual in such a position, and the resultant central volume increase triggers atrial and carotid stretch receptors to activate a baroreceptor-mediated decrease in cardiac output and, eventually, inhibition of antidiuretic hormone (ADH) production, ultimately resulting in diuresis. This effect can be dramatic; in studies during the Space Shuttle era, measurements of central venous pressure (CVP) at the cardiac atrium obtained from indwelling venous catheters in three astronauts showed a rise from a standing-position pressure of 5 to 6 cm H_2O to 10 to 12 cm H_2O after the astronauts took their seats before launch.[20]

During launch, crewmembers experience multiple vibroacoustic and acceleration stimuli. The noise and vibrational motion inside the spacecraft, although dramatic, are largely dampened by the crewmember's seat, spacesuit, and helmet, so that astronauts are able to hear radio-voiced commands, speak, monitor instruments, and write on a kneeboard. The overlying sustained 3 to 4 +Gx acceleration to orbital velocity is clearly noticeable, however, in the sensation of heaviness in the limbs and anterior-to-posterior chest pressure. The CVP at this time increases up to 15 to 17 cm H_2O[20] (Video 126-1).

EFFECTS OF MICROGRAVITY

The human cardiovascular system evolved with venous valves and neuroendocrine responses to ensure adequate cerebral blood flow while standing upright under Earth's gravitational influence. Likewise, inner ear otolith organs determine head and body position and sense motion relative to a gravitational field. These physiologic mechanisms remain active when the spacecraft attains LEO, the rocket engines shut down, and the spacecraft begins its free-fall trajectory. Precisely at this moment of engine cutoff, the influence of gravity on the human body vanishes. The astronaut can sense a rapid, remarkable cephalad shift of their intravascular fluid volume, unweighting of limbs, and lifting of intraperitoneal organs, not unlike what one feels in the first few seconds of free-fall on Earth. Crewmembers immediately perceive pressure at the base of the skull and a sense of levitation from the seat as they rise against the pressure of seat restraints. The vestibulo-ocular reflex, which integrates sensed gravitational cues from otolith organs with eye movement and visual input of perceived spatial position, is now essentially uncoupled. Without sensed gravity to help establish a local vertical, crewmembers may feel as if they are hanging upside down, and any head movement can induce a sensation of tumbling.

In these first minutes in orbit, the astronaut may experience sinus congestion, similar to what may be felt with a slight head-down posture on Earth, manifesting in some as a headache. Facial edema, particularly in areas of loose skin such as the eyelids, and an increase in conjunctival injection are often soon visually apparent (Figure 126-6). In fact, forehead soft tissue thickness measured in microgravity with ultrasound showed a 9% increase after several weeks of spaceflight, relative to measurements obtained on Earth.[155] Reductions in leg circumference can be noted after only a few hours in microgravity, and jugular venous distention is also often discernible. Remarkably, CVP decreases to 0 to 3 cm H_2O on arrival into microgravity,[20] unlike the CVP rise associated with the recumbent position on Earth. Decrease in CVP may be caused by a number of mechanisms, including decreased intrathoracic pressure, decreased external cardiac constraint, diuresis, and intravascular volume depletion,[63] as discussed in detail later.

Human physiologic adaptation to weightlessness continues over days and weeks of life in space, as microgravity induces changes in fluid volume and body morphology, and neural and hormonal regulation ensues. These changes can be detected in both short-duration flights of 18 days or less and long-duration flights of 1 month or more. Experience with long-duration flight has led to recognition that these physiologic changes ultimately result in a newly adapted, weightless state of the human organism, a condition explored further in this chapter.

Space Adaptation Syndrome

Motion sickness is one of the most prominent symptom complexes affecting astronauts on entry into microgravity and may occur within minutes of engine shutdown. Conventionally named space adaptation syndrome (SAS), space motion sickness consists of headache, nausea, vomiting, malaise, loss of appetite, and occasionally disorientation. SAS is common and has been studied extensively in attempts to predict susceptibility and determine prophylaxis and treatment measures. The cause of SAS is believed to be a mismatch between movement of the visual surround and neurovestibular stimulation, with headward fluid shifts as a possible contributor. This discordance between visual and neurovestibular inputs is often provoked by head motions and can be attributed to the microgravitational lifting of otoliths away from sensory hair cells in the vestibular labyrinth of the inner ear. As the head moves, the human body expects a resultant activation of the inner ear hair cells by otolithic movement; in microgravity, floating otoliths fail to activate the hair cells, leading to this mismatch and the sequelae of SAS. It has been postulated that cephalad fluid shifts may result in increased intracranial pressure that alters the responses of vestibular receptors, further contributing to the syndrome.[90] Confining launch restraints and the small internal volume of early Mercury, Gemini, and Soyuz spacecraft likely prevented freedom of movement sufficient to elicit symptoms of SAS; thus, the syndrome was not appreciated until the Apollo program. As many as 60% to 80% of all astronauts are affected by some degree of SAS, although the vast majority of cases resolve within 48 hours. Rare cases have been known to persist for up to 14 days before resolution of symptoms.[90] Incidence is decreased for repeat flyers, suggesting either some degree of retained adaptation or use of cognitive strategies to mitigate provocative movements.

Most crewmembers are able to perform operational tasks despite symptoms; early in the Space Shuttle program, only 13% were unable to complete their scheduled work in the first hours of flight because of SAS.[48] Some suggest that adrenergic stimulation from the excitement of the first minutes in space can suppress symptoms.[47] Limiting head movement can also minimize symptoms, but this is often difficult with the required work and occasional unintentional tumbling early in flight before controlled motions are mastered in microgravity.

Thornton and Bonato[227] review the differences between terrestrial and SAS-related motion sickness. Of particular interest is the different character of emesis in space. SAS-induced emesis may be described as “wet burps,” which may reflect the lack of a distinct gastric air-fluid level in microgravity. SAS emesis can also occur without prodrome and with rapid onset. Containing small amounts of vomitus in the microgravity environment is relatively simple with readily accessible gauze-lined emesis bags, which take advantage of the predominance of surface tension forces to wick and trap fluids.

Unfortunately, there are few training-related means of reducing the incidence of SAS. Preflight exposure to cross-coupled angular acceleration stimuli, such as off-axis rotating chairs, have been unsuccessful in preventing SAS.[77] Likewise, bouts of aerobatic and parabolic flight during preflight training have not reduced the incidence of SAS in the astronaut population.[93] Preflight adaptation training by exposing crews to novel gravitoinertial surrounds and stimulations in the laboratory has shown some success in reducing motion sickness during parabolic flight and simulator exposure[29,71] but has not been shown to be successful for SAS prophylaxis.

Pharmaceutical agents remain the most effective treatment for SAS.[47,103] Metoclopramide, scopolamine (oral and transdermal), promethazine (oral, rectal, and intramuscular), and meclizine have all been used as SAS pharmacologic countermeasures. Intramuscular promethazine was the preferred in-flight medication to combat SAS during most Space Shuttle missions and was routinely given before sleep so the sedative side effect would not affect operational performance.[103] Because the onset of SAS occurs shortly after orbital insertion, prophylactic medications are also frequently taken before launch. Dextroamphetamine can be added to promethazine or scopolamine oral preparations to counteract associated drowsiness and sedation, with the added benefit of further antiemetic properties. Of all the medications used to treat SAS at clinically useful doses, ground-based studies showed that meclizine seems to have the least influence on cognitive function, and meclizine taken 12 to 14 hours preflight and repeated 2 hours before launch has recently become the prophylactic pharmacologic countermeasure of choice among U.S. crewmembers to prevent SAS symptoms.[183]

Currently, ISS mission planners allow 48 hours for an adaptation period for SAS resolution before scheduling heavy operational work or mission-critical activities such as spacewalks. SAS is now accepted as an unavoidable consequence of the first few days of microgravity adaptation; work schedule planning and medications to alleviate symptoms are the main means of preventing further mission impact.

POSTFLIGHT

Eventually, humans acclimate to a new functional normal during their time in microgravity, but this transformation results in deconditioning of some physiologic systems required for basic performance on Earth. Returning astronauts may suffer from neurovestibular balance disorders, cardiovascular deconditioning, bone loss, and muscle atrophy. The postflight state is influenced by duration of flight, in-flight countermeasure activities (e.g., exercise), nutritional state, and individual variability. With dedicated countermeasure performance, many long-duration spaceflight crews have demonstrated the ability to function independently immediately after return to Earth, being able to stand, ambulate, and perform light work at the landing site. Others will require significant assistance because of symptomatic orthostasis and disruption of balance. Discussions later in this chapter summarize the physical findings the aerospace practitioner can expect to see in crewmembers immediately on their landing after a long-duration flight.

PHYSIOLOGIC CONCERNS OF SPACEFLIGHT

The following sections delineate many common medical findings in humans after spaceflight, focusing specifically on various physiologic systems and microgravity sequelae. Expression of many of these symptoms is proportional to the duration of flight; longer periods in weightlessness result in more extreme physiologic adaptations. As with many extreme environmental scenarios, the human body adapts, often remarkably, to tolerate and even thrive in the microgravity environment.

CARDIOVASCULAR ISSUES

Headward fluid shifts, tissue unloading, and an almost complete lack of postural or ambulation loading drive the majority of changes in the cardiovascular system in weightlessness. This nullified hydrostatic and hypokinetic state, different from recumbency or even head-down positioning on Earth,[176] is unique in human experience. The cardiovascular system seems to function well with these changes, but the adaptation process has not been completely characterized, and its long-term systemic and end-organ effects have not been fully identified or completely understood.

As previously mentioned, on cessation of acceleration forces at the moment of main engine cutoff, the astronaut’s intravascular fluid redistributes cephalad. Unweighting of tissue structures induces thoracic expansion, which is believed to increase venous compliance, allowing thoracic venous accommodation of the onrush of blood volume (up to 2 L) from the lower extremities and splanchnic vascular beds into the thorax.[20,228] The CVP decrease previously noted is accompanied by cardiac accommodation of the increased fluid load, evidenced by increased venous return, enlarged cardiac chamber volumes, and significant increase in cardiac output.[232] This indicates that despite decreased CVP, a central transmural venous pressure increase occurs in microgravity,[63] most likely caused by decreases in pleural and lung parenchymal pressures and release of pericardial constraint.[84,235] Thus, fluid shifts from the lower to the upper body increase the pressures across the cardiac chamber walls and thus increase cardiac preload. Frank-Starling–induced elevations in stroke volume and cardiac output follow, with cardiac output increases of 15% to 22% maintained 1 week into flight[175] and for up to 40% after 3 months of microgravity.[173] The presence of jugular venous distention in some crewmembers in microgravity despite a reduction in CVP is not completely understood but clearly shows that this clinical sign cannot be relied on as an indicator of right-sided heart dysfunction in the space-adapted individual in the same way as it would be in a terrestrial patient.

Monitoring of blood pressure (BP) and heart rate (HR) in long-duration astronauts on the ISS has shown either no change[95] or decreases of 10 mm Hg or less in mean or diastolic arterial pressures[9,173] compared with preflight recumbent or ambulatory conditions, indicating a drop in peripheral vascular resistance (PVR) to accommodate the increase in cardiac output previously noted.[173] However, unlike in recumbency on Earth, circulating noradrenaline and sympathetic nerve activity remain high in microgravity,[58,59] so this decrease in PVR must be controlled by mechanisms other than a baroreflex-induced decrease in sympathetic nervous activity.[173,174]

Investigations since the Apollo program have established that reduction in intravascular fluid volume stabilizes at approximately a 15% decrease from preflight in approximately 5 days.[4] As a response to the increase in central vascular volume, immediate diuresis would be expected, but this has not been observed in space. Instead, total body water is preserved.[132] This finding suggests that Starling forces induce fluid shifts from intravascular into interstitial and intracellular compartments,[132,176] resulting in a novel cardiovascular euvolemic state in microgravity. Hemoconcentration resulting from this decrease in intravascular water is believed to trigger a process, termed *neocytolysis,* in which newly released red blood cells (RBCs) are selectively removed from the circulation to reestablish a normal hematocrit.[6] Loss of young, larger RBCs results in overall RBC mass reduction that stabilizes at 10% to 15% loss by 1 week into flight.[5] Although astronauts seem to function well in space after these trends stabilize, the adapted state may be inadequate to maintain a normal cardiovascular response to standing in normal Earth gravity on return from spaceflight. Much research has been devoted to this issue of orthostatic intolerance after landing.

As noted earlier, return of gravity can pool blood in the lower body and lead to orthostatic intolerance during and initially after landing. As many as 20% of astronauts after short-duration flight and up to 80% after long-duration flight[161] experience postflight presyncopal symptoms, with a somewhat higher incidence among female astronauts.[233] Low PVR, hypovolemia, and blunted adrenergic response are considered to contribute to this phenomenon.[141]

Lower-body negative pressure (LBNP) techniques, in which the legs and lower torso are inserted into a small vacuum chamber and exposed to a reduced pressure (usually 1 psi below ambient), were used on orbit during the Skylab and Space Shuttle programs as a tool to simulate the orthostatic stress of landing (Figure 126-7). A series of investigations found that a 5-hour session of in-flight LBNP, in combination with hypertonic saline ingestion, maintained HR and systolic BP similar to the preflight response to orthostatic stress.[32] LBNP is considered a possible peripheral neurovascular stimulant that may help prepare crews for return to Earth and resultant hydrostatic stress, but its benefit was found to extinguish within 2 days, leaving astronauts just as susceptible to orthostasis after that period. Given the significant amount of time required to use LBNP as a prelanding countermeasure and the extinguishing benefits, LBNP is not used by U.S. astronauts,[31] although the Russian program has continued to use LBNP exposure for its cosmonauts in the few weeks before return to Earth.

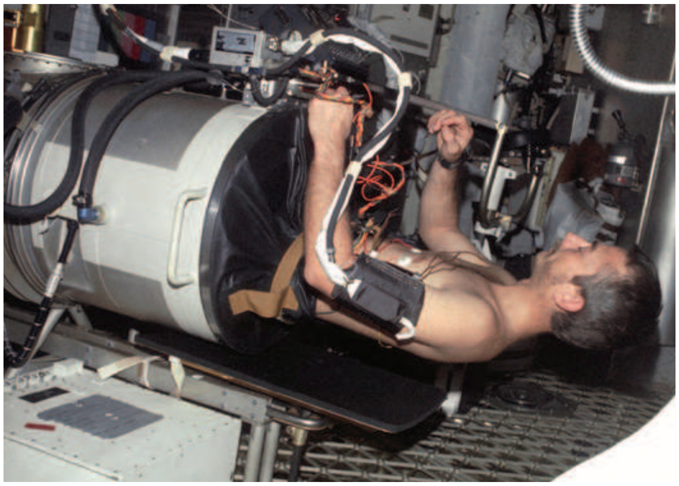


FIGURE 126-7 Astronaut Owen Garriott using the lower-body negative pressure (LBNP) device in an experimental protocol aboard the Skylab Space Station during the Skylab 3 mission in 1973.

Other means of reducing postflight orthostatic intolerance have been adopted. An inflatable lower-body antigravity (anti-G) suit was used during Space Shuttle launch and landing to prevent blood from pooling in the lower body during +Gz (head-to-toe) acceleration exposures. This was particularly relevant to Space Shuttle crewmembers who were returned from short-duration flight in an upright seated position, rather than in the recumbent position in U.S. and Russian capsules. Arterial pressure was higher in astronauts who inflated their anti-G suit during acceleration exposure than in those who did not inflate them.[184] In addition, a liquid cooling garment was worn inside later Space Shuttle reentry suits to reduce thermal stress and allow astronauts to regulate their individual body temperature, which improved overall tolerance of the physiologic stresses of reentry. Use of this liquid cooling garment, in combination with the anti-G suit, resulted in a significantly lower average HR than in astronauts who did not use either countermeasure.[184] In the Russian space program, Soyuz crewmembers currently wear an elastic compression garment during reentry and on the days immediately after landing, to provide venous compression and improve cardiac return from the extremities. A graded compression garment has also been developed to apply progressive pressures from the feet to the abdomen for use after long-duration and exploration missions; it has been shown to prevent tachycardia and to increase total peripheral resistance when standing after short-duration spaceflight.[216]

Crewmembers return to Earth relatively dehydrated compared to the terrestrial norm. “Prehydration,” or fluid loading, with oral water and salt in the hours before return to Earth is another effective countermeasure against orthostasis.[21] Alfrey and colleagues[5] described a 27% decline in stroke volume and 14% drop in cardiac output in astronauts who did not fluid load before reentry. Effective fluid loading countermeasures have been demonstrated to help prevent this reduction of stroke volume and cardiac output.

By 2 weeks into a long-duration mission, aerobic capacity has been shown to decrease by 17% from preflight values. Although it can trend toward preflight values with sufficient exercise, aerobic capacity can decrease again to 15% to 22% below preflight baseline after landing.[141,164] With a supervised physical rehabilitation program, aerobic capacity returns to preflight levels by 30 days after landing.[164] Considerable variability exists between crewmembers in alterations of aerobic capacity during and after flight, apparently related to intensity of aerobic exercise while in orbit.[164] Intense exercise countermeasures during flight can effectively prevent postflight aerobic deconditioning. Even with in-flight exercise, diminished aerobic capacity can be expected both during flight and in the first weeks after landing from long-duration flight, most likely as a result of intravascular volume losses, which can take several days to return to normal, and from dilutional anemia that results from the intravascular fluid replenishment, recovery from which can take weeks.[211]

Dysrhythmias

Since the beginning of human space exploration, there has been concern about increased risk of dysrhythmias in microgravity. One instance of paroxysmal supraventricular tachycardia resulted in early termination of a Russian space station mission,[171] and another episode was implicated in termination of a Russian spacewalk.[83]

Short-duration spaceflight has not been shown to predispose crews to clinically significant dysrhythmias,[196,82] although rates of clinically benign ectopy can increase during physiologically demanding activity, such as spacewalks.[83] Preliminary results of Holter monitoring of astronauts on long-duration missions aboard the ISS indicate no evidence of clinically significant electrocardiographic abnormalities[88] (Levine B: Personal communication, 2013), suggesting that spaceflight conditions do not cause an increase in ventricular dysrhythmia risk. In-flight ECG monitoring is currently performed in the U.S. program only during exercise testing and spacewalks, as an objective measure of activity level during these events.

Cardiovascular Fitness

The potential for cardiac atrophy from lack of physical demand in microgravity has been a concern. Decreases in cardiac mass measured with magnetic resonance imaging (MRI) in astronauts after short-duration spaceflight and in terrestrial bed-rested subjects[185] suggested that significant cardiac atrophy could occur after 1 week or more of exposure to microgravity. However, rapid recovery of left ventricular (LV) mass to preflight values, detected in crewmembers after 4 days of short-duration flight, suggests that LV mass changes can be explained by dehydration or microgravity-induced fluid shifts alone.[223]

Reduction in postflight aerobic fitness has been noted since the Gemini program, with decreases in peak O_2 consumption ($\dot{V}O_2max$) of up to 22% relative to preflight values.[30,141] Generally, crewmembers experience a 15% to 20% decrease in $\dot{V}O_2max$ over the first month of flight, which can improve over several months with adequate access to in-flight aerobic exercise, approaching preflight levels in some individuals. However, postflight $\dot{V}O_2max$ usually remains below preflight levels; after a 6-month mission, average $\dot{V}O_2max$ significantly decreased to 10% to 15% below preflight values.[66,164]

Aerobic exercise aboard the ISS is provided by using a cycle ergometer and a treadmill with a load-bearing harness (Figure 126-8 and Video 126-2). Thirty minutes of aerobic exercise with protocols that generate O_2 consumption at 75% of a crewmember's $\dot{V}O_2max$ are scheduled 6 days per week on the ISS. Participation in robust aerobic exercise during spaceflight has also been increasingly recognized as a mitigator of postflight orthostasis after even short-duration exposure to microgravity.[137]

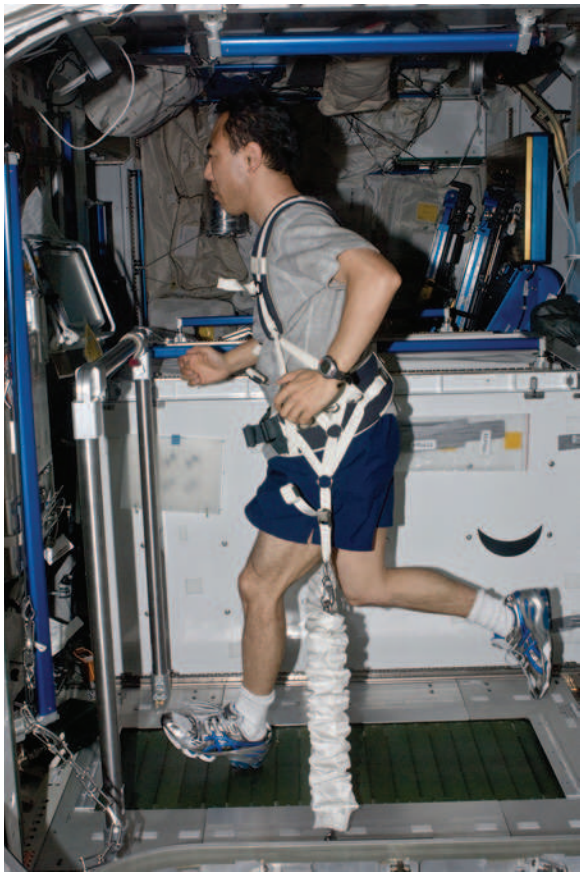


FIGURE 126-8 Astronaut Satoshi Furukawa using the T2, a vibration-isolated treadmill used extensively for exercise aboard the ISS. Note the white loading harness and bungees that pull the astronaut onto the treadmill with a force equivalent to up to 80% of the astronaut's body weight.

Monitoring and Treatment

Screening potential spaceflight crewmembers for cardiac disease is the most effective means of preventing cardiac issues from having an impact on space missions. Hypertension, cardiac structural abnormalities, evidence of inflammation of the myocardium or pericardium, cardiac tumors, and symptomatic atherosclerotic disease are considered disqualifying for selection as a career astronaut. Tachydysrhythmias and conduction defects can be disqualifying, depending on symptoms, persistence, and associated underlying conditions. Catheterizations are not routinely used because of cost and risk considerations, but a coronary artery calcium score (CACS) determined by electron-beam computed tomography (CT) scanning, in conjunction with a Framingham risk evaluation, is currently used to stratify astronauts into low-, medium-, or high-risk categories.

After selection, a program of exercise, nutrition counseling, and annual monitoring is followed to detect, prevent, and delay progression of atherosclerotic disease in susceptible individuals. Members of the astronaut corps are generally fit, but for some astronauts, age alone constitutes an independent risk factor; average age in the astronaut corps was 44 years for women and 45 years for men in 2002.[84] When disease resulting in flight disqualification is identified in an active astronaut, definitive treatment can be appropriate in select cases and can result in return-to-flight status. For example, radioablation of ectopic foci and accessory conductive tracts has been performed in astronauts diagnosed with new-onset atrial fibrillation and resulted in successful subsequent spaceflights. In addition, detecting early stages and progression of atherosclerotic disease has been a major component of an astronaut's annual physical examination. Risk can be decreased by reducing low-density lipoprotein (LDL) to less than 100 mg/dL by means of diet and medications and by increasing $\dot{V}O_2max$ as determined by exercise stress testing. While hypertension is considered disqualifying on selection, single-agent antihypertension control can be instituted in an already-qualified astronaut. Risk of a life-threatening cardiac event during a mission still exists because no prevention program will be perfectly sensitive to detecting and mitigating cardiac disease.

Future Research

The extent to which redistribution of fluid within body compartments during spaceflight affects long-term human health is unknown. It has become apparent that fluid shifts and tissue unweighting in microgravity impose a profound change on cardiovascular, neuroendocrine, musculoskeletal, and hematologic systems. There is further evidence for effects even at the cellular level, in size, contour, shape, and distribution of renal microtubules and changes in mechanotransduction of various cell types.[172] A recently discovered phenomenon, known as visual impairment and intracranial pressure, may result from the unique vascular response to the fluid shifts and tissue compression changes associated with microgravity. Likewise, we are only just beginning to understand the effects of ionizing radiation on vascular viability and rate of cardiovascular aging. Investigations are currently underway to detect evidence of insidious radiation damage to blood vessels that can manifest as intimal thickening, changes in vascular reactivity, or increases in O_2 radicals in blood or urine.[10]

VISUAL IMPAIRMENT/INTRACRANIAL PRESSURE SYNDROME

Concerns related to the effects of long-duration spaceflight on human vision have been expressed for decades. Some crewmembers have historically reported minor changes in visual acuity during or after spaceflight, with symptoms often resolving after return to Earth. Prolonged vision changes that required more significant or chronic corrections have been largely attributed to age-related presbyopia. However, toward the end of the Space Shuttle era, a postflight funduscopic examination and fluorescein angiography of a visually affected long-duration ISS crewmember identified findings that could not be attributed to presbyopia alone.[2] Subsequent medical surveillance since this seminal case has included more detailed eye examinations and advanced imaging technologies to further characterize these findings, including pre- and postflight imaging, regular visual acuity and field testing during flight, ultrasonography of the globe and optic nerve, and similar detailed examinations throughout the pre-, in-, and postflight periods. Findings show that a significant proportion of astronauts exhibit combinations of hyperopic shifts, optic disc edema, posterior globe flattening, optic nerve sheath distention, nerve fiber layer thickening, choroidal folds, cotton-wool spots, and elevated postflight cerebrospinal fluid (CSF) pressure, with variable expression of any given finding.[146]

This constellation of findings was noted to have similarities with idiopathic intracranial hypertension (IIH, formerly "pseudotumor cerebri"), which raised concern for the potential of increased intracranial pressure (ICP) secondary to microgravity exposure leading to ocular and visual alterations.[125] Previous evidence from Russian cosmonauts serving aboard the Mir station suggested that ICP elevation may result from long-duration spaceflight, as indicated by postflight optic disc edema.[166] Intracranial hypertension was thus rapidly identified as a possible causative factor for these ocular structural findings, and a provisional name of *visual impairment/intracranial pressure* (VIIP) was coined to denote this variable yet prevalent phenomenon. Interestingly, not all affected astronauts expressed persistent visual symptoms; some cases resolved on return to Earth, whereas other astronauts retained visual acuity changes or mild elevations of ICP for months to years after landing.[2,146] Because of this variability and the potential for significant disability given the wide range of ocular and neuroanatomic findings, VIIP has become recognized as one of the top health risks of human spaceflight, and substantial resources have been allocated to study the phenomenon.

The etiology of VIIP is unclear, although elevated ICP is thought to be a significant contributor. ICP elevations likely occur secondary to a number of possible factors. The cephalic fluid shift caused by microgravity is thought to cause cerebral venous congestion leading to increased ICP. Intense resistive exercise performed to prevent bone and muscle loss is capable of causing intermittent but potent spikes in intraabdominal and intrathoracic pressures; these intermittent Valsalva-like events may have a smaller but additive effect on ICP.[154] Insufficiency of internal jugular venous valves has been proposed as a factor predisposing to ICP elevation for some groups of terrestrial patients[53,169] and might play a role in ICP alterations in the astronaut population. Elevated concentrations of CO_2 in the vehicle atmosphere may be another contributing factor, because CO_2 is known to cause cerebral vasodilation and subsequently ICP elevation. As previously described, permissible CO_2 concentrations have undergone scrutiny over the years because symptoms suggestive of CO_2 toxicity have been reported at lower concentrations. One investigation also suggests that exposure to chronically elevated CO_2 level during flight may lead to the impaired cerebrovascular CO_2 reactivity seen after landing from long-duration missions.[244] The duration of spaceflight may affect cerebral autoregulation. Cerebral autoregulation in response to LBNP was preserved 2 weeks into short-duration spaceflight[98] and after long-duration flight,[244] but cerebrovascular sensitivity to CO_2 after long-duration spaceflight showed significant reduction in cerebral blood flow response relative to before flight.[244] Increased dietary sodium has also been suggested as a contributor to ICP elevation, precipitating a substantial reduction in allowable sodium levels in the ISS diet.[128] Also considered is the interplay of CSF production and ICP in space. Recently returned astronauts with significant posterior globe flattening had, on average, 70% increased CSF production relative to preflight values, as measured by MRI,[124] suggesting a more nuanced contribution to ICP increases than that caused by venous congestion alone.

Although these clinical and imaging data suggest ICP elevation as a likely contributor for VIIP development,[146,125,124] actual in-flight ICP remains speculative. Ultrasound measurements of optic nerve sheath diameter and globe dimensions show changes that suggest in-flight rise of ICP, but data are not conclusive. Postflight lumbar punctures have been performed on a few affected individuals and showed normal to mildly increased opening pressures.[146] Ocular findings have been out of proportion to these ICP values, but the timing of the lumbar puncture (performed when the busy postflight schedule allowed) may have missed elevations that had since normalized. In-flight lumbar puncture in microgravity has not been performed for many reasons, including sterility concerns and operational risks; even if single ICP values could be obtained from in-flight lumbar puncture, there is no guarantee that measurements would be representative of the ICP during an entire flight duration. Options exist for continuous ICP measurement over several months and involve parenchymal or subdural implantable devices, but these methods carry an additional set of risks.

Fluid shift contribution to increased ICP has been demonstrated in several animal and human spaceflight analog studies. An immediate increase in ICP was seen in rats during analog testing using vacuum drop tests,[78] and a rhesus monkey flown on a Russian biosatellite showed increased ICP within minutes of arriving in microgravity.[229] Elevated human ICP has likewise been seen immediately on head-down tilt,[131] and a recent study showed ICP to be higher both in the supine position and during parabolic flight relative to the upright seated position.[131] Given the inability to "sit" or "stand" in space, the time-averaged ICP in microgravity may be higher than on Earth, where humans spend most of the day upright.

Clinical signs and symptoms of VIIP syndrome vary widely among affected individuals. Interestingly, alterations of ocular anatomy are more likely to occur in the right eye.[146] Female astronauts have presented thus far with milder signs and symptoms than have males,[13] and thus far there have been no reported cases of frank optic disc edema in female astronauts. This gender difference may be related to differences in vascular compliance; overall, male astronauts with VIIP findings have been older than their nonaffected counterparts. Age might contribute to decreased vascular compliance. Affected crewmembers have shown alterations of their folate and vitamin B_{12} one-carbon metabolism pathway, suggesting this trait may predispose a subset of astronauts to developing VIIP.[245] A spectrum of expression of VIIP-related findings suggests biologic variability between individual responses to spaceflight and warrants determination of potential risk factors. A case study of a repeat long-duration astronaut indicated that previous spaceflight experience with ocular abnormalities may be a predictor for developing VIIP findings during a subsequent mission, and repeat flight may have a cumulative effect on the VIIP process.[145]

Medical management of VIIP syndrome is multifactorial, with the goal of preserving vision by managing increased ICP, reducing dietary sodium, avoiding medications or exposures that might increase ICP, and judicious use of medications to lower CSF production. Terrestrially, ICP may be lowered with carbonic anhydrase inhibitors to protect optic nerve function. Although this drug class may have utility in VIIP patients, it could worsen VIIP manifestations by lowering intraocular pressure out of proportion to ICP and thus increase the pressure differential experienced by orbital structures. Corticosteroids may be used in addition to acetazolamide in patients with severe papilledema, although concerns surround the potential for rebound of elevated ICP after cessation. Although not routinely used in terrestrial clinical populations with increased ICP, omeprazole may be considered for treatment of VIIP findings because it is known to decrease CSF production in animal models.[102,143] Medical evacuation from space would be considered in extreme cases, such as impeding retinal compromise or new onset of neurologic findings such as diplopia.

MUSCULOSKELETAL ISSUES

Life and work during a space mission generally demands little from the postural muscles and load-bearing skeleton of the human body. Prolonged, unmitigated exposures to microgravity can therefore result in significant muscle atrophy and bone loss in an otherwise healthy astronaut. Even with regularly scheduled resistive and aerobic exercise, whole-body dual-energy x-ray absorptiometry (DEXA) scans of astronauts returning from long-duration flight in both the Russian and the U.S. space program have shown site-specific areal bone mineral density (aBMD) loss rates of 1% per month in the spine and 1.5% per month at the hip,[129] rates higher than those seen in postmenopausal women.[167,194] A review of the first 23 missions to the ISS showed that, although no astronaut had postflight aBMD values commensurate with osteoporosis, several astronauts evaluated had statistically significant aBMD decrements in the lumbar spine and/or hip (femoral trochanter, femoral neck, or total hip), with some showing aBMD decreases greater than 10% in both areas.[136,179] Recovery to preflight levels of aBMD after return to Earth from the standard 6-month mission in microgravity can take up to 1 year, and 3 years in some cases.[204,207] Removing gravitational compression and postural muscle tension on the human skeleton seems to be the cause of these losses, with studies providing evidence for upregulation of osteoclastic activity in space that outstrips the normally balancing effects of osteoblasts.[225]

Likewise, significant muscle atrophy and functional declines have been measured in the lower back, hips, and lower limbs of individuals after missions longer than 6 weeks. If atrophy is unmitigated, its magnitude appears to correlate with flight duration, associated with a shift of muscle fiber type from slow type I to fast type II phenotypes.[61,76] Muscle volume losses of 4% in the psoas and 17% in the gastrocnemius, as measured by MRI,[134] as well as loss of knee flexion and extension isokinetic torque of 31% and 27%, respectively,[138] have also been observed after long-duration missions.

Atrophy of bone and muscle, which limit function on return to Earth's gravitational environment, can affect in-flight performance. Typical workday activities in space place little demand on the musculoskeletal system, but translating (moving) massive objects across modules, using tools to maintain station systems, and operating mechanisms such as hatches require in-flight maintenance of strength and endurance. Crewmembers must learn to balance any pushing or pulling force with a counterforce in their work, often by the use of postural muscles to brace themselves against a structure. EVAs in particular demand significantly more strength and endurance than usual mission activities, particularly in the upper extremities. For example, two crewmembers conducting a Salyut spacewalk after almost 2 months in space had difficulty opening their hatch and reentering the station airlock.[170] After this event, the Russian program established an upper-body exercise regimen and arm ergometry test that cosmonauts must now pass before EVA.

Various resistive exercise devices have been employed on spacecraft in an attempt to minimize microgravity-induced bone and muscle losses. Sufficiently high loads to the musculoskeletal system, applied through a series of exercises that focus on postural muscle groups and load-bearing skeletal sites, are essential. In one study, a device capable of generating 300 lb-equivalent loads used for 1 hour a day in a variable exercise regimen, including squats, dead lifts, and heel raises, neither prevented loss of muscle mass or decreases in force or power,[76,230] nor preserved bone.[179] A more recent device, launched to the ISS in 2008, uses adjustable piston-driven vacuum cylinders and an inertial flywheel to simulate the constant mass and inertia of free weights on Earth. The Advanced Resistive Exercise Device (ARED) delivers loads ranging from 20 to 600 lb equivalent, and its daily use on board during 6-month missions has demonstrated the capability to reduce aBMD loss significantly in the pelvis, hip, and trochanter, as well as preserve lean tissue mass[210,212] (Video 126-3).

Preservation of bone and muscle also depends on adequate energy intake and vitamin D supplementation.[212] The required daily exercise during a space mission may place a greater metabolic demand on the human body than working and living on Earth. Absence of ultraviolet light exposure during spaceflight in conjunction with decreased vitamin D intake can significantly reduce body vitamin D stores, further adding to bone loss. Stein and colleagues[215] demonstrated that astronauts on short-duration Space Shuttle flights averaged a negative energy balance while in space, consuming only 70% of predicted caloric requirements, likely because of food palatability and compressed schedules that limit time for food preparation. Crews are more likely to attain the recommended caloric and mineral intake with present-day menus and their wider variety of food choices, along with readily available vitamin D supplementation.

The bisphosphonate drug alendronate has also been evaluated during spaceflight as a potential pharmacologic adjunct to exercise in preventing bone loss. In a recent study, exercise in addition to alendronate attenuated all indices of bone loss, including aBMD, trabecular alterations, serum markers of bone turnover, and urinary calcium excretion.[135] Because alendronate was evaluated during the time ARED use became routine, the degree of aBMD preservation attributable to the drug alone cannot be ascertained at this time.[179]

Given the benefits of exercise in maintaining musculoskeletal health, exercise hardware systems will continue to be a critical component of future spacecraft. However, exercise hardware development is not without its challenges. One of the most important issues in providing effective exercise countermeasures is the need for vibration isolation of any exercise equipment from its vehicle interface to avoid unwanted mechanical stress on spacecraft structure. Efforts are underway to develop more compact systems that can work within the smaller spacecraft interiors expected for exploration-class spaceflights, and that demonstrate greater reliability for these missions where resupply is impractical, yet still provide musculoskeletal stresses similar to a 1-G environment. Investigations are also underway to develop resistive exercise protocols that require less crew time but maintain bone and muscle fitness, by incorporating more explosive movements of the large muscles of posture and locomotion. Still to be determined are the long-term effects, if any, of radiation on bone remodeling from the higher radiation doses that accrue with flights outside Earth's GMS.[239]

The long-term health effects of microgravity-induced bone loss are unknown, particularly when considered in combination with expected age-related losses. Quantitative CT used to evaluate volumetric BMD in astronauts shows slower recovery rates in the more metabolically active trabecular compartment of bone in astronauts after long-duration flight, despite preservation of aBMD (measured by DEXA), suggesting that bone remodels itself with altered architecture during microgravity exposure.[129,158] It is unclear how this alteration affects fracture risk, or whether it reduces age of onset or characteristics of age-related osteoporosis.[179,207]

PSYCHIATRIC AND BEHAVIORAL HEALTH

Most astronauts find months of life and work in space to be an enriching experience, but there are many stressors inherent to spaceflight that can threaten psychological health of spaceflight crews. Identifying and reducing these stressors are essential to mission success, particularly for deep-space exploration. Significant effort is put toward maintaining behavioral and psychiatric health, supporting crew cohesion and productive crew-ground interactions, and working with ground control teams to minimize crew overwork and circadian dyssynchrony (asynchrony).

Some negative stressors inherent to spaceflight are common to analog environments, including remote polar or subsea stations. Studies of populations deployed to these environments are useful in helping define risks to behavioral and mental health for astronauts who experience similar conditions, and in evaluating methods and products to mitigate these risks. Stressors inherent in both spaceflight and analog missions include confinement, stimulus reduction, social crowding, workload responsibilities, and potential for unhealthy crew social dynamics.[92,182] Another similarity is physical separation of deployed teams from their nuclear social networks, during preparation and for duration of the mission itself. A spaceflight of 6 months is the culmination of 1.5 to 2.5 years of training that includes frequent travel to international training centers. Family crises and world events, completely unrelated to the mission, can affect crews; one cosmonaut withdrew from ground communication for several days on receiving news of a family member's death during a long-duration flight.[35]

Other stressors are found only in spaceflight. Microgravity is both a source of immense fascination for the crew and a complicating element in planning and executing most daily activities. At the same time, the microgravity environment initiates a process of physiologic adaptation with unknown changes in sleep demand and cognitive effort needed to perform several or complex tasks.[92] Other characteristics of the typical spacecraft environment include continuous cabin fan noise, dim lighting conditions, altered circadian cues (including 16 day-night cycles per workday in orbit), variations in air contaminants, and high levels of ionizing radiation exposure.

The spaceflight experience also carries unique positive characteristics, including the qualities of life in microgravity, awareness of membership in an elite group,[221] and, perhaps most important, the visual impact of Earth as seen from orbit[96] (Figure 126-9). Even so, behavioral or psychological challenges can arise during flight. A review of Space Shuttle postmission debriefs conducted from 1981 to 1989 found a total of 34 comments regarding behavioral symptoms, most often anxiety and annoyance;[16] symptoms of stress, tension, mood elevation, and depression were also mentioned in a subsequent review.[208] Crews on Russian and U.S. long-duration missions have shown signs and symptoms associated with negative intracrew interactions, a tendency toward detachment from ground-support teams, symptoms of depression, irritation, boredom, and both overwork and underwork.[37,114]

Psychosocial aspects unique to a long-duration crew aboard the ISS include its small size, with typically only three to six members, its international makeup, and the high levels of education and motivation in each of its members. In addition, the crew in LEO is in near-continuous communication with ground specialists, allowing mission control centers to constantly supervise work activity on board, which can at times undermine crew autonomy or increase workload (Figure 126-10).

Given the demanding schedule and operational requirements of spaceflight activities, disruptions of sleep and sleep quality are known to occur during both short- and long-duration missions. Crews are often required to shift their own circadian rhythms, particularly before dynamic events such as launch, landing, and docking of newly arriving spacecraft. Since 1990, astronauts have undergone scheduled exposures to bright light during preflight quarantine to induce circadian shifts as needed to accommodate launch times or in-flight shift work.[218,238] Sleep-shift protocols that delay, rather than advance, the circadian rhythm are practiced aboard the ISS as needed before major dynamic operations, because most individuals find that delaying their circadian cycle (in other words, going to bed later than usual) is more tolerable than a phase advance (going to bed earlier than normal). "Slam-shifting," a sudden shift in the sleep/wake cycle for one-time critical operations, is sometimes practiced. One study identified that as many as 13% of flight days were spent in slam-shifted cycles.[123] These disruptions of sleep cycles can add to social or operational pressures and further precipitate crew irritation or stress.

FIGURE 126-9 Astronaut Tracy Caldwell-Dyson takes a break from work aboard the ISS to view Earth through the largest windows ever constructed for spacecraft. The view of Earth from space is often cited as one of the most enjoyable aspects of spaceflight.

FIGURE 126-10 Mission Control at the Johnson Space Center in Houston. Flight directors lead a team of engineers and scientists to monitor the ISS and work with the in-flight crew to successfully complete space missions. *(Courtesy NASA.)*

Dedicated efforts are made to reduce the psychological stressors of spaceflight. Selection of astronauts is in part based on their demonstration of qualities suitable for work in the spacecraft environment, including strong teamwork and coping skills, high level of motivation, competency, self-discipline and sensitivity, and history of good social relationships.[181,208] After selection, annual psychiatric and psychological interviews are conducted to confirm fitness for duty. Once astronauts are assigned to a mission, a dedicated behavioral health and performance program provides active support to crewmembers and their families. This focuses on behavioral health, fostering positive crew and family interactions, while monitoring and mitigating disruptions to sleep and circadian rhythms.[15,209] During a mission, astronauts speak privately every 2 weeks with the same psychiatric and psychology professionals with whom they trained before flight. In addition, crewmembers are able to use e-mail and connect to private calls or videoconferences with their families. These are some of the many advantages of near-Earth spaceflights that are capable of real-time voice and video communication.

As with any deployed team, optimizing crew workload is a common challenge on ISS missions. This includes managing the risk of performance being compromised by extended duty days, irregular work schedules, and high workload.[149] In LEO, managing work/rest schedules requires satisfying research and maintenance requirements, planning sleep shifts, and accommodating off-duty time. Preestablished scheduling guidelines include methods of increasing crew autonomy and establishing "payback," utilizing increased time off for more efficient operational activities; this provides a means of managing workload during periods with an accelerated operational tempo. Psychiatrists and psychologists have also derived flight rules defining appropriate sleep-shifting schedules both to prevent accumulation of fatigue and to shift the crew's circadian nadir away from times when they must perform critical activities, such as robotic operations. Sleep aids are available in onboard medical kits to support circadian shifting.[62,190] Symptoms of depression, if detected through crew medical conferences, are generally treated with conservative measures that may include focusing on ensuring adequate rest in the schedule and providing positive feedback to the crew.[62] Antidepressants are available in the space medical supply for severe or prolonged symptoms, but their use has not been required to date. Similarly, antipsychotic and anxiolytic medications are available, but have yet to be needed during spaceflight.[62]

In the future, crews of missions beyond Earth orbit will experience a level of isolation unique in human existence. The distances involved will prevent real-time communications with any terrestrial social network; the round-trip journey for a signal from Earth to the Martian surface could take 45 minutes. The view of Earth, one of the most pleasing aspects of today's missions, will not be present for much of the time during these expeditionary flights. In anticipation of these challenges, current research focuses on further defining personality traits and personal skills that, according to our long-duration experience to date, would be desirable characteristics for future expeditionary crews. Our experience continues to provide lessons in the value of increasing crew autonomy and in ensuring adequate sleep and work/rest balance. Self-assessment tools are being developed to assist remote crews in prevention, assessment, and management of sleep disruption, fatigue, and psychosocial problems.[27] The lessons learned in continued human missions to LEO will be invaluable in designing for the psychological challenges of the longer expeditionary missions of the future.

DERMATOLOGY AND HYGIENE

After a few days in microgravity on orbit, the normal skin desquamation process becomes a fascinating spectacle. In fact, an astronaut recently returned from a long-duration mission can be identified by both the lack of calluses on the foot sole and the addition of a corn on the foot dorsum, formed after months in flight by repeated abrasion of wedging the feet under handrails for positional stability.

Understanding underlying physiologic changes in human skin is largely subjective. Crewmembers often describe delayed healing time for small cuts and abrasions, but this has not been objectively documented. One experiment found delayed epidermal cellular proliferation and loss of elasticity in one crewmember during flight compared with after flight. Coarser skinfolds in magnified images showed changes similar to those in aging skin, and results suggest slowing of epidermal cell turnover from the basal layer to the stratum corneum during flight.[231] The effects of ionizing radiation, immune changes, or microgravity are unclear with respect to these findings.

As discussed earlier, water is a precious commodity during spaceflight. Because of the microgravity challenges of containing free-floating water and detergents, there are no running water showers or handwashing stations in current spacecraft. Personal hygiene consists of using an array of moistened wipes, dry wipes, and towels. Without laundering available, crews simply wear the same clothing on a rotation schedule and discard their clothing after a certain duration of use (every 2 weeks on the ISS). While this might suggest that the ISS would be a rather pungent environment, body odor is generally minimal: clothes remain generally clean because of lack of terrestrial dirt, and in microgravity, without body weight pressing down on socks or underwear, clothes generally "hover" around the body and can remain surprisingly clean despite normal perspiration. Even so, it is worth noting that contact dermatitis, folliculitis, and fungal infections have occurred on long-duration missions. Whether this level of hygiene predisposes astronauts to superficial skin infection is not clear.

Skin trauma is one of the most common injuries reported by astronauts during spaceflight.[202] Although generally minor, skin injuries can affect crewmembers on an almost daily basis. Constant handling of rough materials, such as Nomex fabric or Velcro, will abrade a crewmember's hands and fingers. Minor cuts are common because the hands are so frequently used for mobilization and stabilization. Primary irritant contact dermatitis is also common; EVAs, exercise tests, and scientific payloads require repeated application of electrocardiographic electrodes, the usual cause of this condition. Other dermatologic conditions include tinea pedis and cruris infections.[80]

Skin trauma from working in a spacesuit deserves special attention. A pressurized suit glove acts as a rigid, jointed shell, and pressure points on the hands and elsewhere are common during and after spacewalks.[202,219] This pressure has been known to cause bruising of hand extensor tendons and digital neuropraxia. Strong, continuous fingertip pressure in the suit glove is desired by spacewalkers to enhance tactility throughout arm range of motion; however, this pressure typically causes trauma to the nail bed matrix. Protective measures, such as applied synthetic enamels or cyanoacrylate adhesives, have been tried with variable success. The moist environment of the glove after hours of continuous work and perspiration can contribute to fungal growth, so that onychomycosis may have to be treated in the compromised nail bed.[219]

With the absence of atmospheric filtering, excessive ultraviolet (UV) radiation exposure is another risk inherent to spaceflight. Although most windows aboard spacecraft have coatings to prevent transmission of UV radiation, some are uncoated to fulfill science requirements in Earth and astronomy observations. Skin exposed to solar nonionizing radiation transmitted through non–UV-protected windows can sustain first- and second-degree burns within seconds. As a result, even momentary exposure to light through these windows is avoided.

In-flight treatment for dermatologic problems includes topical corticosteroids and antibiotics, using terrestrial treatment regimens in consultation with ground medical support. Photography is perhaps the most useful diagnostic tool for dermatologic conditions, particularly to assist ground specialists in diagnosing persistent cases.

TRAUMA

Minor trauma, such as contusions, abrasions, and superficial cuts, particularly to the hands as previously noted, is a common occurrence during spaceflight.[202] Activities most often associated with these injuries include interfacing with exercise machines, stowage operations, and translation about the spacecraft. Work in space occasionally requires moving massive objects such as experiment racks, so that more serious crush injuries, particularly to the hands, are possible because of an object's inertial mass. Scalp lacerations, while uncommon, are possible as well from accidental collision against spacecraft structure during crew locomotion, particularly early in a mission before crewmembers have developed their skills in microgravity. Soft tissue trauma also frequently occurs from prolonged work inside the rigid confines of the pressurized spacesuit, which can cause contusions of the hands and feet.

The hardware available to repair lacerations on orbit is similar to that found in the terrestrial clinic or wilderness medical kit. A limited variety of nylon and polypropylene sutures, curved needles, hemostats, forceps, iris scissors, cyanoacrylate tissue adhesive, staples, and bandages are available on board. The main challenge in repairing a wound in space, however, is in setting up the worksite and stabilizing hardware. Maintaining sterile technique is more difficult without the organizing assistance of gravity: Velcro, bungees, magnetic strips, and tape are most often used for this purpose. Wound irrigation is necessary, particularly because of the increased airborne particulate burden inside spacecraft, but irrigation can be challenging in microgravity. Towels or loose-weave gauze can capture fluids effectively in space (but may require some planning).[22]

In vitro studies have shown variable results with respect to integrity of the healed wound in space; abnormal cellular migration and collagen formation, as well as increased inflammation, have been observed.[189] Likewise, aerospace practitioners and astronauts have subjectively noted increased rates of wound infection of up to 50%.[159] The greater challenge with more serious wounds may be to balance complexity of repair with in-flight medical capability and degree of disability if definitive treatment is not rendered in flight. Considerations include limited capability of the medical kit and skill level of the medical officer, who may not be a physician and is likely to have had only a few hours of training. In these situations, photographic documentation and conferences with ground specialists may be essential.

Inertial loads from impacts and sudden motion can be significant, given the speeds crewmembers can attain in translating through the relatively large volume of the ISS and the torque that can be imparted to a single hand or foot acting as a pivot to stabilize the entire mass of the body (Video 126-4). Foot restraints are used throughout the ISS, both inside and outside the spacecraft. Although essential to stabilize oneself for close work, foot restraints can allow transfer of torsional stress, for example, to the knee if only one foot is constrained. Minor knee injuries, such as ligamentous sprains, have occurred. No fractures have occurred in space, but accidental entrapment of a finger or foot between panels or handrails during translation could result in hand or foot fractures. One phalangeal dislocation occurred on Skylab by such a mechanism; fortunately, one of the other crewmembers happened to be a physician and was able to perform rapid relocation of the digit.[38]

Back pain is also a common complaint among crewmembers. In a review by Kerstman and colleagues,[117] back pain was shown to affect 52% of U.S. astronauts during spaceflight. Onset typically occurs on the first flight day, usually resolving by the second day. The pain has features of musculoskeletal or myofascial origin, is typically located in the lumbar region, and is most often mild. Radicular symptoms can also arise, with lancinating pain and patchy anesthesia over the lower extremities. Spinal column lengthening likely contributes to in-flight back pain as a natural consequence of body gravitational unloading.[117,200] One study demonstrated that crewmembers can expect to “grow” 2 to 6 cm (0.8 to 2.36 inches) above their terrestrial preflight height during even short-duration missions, demonstrating significant spinal elongation in microgravity and resultant strain on spinal nerves.[19,220] In general, back pain is alleviated by donning the treadmill harness to impart a compressive load on the spine,[117] or more often by pulling the knees to the chest in a fetal position and stretching the lumbar musculature. Some crewmembers use a strap behind the back and around the knees to maintain this fetal position during sleep.

Herniated nucleus pulposus (HNP) is increasingly recognized as a postflight injury among astronauts, after both short- and long-duration missions.[110] The incidence of both cervical and lumbar HNP is four times higher in astronauts than in terrestrial cohorts in the 12 months after landing.[110] Although the mechanism of this increased incidence is unknown, persistent spinal elongation in microgravity, with intermittent compression from using a harness for treadmill exercise or from loading while performing squats on the ARED, may predispose an astronaut to annulus stress that manifests as postflight herniation.[117] To protect themselves from vertebral disk injuries, astronauts limit heavy load lifting or overaggressive return to nominal exercise regimens for at least 2 weeks after landing.

Shoulder injury constitutes another prominent musculoskeletal risk for astronauts. Spacewalk training in particular is known to be associated with shoulder rotator cuff injuries.[219] Astronauts spend at least 72 hours overall in the spacesuit in an underwater simulated weightless environment before a mission, working continuously for 5 to 6 hours at a time. Training for and executing spacewalks in the U.S. program generally involves repetitive, sustained shoulder joint stress, pulling and pushing the mass of the body and the 250-lb spacesuit during translation. Since the spacesuit is designed around a hard upper-torso component, rotator cuff impingement against the arm opening can occur with arm abduction, flexion, and external rotation (Video 126-5). Rotator cuff tendon strains, subacromial bursitis, and labral tears, usually of the superior labral anterior-to-posterior (SLAP) variety, are relatively common and occur more frequently in astronauts training for and performing EVA[200] (Scheuring R: Personal communication, 2015). In-flight acute shoulder injury is rare but has occurred during the process of spacesuit doffing and during Apollo lunar surface operations after using a drilling tool on the lunar surface.[14]

One of the periods of greatest risk of musculoskeletal injury for astronauts is the preflight period. An injury during this time can have significant impact on the upcoming mission or even prevent an astronaut from flying. In one review of an 8-year span during Space Shuttle operations, astronauts sustained ligamentous sprains and fractures during preflight athletic activity and flight training that resulted in 28 orthopedic surgical procedures during this period. Knee injuries accounted for the majority of the surgical conditions, and running, skiing, and basketball were most frequently associated with injuries.[104] Restricting and monitoring athletic activity, along with improved awareness of the risk of preflight injury, are necessary to prevent impacts of such injuries on space missions.

For finger, wrist, and ankle strains during flight, current spaceflight medical kits contain compression bandages and splinting devices similar to those found in a wilderness medical kit. Cold packs are available, although refrigeration in spacecraft is primarily used to support research and will not likely be available on deep-space exploration missions because of weight and power restrictions. It is not known how bone callous integrity and formation rate in response to a fracture would be affected by microgravity, although investigations in simulated microgravity indicate the potential for alterations to the fracture healing process.[178] A graded loading protocol, such as on the treadmill using an adjustable loading harness, would likely need to be devised for rehabilitation and functional recovery.

For shoulder injuries, in-flight evaluation of soft tissue structures using ultrasound has been performed in space,[60] and nonsteroidal antiinflammatory drugs (NSAIDs) are generally used by crewmembers for exacerbations of a chronic problem in flight. Efforts are currently focused on preflight prevention and early detection and intervention for musculoskeletal and traumatic injuries, in particular shoulder injuries associated with EVA training in water immersion. This includes a consistent exercise regimen to help protect the shoulder through a rotator-strengthening and scapular stabilization program. Still, surgical intervention is often required. Sports medicine specialists are assisting in development of a new spacesuit, which will include new torso/upper arm configurations to allow more freedom of shoulder and upper arm movement and minimize the risk of shoulder injury.

IMMUNOLOGY

Radiation exposure, altered sleep cycles, noise, isolation, allergens, and other terrestrial factors are known immunologic stressors experienced during spaceflight. In addition, microgravity itself directly or indirectly alters the human immune system and may disturb its performance and response to host or environmental pathogens.[40,163,226] Dysregulation of the immune system may increase the risk of contracting infectious diseases during spaceflight. Astronauts and cosmonauts aboard the Mir space station experienced a number of infectious disease events, including conjunctivitis, upper respiratory infections,[156] otitis media/externa, pharyngitis, urinary tract infections, and skin infections. Severe infections or those not responding to antimicrobials may require medical evacuation from LEO. In 1985, one Soviet cosmonaut developed prostatitis while on board the Salyut-7 space station and was medically evacuated after his condition did not improve.[112] Even with all these incidents, no increased incidence of infectious disease has been attributable to spaceflight factors, relative to terrestrial experience.

Alterations of blood hematologic components have been documented in crewmembers and persist for the duration of ISS flights. Decreased resistance to viruses is suggested by cytokine shifts observed in humans and animals during spaceflight. Of note, microgravity alters production of the interferon class of cytokines, reducing the quantity of available defenders to combat invading viruses.[140,214] Furthermore, antiinflammatory cytokines in mice[12] and astronauts[42] have been shown to increase in microgravity; this could weaken the inflammatory response needed to combat immunologic insults. Leukocyte cell distribution shifts have also been observed after spaceflight. Granulocyte percentages from peripheral blood have been shown to increase (although this change is likely caused by demargination with the stress of landing) and lymphocytes to decrease after landing, whereas monocytes remain relatively unchanged when measured after short-duration flights.[43] In multiple studies, T-cell function has been shown to be impaired in both short- and long-duration spaceflight.[41,81,121] The extent of immunologic alterations and amount of time needed to return to baseline after flight depend on the flight duration and accompanying stress levels.[69] Stress hormone elevation during spaceflight may also play a role.

It is uncertain whether allergic or hypersensitivity reactions that have occurred in space result from these immune system changes. Basophil granulocytes contain histamine and have an important role in human allergic response, and the increased granulocyte percentage seen in spaceflight may worsen allergic symptoms. Cosmonauts have demonstrated hypersensitivity to several substances during long-duration spaceflights.[69] Likewise, NASA astronauts have reported numerous atypical allergies and prolonged rashes while on board the ISS. There have been 3.29 cases of skin rashes per person-year of U.S. spaceflight,[97] a rate significantly greater than the incidence of skin rashes in the general terrestrial population (0.044 case/person-year),[118] although the terrestrial data may be limited because these data represent only individuals who sought medical attention for skin complaints.

The risk of infectious disease reflects a balance among host defense strength, host exposure, and the pathogen's capability to evade the host defense. There is some evidence that pathogen virulence is influenced by microgravity-induced alterations in pathogen mechanotransduction. For example, *Salmonella* has been shown to increase in virulence during spaceflight.[172] Larger particulates, such as sloughed skin, offer microorganisms a unique harbor in microgravity. While terrestrial gravity causes these particulates to settle, spaceflight gives free-floating particulates enhanced potential for inhalation, resulting in direct contact with an astronaut's ocular, nasal, and respiratory mucous membranes. Cramped conditions in small vehicles may also lead to transfer of microorganisms between crewmembers.[186,187] Fortunately, geographic isolation of a crew in space from the terrestrial population allows introduction of a new contagious process only when a new crewmember arrives, and virulent pathogens are generally hindered from hitchhiking to space by a strict, immediately preflight quarantine process.

Even with these precautions, in at least one instance, a recently arrived Space Shuttle astronaut disseminated an upper respiratory infection throughout the ISS and Space Shuttle crews. Furthermore, the quarantine process cannot be effective against reactivation of latent viruses. An average of 65% of asymptomatic ISS astronauts reactivated latent varicella-zoster virus, and 82% shed Epstein-Barr virus during their mission.[37,162] It is unknown whether microgravity-induced alterations render immune cells less able to defend against growth of cancer cells induced by the increased radiation exposure outside Earth's atmosphere.

Prevention is the primary space medicine countermeasure against immune disorders. Acquired immunity is enhanced with preflight immunizations against specific disease processes, such as influenza and herpes zoster. Routine sampling of ISS air, water, and surfaces is used to gauge potential pathogenic challenges to crewmembers' host defense systems. Monitoring and early intervention through station cleaning also prevent biofilm propagation across spacecraft surfaces, which could otherwise result in unhygienic conditions during spaceflight. Modern spacecraft engineered with efficient ventilation systems, adequate air filters, microbial limits on launching space vehicles, and strict quarantine before launch help prevent occurrence or transfer of infectious illness in space.

For onboard treatment, ISS medical kits contain standard topical, oral, and injectable medications to combat active infections and hypersensitivity reactions. Ground-based flight surgeons assist crews in selection of a medical regimen and evaluate treatment progression.

Future immune countermeasures may include nutritional supplements. Cervantes and Hong[28] suggested that astronauts may be able to maintain a healthy microbiome by supplementing a well-managed diet with probiotic therapies. Immunotherapies, such as recombinant cytokines, monoclonal antibodies and immunoconjugates, immunomodulators, activated immunocytes, and gene therapy, are other countermeasures that could someday be tailored for spaceflight.

Infectious disease, viral reactivation, hypersensitivity, and increased risk of cancers are the principal immunologic concerns during spaceflight. Immunologically impaired crew health could result in poorly executed tasks, early mission termination, or at worst, loss of crew life. Even a minor illness can result in substantial mission costs, particularly if illness forces an evacuation and early mission termination. Fortunately, the incidence of infectious disease remains low.

UROLOGY

Renal stone formation, urinary retention, and urinary tract infections (UTIs) are the most prevalent genitourinary disorders in spaceflight.[113] Specific stresses associated with spaceflight elevate the risk of urinary problems during flight and in the pre- and postflight periods and contribute to formation of renal stones. Decreased urine output from reduced plasma volume or fluid restriction is known to lead to increased urinary solute concentrations. Fluid restriction can result from limited access to water during launch and landing operations, during spacewalks, and during the compressed timeline of on-orbit operations. Intentional fluid restriction may be practiced by some astronauts to avoid use of primitive hygiene systems (e.g., simple absorbency garments) during launch and spacewalks. Increases in urinary calcium and phosphate from leaching of bone, secondary to skeletal unloading in microgravity, and increased dietary salt also contribute to an elevated risk of forming urinary calculi.[236,237]

Kidney stones can be incapacitating and risk crew health, safety, and mission success. The one reported case of a symptomatic renal stone in spaceflight involved a Russian cosmonaut who experienced severe lower abdominal pain that spontaneously resolved, barely preventing an emergency deorbit.[133] The postflight period may also carry increased risk of kidney stones. Some NASA astronauts experienced kidney stones during the 12 months following their spaceflight. These kidney stones might have developed during spaceflight, or the postflight period may continue to increase the risk of kidney stone formation.

American astronauts are screened for the presence of renal stones with ultrasound imaging both at selection and again on assignment to a long-duration mission, and their stone formation risk is evaluated through measurement of stone risk parameters (urinary analytes, pH, urine volume, and supersaturation of calcium oxalate, calcium phosphate, and uric acid).[112] In an effort to prevent stone formation, astronauts are encouraged to ensure adequate oral intake of water, sufficient to maintain urine output of 2 L/day.[112] In addition, the ISS food system is undergoing reformulation to decrease the sodium content of food items, potentially changing the urinary biochemistry environment to make it less conducive to stone development.

During spaceflight, should symptoms suggestive of acute ureterolithiasis arise, remote guided ultrasound imaging can be used to localize and measure ureteral stone size or to detect the presence of obstruction or alternate diagnoses. Urinalysis using standard reagent sticks can be performed in flight; the presence of blood may assist with diagnosis. For confirmed cases, in-flight kidney stone treatment would mirror standard terrestrial symptomatic care. Tamsulosin, ketorolac, morphine, ondansetron, promethazine, and intravenous (IV) rehydration are available in the current ISS medical kit. Medical evacuation may be required depending on the astronaut's condition, ultrasound characteristics, and associated complications, such as concomitant urinary infection or inability to manage nausea and emesis.

Urinary retention has occurred during both short- and long-duration spaceflight. Intentional voiding delay, possibly because astronauts may prefer not to use absorbent pads during launch, landing, and EVA, may contribute to this incidence. In the close quarters of spacecraft, psychosocial ("shy" bladder) influences may be a contributor because there is often limited privacy for bathroom activities. In addition, anticholinergic and sympathomimetic side effects of medications used during spaceflight (e.g., promethazine, pseudoephedrine) are likely a predominant influence on the frequency of urinary retention. Lastly, absence of gravity may be a separate, unique contribution to urinary retention in spaceflight.[217]

Terrestrially, urinary retention is almost exclusively seen in older men secondary to prostatic hypertrophy; in spaceflight, women represent a larger percentage of this incidence. Most spaceflight cases of urinary retention have occurred in the first 30 days of space missions, implicating SAS (and thus antiemetic medication use and dehydration) as a causative factor. Urinary retention during a mission can generally be treated successfully with a variety of urinary catheters (straight or indwelling) available from the current ISS medical kit. Ultrasound evaluation of bladder urine volume to determine the need for catheterization has been shown to reduce the rate of this procedure terrestrially, reducing bladder or ureteral irritation and nosocomial UTIs; this may be a useful adjunct in spaceflight.[108]

Cystourethritis has occurred during both short- and long-duration spaceflight. Stepaniak and colleagues[217] reported two cases associated with bladder catheterization during shuttle flights. Urinary infection has also been described in one returning Apollo astronaut,[14,112] and one Russian medical evacuation resulted from a case of septic UTI/prostatitis during a long-duration space mission.[112] Currently, in-flight UTI diagnosis is limited to symptoms and simple urinalysis; urine cultures cannot be obtained, and UTIs must be treated empirically. Onboard antibiotic stores are limited, and even a single case of UTI can significantly deplete resources.

GYNECOLOGY AND REPRODUCTIVE ISSUES

Women comprise about 10% of the more than 500 individuals who have flown into space on short- or long-duration missions and have participated in most aspects of spaceflight activity. The flight experience of female astronauts has not revealed any obvious difference in incidence of gynecologic disorders in space or after landing, compared with the terrestrial population.[105,113] However, a 2014 review of the life sciences literature on physiologic adaptation to spaceflight highlights the lack of spaceflight data pertaining to reproductive physiology.[153]

Medical screening criteria for women are the same as for men, with the exception of breast and reproductive system evaluations. Maintenance of women's health before spaceflight follows the same practice as does terrestrial clinical medicine, with the addition of pelvic and abdominal ultrasound examinations every 5 years. On assignment of a female astronaut to a long-duration flight, repeat pelvic ultrasound is required to detect uterine or ovarian abnormalities that may have arisen since the astronaut's selection. During flight, menstrual suppression has been practiced by about one-half of female U.S. astronauts, including use of levonorgestrel implants and oral contraceptives.[99,105] For astronauts who choose to continue menstruating during their mission, standard sanitary products are available, and control and disposal of blood and blood products present no greater challenge in microgravity than do other routine hygiene activities. Although microgravity seems to have no effect on menstrual cycling,[105] the overall picture of spaceflight effects on the hypothalamic-pituitary-ovarian cycle has yet to be fully explored. Microgravity poses a greater challenge for women to maintain hygiene after urination. Toilet designs use controlled airflow in place of gravity to collect and contain urine, and each female astronaut uses a custom-made funnel to interface with the perineum. Surface tension causes urine to collect in the vaginal orifice or remain in the distal urethra in some women, making postvoid drying more difficult. Although this phenomenon is described as a relatively minor annoyance,[105] care must be taken to prevent UTI.

Preventing bone loss during a long-duration spaceflight is of particular interest given the existing risk of bone mineral density loss in postmenopausal women. Fortunately, there appear to be no gender differences in musculoskeletal response to microgravity or in the efficacy of in-flight impact and resistive exercise in preventing bone loss.[210] Adherence to the in-flight fitness regimen, estrogen replacement for postmenopausal astronauts,[39,50] and particular attention to adequate dietary calcium and vitamin D supplementation are necessary to help mitigate potentially additive osteoporotic influences. Bisphosphonates have been shown to be effective as an adjunct to in-flight exercise[135] and are also offered to postmenopausal astronauts to help protect the skeleton in microgravity.

The cardiovascular response to spaceflight shows gender differences. In stand tests conducted after Space Shuttle missions, women had a significantly higher incidence of presyncope than did men, likely from gender differences in cardiovascular response to orthostatic stress. Men tend to respond with a greater increase in PVR, whereas women respond with a greater increase in HR.[68] Recent improvements in maintaining cardiovascular and musculoskeletal fitness during long-duration missions seem to be reducing the incidence of postflight orthostatic intolerance for both men and women.

Radiation exposures are expected to put women at greater risk than men to long-term health. For all ages, ionizing radiation exposure limits for female astronauts are lower than for their male counterparts. The additional cancer risk for the breast and ovaries, the higher lung cancer risk from radiation in women, and other gender-related differences in natural cancer incidence contribute to an overall higher risk of cancer for women exposed to space radiation.[168] Likewise, the longer female life expectancy allows more time for any postflight carcinogenesis to arise.[75] Organ dose equivalents are a few percent higher for women because of their smaller body mass shielding. Female astronauts are therefore at a greater risk of exceeding their career limits for radiation than men are for an equivalent radiation exposure.[45]

Pregnancy is contraindicated for most astronaut training activities and is disqualifying for spaceflight. Flying time in high-performance jets is limited to no later than the first trimester of pregnancy (although most women choose not to fly at all during pregnancy) because of the concern for exposure to high-acceleration forces with potential ejection, possibility of accidental cockpit decompression, hypoxia, and similar risks. Other training risks include rapid pressure changes with standard spacesuit training in vacuum chambers, which includes exposure to 100% O_2, hypercarbia, and depressurization deltas as great as 10 psi from ambient. These and other exposure risks, such as the risks of increased radiation on a developing fetus, require deferral of spaceflight training and flight itself for the pregnant astronaut. Additionally, abnormalities have been demonstrated in neonatal mammals exposed to microgravity during their fetal development, with observable detriments to function of the neurovestibular system.[195]

Spaceflight is not known to affect postflight female fertility. No differences in rates of conception or spontaneous abortion have been detected in female astronauts compared with age-matched controls. Because many female astronauts choose to delay their first pregnancy until completion of selection, training, and their first spaceflight, they often require assisted reproductive technology due to advanced maternal age. Success rates for both natural pregnancy and assisted reproduction in these women are similar to those of nonastronaut cohorts of the same age.[105] For long-duration missions, preflight cryopreservation of embryos or oocytes may be desired, because of the risk in delaying a pregnancy after years of flight and training and the risk of ionizing radiation exposure to reproductive organs.

OPHTHALMOLOGY

Ophthalmologic issues have historically been the predominant reason for disqualification from astronaut selection.[109] Visual acuity, stereopsis, color vision, intraocular pressure (IOP), the fundus, and visual fields[150] are all evaluated in the selection physical examination. However, acuity standards for selection have relaxed since the beginning of the U.S. space program. Currently, visual correction with laser-assisted in situ keratomileusis (LASIK) or photorefractive keratectomy (PRK) is acceptable.[74] Visual correction is common during spaceflight; approximately 80% of U.S. astronauts wear some form of correction. Contact lenses are often worn, and one astronaut has also flown in space with bilateral intraocular lens replacements, demonstrating no instability or visual defects.[147]

The spaceflight environment can be particularly hazardous to the eyes. Microgravity increases the risk of ocular injury from free-floating objects, and increased particulate burden in cabin air raises the incidence of foreign body contamination. Cephalad fluid shifting also contributes to elevated IOP early in flight,[54]

FIGURE 126-11 Astronaut Jean-Francois Clervoy demonstrates an effective spaceflight technique for removing foreign bodies from the eye using potable water. In microgravity, the water forms a dome that bathes the eyes and can in most cases remove offending particulates.

likely from engorgement of the choroid,[148] and is probably a component of the spectrum of physiologic changes that affect visual acuity, retinal vascular diameter, and the optic nerve, the constellation of findings known as the VIIP syndrome (see Visual Impairment/Intracranial Pressure Syndrome, earlier).

Minor ocular injuries resulting in limited, self-resolving ocular irritation are a daily fact of life in space. Crewmembers often find their work or mealtime halted for a few moments to remove a foreign body from the eye or to flush out mild irritants (Figure 126-11). As previously noted, certain activities, particularly entry into a newly arrived spacecraft and packing operations, pose a particular risk for corneal abrasions from atmospheric debris. Crews are required to wear safety glasses during initial entry into newly arrived cargo vehicles, because dust, lint, or metallic debris can easily escape detection during ground cleaning operations and find its way into the open cabin space on arrival to microgravity. The momentum of free-floating particles on impact with the cornea is generally insufficient to embed a foreign body; however, larger free-floating objects, such as tools or writing pens, particularly if tethered to the worksite or crewmember and hovering near the face during work, can pose a risk of eye injury. Corneal abrasions from larger free-floating objects, requiring ocular examination and antibiotic ointment, have occurred several times during flight and usually resolve in 24 to 36 hours. Elastic straps can be particularly hazardous because of the significant potential energy that can be released when they are stretched and snapped back into place. In one crewmember, failure of an exercise cord with elastic snap-back caused traumatic iridocyclitis with photophobia and a decrease in visual acuity in the affected eye, resolving 10 days after the initial injury.

Greater particulate contamination in spacecraft air and difficulty maintaining appropriate hygiene in microgravity without running water likely create increased microbial load to the eye. Sterile infiltrative keratitis has occurred during missions, diagnosed after flight upon discovery of corneal infiltrative scarring. Erythromycin ointment and ciprofloxacin 0.3% solution are available in space medical kits; however, even these medications pose challenges in the microgravity environment. In the absence of gravity, liquid ophthalmologic medications must be placed directly onto the eye so that surface tension can wick the solution onto the conjunctival surface. Therefore, contamination of the dropper tip is likely, and some wasted dropper fluid is expected. Ophthalmologic ointments are preferred for use in space;[74] otherwise, extra liquid preparations are provided to compensate for waste and contamination.

Work aboard the ISS at times requires proximity to caustic materials, such as the sulfuric acid used to treat urine before recycling it into potable water, or lithium hydroxide, a common particulate in CO_2 removal systems. Eye protection is mandatory for procedures involving these systems. If an ocular toxic exposure occurs that requires copious irrigation, an eyewash station using potable water and modified swim goggles can flush potable water across each eye at a rate of 1 L/min[203] (Figure 126-12).

For more complex and serious ocular injuries, communication with the ground is essential. Still images are valuable for ground evaluation of injury severity. Ultrasound is useful in visualizing injuries and can be immediately streamed to the ground to facilitate real-time guidance and diagnostic assistance. Comprehensive ocular examinations using B-mode ultrasonography and Doppler with remote guidance are regularly performed as part of a periodic in-flight ocular examination.[33] Optical coherence tomography, currently aboard the ISS, can also be used to measure the size and depth of corneal foreign bodies, abrasions, and ulcers.

Radiation poses an ophthalmologic risk during spaceflight. Window coatings protect astronauts from the intense solar UV radiation found outside Earth's atmosphere, but as noted earlier, some windows are uncoated to allow transmission of the full electromagnetic spectrum for scientific observations. Exposure of the eye to radiation transmitted through these windows has occurred and quickly results in UV keratitis.[133] "Light flashes," noted by crewmembers in a darkened compartment or when the eyes are closed, occur when energetic particles hit the retina[70] and are reminders of the ionizing radiation environment in space. The GCR component of space radiation has been found to be associated with increased risk of cataracts,[44] although a subsequent 5-year NASA study of lens opacities in astronauts showed no increase in cataract progression with subsequent flight exposure.[34] These findings are relevant to the long-term ophthalmologic health of pilots participating in high-altitude sorties as well

FIGURE 126-12 The emergency eyewash system, demonstrated here by astronaut Tom Marshburn during training. The system can flush the eyes with copious amounts of potable water to remove chemical contaminants, capturing the waste water in a plastic waste bag (not shown).

as of astronauts, particularly those who will crew vehicles traveling beyond LEO on extended exploration missions.[111]

OTOLARYNGOLOGY

Microgravity, cephalic fluid shifts, pressure and PO_2 differentials, and noise exposure are all components of spaceflight that can affect otolaryngologic structures. Nasal congestion is a common occurrence during spaceflight. Mucosal swelling associated with cephalic fluid shifts is believed to be the major contributor to congestion symptoms; pharyngeal mucosal hyperemia has been noted in the normal physical examination in space.[87] Although the cephalic fluid shift can increase the thickness of soft tissue structures above the clavicles, in microgravity there is no posterior displacement of the tongue, soft palate, uvula, or epiglottis, as occurs with the supine position terrestrially. Snoring and signs of obstructive apnea were reduced relative to pre- and postflight measurements and were almost eliminated during in-flight sleep studies conducted on the Space Shuttle,[52] perhaps a welcome finding given that living in close quarters is a common characteristic of life aboard spacecraft.

Nasal congestion can wax and wane in severity during a mission. Symptoms seem to vary depending partly on adequacy of air ventilation in the spacecraft interior, suggesting an allergic component to congestive episodes. Blockage of air ventilation filters has been associated with increased congestion among the crew, and symptom improvement has been noted after filter cleaning.[157] Although symptoms associated with upper respiratory infection (URI) are rare during spaceflight since the establishment of preflight quarantine, they remain the most common otolaryngologic complaint.[3] URI symptoms can be particularly troublesome when compounded with nominal nasopharyngeal congestion. Of note, rhinorrhea is not often seen in space because of the lack of gravity-assisted drainage of nasopharyngeal secretions; this may also exacerbate congestion. Intranasal oxymetazoline, cromolyn, mometasone, and oral antihistamines are available in the onboard medical kits.

Upper respiratory congestion is of particular concern during procedures that require cabin or spacesuit atmospheric pressure changes. Atmospheric pressure in the Space Shuttle and the U.S. airlock on the ISS is decreased during preparations for EVA as part of a DCS prevention protocol. Within the airlock, pressure in the U.S. and Russian spacesuit is reduced from the normal 14.7 psia to 4.2 to 5.8 psia, to allow sufficient flexibility of limbs and hands as the airlock pressure drops for hatch opening into the vacuum of space. Repressurization after an EVA has led to sinus pressure discomfort from blockage of sinus ostia or eustachian tubes in some congested astronauts. EVA crewmembers use yawning, swallowing, and jaw movement to help maintain patency of the eustachian tubes. Valsalva maneuver, if needed, must be done by compressing the nares against a foam block on the inside of the helmet, since EVA crewmembers are unable to access their face with their hands when fully suited. After an EVA, swollen mucosal blockage can again cause problems, and EVA astronauts have reported pain from barotitis media. Prophylactic nasal decongestants (oxymetazoline and/or pseudoephedrine) may be used before EVA as a preventive measure, and symptomatic crewmembers can pause the airlock pressurization process to allow equalization of inner ear pressure when necessary. URIs can exacerbate any of these symptoms, so astronauts with URI symptoms are not allowed to perform EVAs.

Delayed barotitis, also called O_2 otitis, can occur during aviation training or after an EVA as a result of breathing pure O_2. In this condition, O_2 is absorbed in the middle-ear mucosa for up to 24 hours after low-pressure exposure. This absorption creates enough negative middle-ear pressure to cause discomfort if active equilibration is not initiated. Most often, astronauts report delayed barotitis symptoms after a sleeping period when they are not sufficiently awake to perform these equilibration maneuvers.

Otitis externa is occasionally seen with use of earplugs on orbit. If required, antimicrobial ear drops must be administered by inserting a wick or by allowing the dropper tip to touch the external canal skin such that the surface tension allows the drops to adhere to the skin. Dry cabin air and cephalic fluid shifts are believed to predispose crews to epistaxis, another common otolaryngologic malady that occurs during spaceflight.[3] Anterior nasal bleeds can occur spontaneously or after nose blowing, and crewmembers are trained to apply direct external nasal pressure at the onset of epistaxis. If bleeding continues, oxymetazoline, silver nitrate, and lidocaine with epinephrine are available in the ISS medical kit. Posterior epistaxis has not occurred in space, although nasal packing and Foley catheters are available to tamponade a posterior bleed. Sulfamethoxazole-trimethoprim is available as prophylaxis against infection from nasal packing.

Astronauts are exposed to a wide range of vibroacoustic energy throughout a mission, so hearing protection is another aspect of spaceflight preventive care. Astronauts are exposed to significant noise from rocket engines during a launch phase, and high sound pressure levels also occur on board space vehicles in orbit. The main contributors to noise in the orbital phase of flight are internal to the spacecraft and include exercise hardware, experiment payloads, ventilation fans, and pumps. Astronauts living aboard space stations experience less noise than exists in many industrial workplaces, but because they live and work continuously in the same environment for 6 months or more, progressive hearing loss from occupational noise exposure is considered a risk of long-duration spaceflight. Also at risk are speech intelligibility, sleep, and detection of alarms. Space station hardware degradation and refurbishment and the constant addition of new experiment payloads result in a dynamic acoustic environment, which necessitates occupational hearing surveillance for crewmembers and monitoring of the spacecraft sound environment for long-duration missions. Onboard hearing assessments, acoustic dosimeters, and pre- and postflight audiometry are all utilized for hearing surveillance and to identify the need for improved hearing protection. Hearing-protective devices that have been used in spaceflight include foam earplugs, custom-made earplugs, passive headsets, and active noise reduction headsets. Rather than personal hearing protection, however, engineering controls are the preferred means of protecting crews from excessive noise levels.

DENTAL CONCERNS

Dental problems requiring urgent or emergency care pose significant risk in the space environment. Although no specific aspect of spaceflight presents a unique risk for dental issues, the freedom of multidirectional movement possible in spacecraft (and therefore inadvertent impacts between an astronaut and vehicle component) and the tendency for crewmembers to hold small tools or other items in their teeth while using their hands for movement control can increase the risk of tooth fracture. Launch vibrations have been known to dislodge dental crowns, and airlock operations place astronauts at risk for barodontalgia. Apical abscesses, usually from undiagnosed dental caries, can lead to significant and even debilitating pain.

Proper preflight monitoring and dental care are paramount for limiting the risk of mission impact caused by dental problems. Astronaut oral care is focused on prevention. Crewmember dental examinations and prophylaxis are required annually after selection and 1 to 3 months before launch, so that any issues identified can be remedied. During flight, the dental kit includes a dental mirror, explorer, spoon evacuator, and elevator. Dental extraction forceps and anesthetic are also available for intractable tooth pain. Carpules for anesthetic injection are prefilled with bupivacaine for ease of preparation in microgravity. Before launch, crewmembers learn simplified examination techniques and simple dental procedures, including temporary filling application, crown cementation, and dental injections.

GASTROINTESTINAL ISSUES

Once SAS symptoms have resolved, gastrointestinal (GI) function seems to readily adapt to microgravity (see Space Adaptation Syndrome, earlier). However, some mild GI symptoms often arise during space missions. Constipation is a common GI complaint among spaceflight crews early in flight; although the cause is unknown, some degree of decreased gastric motility is likely present in flight.[86,191] Onboard medical kits include fiber supplements and bisacodyl to be used as needed, and constipation usually self-resolves within 1 week of arrival in microgravity.[157] Reflux symptoms are also common and are usually associated with ingestion of a large meal or large bolus of fluid; symptomatic treatments are available to crewmembers as needed.

Because GI diagnostic and surgical capabilities aboard spacecraft are extremely limited, preflight GI abnormalities are identified and addressed before flight. Screening colonoscopy, endoscopy, and abdominal ultrasonographic examinations are performed on astronaut candidates at selection. Abnormalities, such as the presence of gallstones, biliary sludge, or evidence of inflammation in the gallbladder or pancreas, discovered in the intervening period between selection and assignment to a flight, would require aggressive treatment or surgical resolution and a sufficient recovery period before medical approval would be given to return to training or flight. Management of abdominal pain of unknown etiology has generated much discussion among aerospace practitioners and surgical specialists.[11,24] Acute surgical abdomen, such as suspected appendicitis or cholecystitis, is a significant concern. It may be possible to perform some degree of diagnostic imaging for such conditions using the onboard ultrasound, but true suspicion of these conditions would likely result in early termination of the mission and return to Earth. Before evacuation, the mainstay of treatment would be antibiotic control of the infection with the goal of walling off an abscess to prevent rupture until return to Earth for definitive treatment. IV access capability, parenteral fluids, antibiotics, and pain control are available on board, along with sufficient hardware to attempt percutaneous drainage of an abdominal abscess under remote guidance and ultrasonic visualization.[119]

Prophylactic appendectomies or cholecystectomies are performed by some sponsoring agencies before Antarctic and other long-duration deployments. These have not been instituted to date in the U.S. space program,[26] however, in part because of the increased risk of small bowel obstruction from adhesions. The risk of an acute abdomen has been considered sufficiently low given the prescreening mandated before flight, and because the average age of the astronaut corps is higher than the age of peak incidence for appendicitis.[7] This practice may be reconsidered because improvements in laparoscopic technique decrease the terrestrial surgical risk to an otherwise healthy astronaut in training, and as space missions extend deeper into space,[11] an emergency return to Earth becomes less tenable.

NEUROLOGY

The human nervous system undergoes a rapid adaptive process in microgravity and again on return to Earth. Neurovestibular changes, oculomotor function (e.g., eye-hand coordination, gaze tracking), alterations in sensory perception, changes in proprioception, and cognitive effects (e.g., three-dimensional visual perception, mental spatial representation), must all take place for an astronaut to function effectively in microgravity.[193] In microgravity, the influence of otolith inputs on central nervous system (CNS) sensory integration seems to decrease,[115,116,177] resulting in a central prioritization of retinal information over vestibular inputs.[67] At the same time, semicircular canal inputs seem to be unaffected.[115,116] Degradation of proprioception[115,116,177] and remembered limb position,[234] loss of orientation without visual cues,[85] and decrements in visuomotor tracking[122] and dual-task performance[151] have been detected during in-flight investigations.

Given these profound and rapid CNS alterations, it is not surprising that neurovestibular complaints are the most common neurology-associated issues that arise during spaceflight, and their severity seems to be correlated with flight duration.[243] Neurovestibular symptoms are often reported with motion sickness early in flight, contributing to the diagnosis of SAS (described earlier). Symptoms similar to those of SAS, collectively termed *entry adaptation syndrome,* often occur in crewmembers immediately on return to Earth. In addition, astronauts have demonstrated marked ataxia, poor hand-eye control, and neurovestibular confusion with any rapid motion in the early postlanding period. In-flight exercise is increasingly recognized as providing some mitigation of postflight neurovestibular ataxia.[8] A dedicated rehabilitation regimen in the first few weeks after return to Earth is intended to return crewmembers to their preflight fitness levels and to reduce the risk of injury as they assume activities of daily life on Earth. Exercises are tailored to one's individual neurovestibular state, with increasingly challenging exercises that promote multisensory integration.[243] The rehabilitation period also allows the clinician to observe the return of neuromotor function during the readaptation process.

Regarding specific neurologic complaints, headaches occur fairly frequently in flight; up to two-thirds of astronauts have reported at least one headache during their mission, even when they had not suffered from headaches on Earth. Headaches occurring early in flight are likely associated with fluid shifts, SAS, or caffeine withdrawal. The etiology of headaches after the first days of a mission are not as easily discernible, although elevated CO_2 levels have been suggested as an inciting factor in some crewmembers.[130] Cerebral venous hypertension has been postulated as well and may be caused by venous insufficiency similar to that suspected with acute mountain sickness or in patients with idiopathic intracranial hypertension.[240] In-flight medical kits contain acetaminophen and NSAIDs; these medications can be used as needed for headache.

Objective measures of crew performance rarely show decrements aside from those related to fatigue. Crewmembers often note a need for increased focus to prevent mistakes early in flight; fatigue, sensory overload, and other nonspecific stressors likely contribute to this effect, making any specific effects caused by microgravity difficult to extract.[152] CO_2 may be a contributor in some cases; some crewmembers have reported that only after CO_2 levels were reduced for operational reasons did they notice greater mental clarity and ease in performing their daily tasks.[130] Space radiation damage to the CNS is increasingly recognized as a potential risk to astronauts on extended-duration missions beyond LEO. Rodents exposed to heavy ion radiation have shown CNS pathology similar to that seen with aging, including CNS atrophy[222] and decreases in cell division with alterations in behavior.[224]

ONBOARD MEDICAL CAPABILITY

In-flight medical capability has expanded to keep up with the growing demands associated with increasingly complex mission profiles. The current makeup and function of the ISS Health Maintenance System (HMS) distinctly express the aggregate experience of space physiology and space medicine, coupled with operational and technologic advances of recent decades. Similar to wilderness medical supplies, the contents of the HMS are based on predictions of most likely medical threats and the medical expertise of the in-flight crew, and must be reasonable in terms of weight, volume, safety, and compliance with vehicle requirements. In addition to general limitations on mass and volume, constraints exist for certain types of equipment, such as alcohol-containing products (because of their volatility and contamination of condensate reclamation systems) and radiographic equipment (because of power use and electromagnetic interference). Shelf life stability is also a critical factor that must be considered due to cost of launch and resupply logistics.

All ISS crews are trained to respond autonomously to emergency medical conditions, such as anaphylaxis, choking, and cardiopulmonary arrest. Because of significant preflight training time limitations, this training is procedure focused, has a very narrow scope, and is limited to lifesaving functions. ISS crewmembers are typically not physicians, and training for select crew medical officers (CMOs) is limited to less than 40 hours in the 15 months before launch. Computer-based "just-in-time" instructional and ground-based physician coaching are available before medical procedures are performed in orbit.

MEDICAL KITS

Medical supplies that are unique to individual astronauts are flown in a personal ISS Medical Accessory Kit (IMAK). These items could include personal medical devices such as shoe orthotics (for use with treadmill exercise), contact lenses and eye care accessories, and medications used to treat chronic conditions. This kit is launched with the crewmember and is also used to deliver resupply items to the various ISS medical kit packs.

The ISS onboard HMS is a collection of consumable and durable equipment used to monitor crew health as well as respond to illnesses and injuries. ISS medical kits are collections of primarily consumable goods, separated into nine different notebook-like packs, ideal for viewing all contents at a glance and for quick transport to a worksite. The packs are color coded, items are individually restrained, and packs are organized in functional groups according to medical problem, frequency of use, or resupply efficiency. The most frequently used medications are stowed separately for easier access.

Kit contents vary according to their intent and function. Over-the-counter and prescription medications are available; many of the medications are designed for astronaut use at their own discretion, with notification of the ground physician for awareness and resupply needs. Diagnostic equipment, including hemodynamic monitors, otoscopes, and ophthalmoscopes, are also readily available to the crew. Less common and injectable medications, dental cements, urinary catheterization equipment, and wound care equipment are stored separately; crewmembers are required to notify ground physicians when using such equipment. Lastly, there is a separately contained Emergency Medical Treatment Pack with advanced life support consumables, including a bag-valve-mask, intraosseous access kit, and injectable medications for treatment of anaphylaxis and cardiopulmonary arrest. Sufficient supplies, as well as an automated external defibrillator (AED), are available to progress through two rounds of assessment and medication delivery for basic life support and advanced cardiac life support algorithms. There is also a rigid plastic platform, fixed to the cabin deck, with integrated restraint straps to stabilize a patient for medical procedures and transport, similar to a medical backboard. It also provides electrical isolation to protect the CMO(s) and space station avionics during defibrillation (Figure 126-13).

Using injectable medications or fluid rehydration poses challenges unique to the microgravity environment. Although modified off-the-shelf bags of normal saline are available on the ISS, medical officers must ensure adequate air/water separation before injection, generating angular acceleration by spinning the IV bag and driving fluid to the outside of the spin radius. A similar technique must be employed with syringes to minimize infusion of air into IV lines or during intramuscular injections.

DIAGNOSTIC IMAGING

Onboard imaging equipment is limited to ultrasound technique. The ISS ultrasound unit is a lightweight, modified commercial device that is primarily used to support scientific research. A phased-array probe and a curvilinear probe are available for

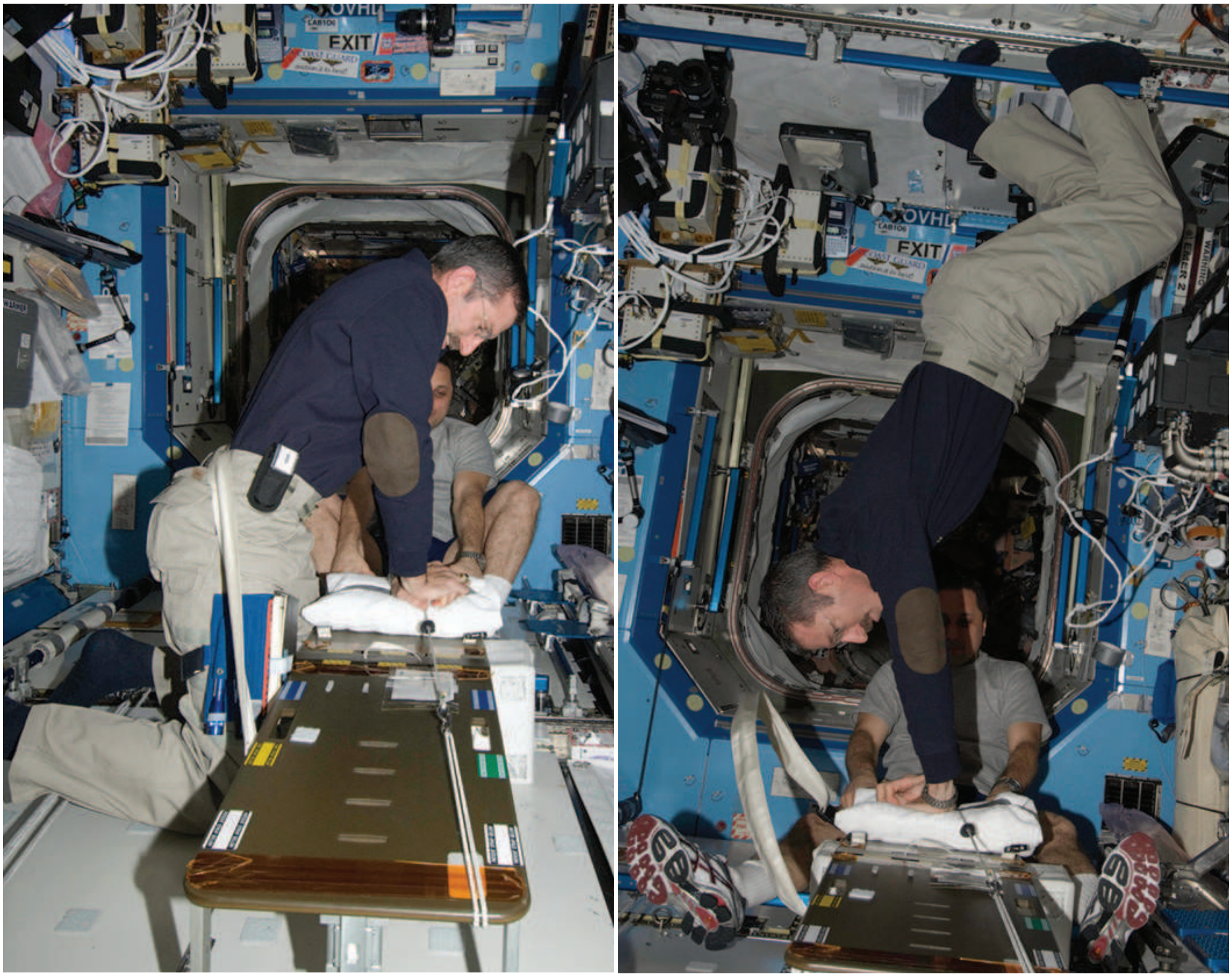


FIGURE 126-13 Astronaut Dan Burbank demonstrates two methods for performing chest compressions during refresher medical training aboard the ISS. The table shown provides patient restraint, a rigid platform for medical interventions as shown here, and electrical isolation from a patient requiring defibrillation.

abdominal, cardiac, thoracic, and transcranial Doppler imaging. High-frequency probes (linear and curvilinear) are available for imaging more superficial structures, such as blood vessels, muscles, and bones. The crew routinely uses ultrasound to perform occupational surveillance of ocular structures, usually guided in real time by imaging experts and ground investigators. Also, the ultrasound and remote guiders are available in the event of a medical contingency to assist with diagnosis and treatment.

Absence of gravity offers both challenges and advantages to ultrasound imaging (Video 126-6). The patient, operator, and hardware must be physically restrained in congruent positions during the exam, typically by using foot restraints. Imaging joints under stress (i.e., valgus/varus) can unintentionally reposition the entire patient rather than the joint alone; occasionally, additional operators, restraints, or ingenuity are needed to ensure patient stability.

Most astronauts undergo preflight photography training. Given the need for photographic documentation of so many aspects of a human spaceflight, high-quality camera equipment is easy to find aboard most spacecraft. Ground-based physicians use downlinked still and video imagery to evaluate skin and mucous membranes, urine reagent stick results, the tympanic membrane, and other visual targets in symptomatic astronauts. Tympanic membrane images obtained with a commercially available otoscope are an important part of the medical examination before and after a spacewalk. Corneal abrasions can also be photo-documented with a high-quality digital camera.

FUTURE CAPABILITIES

Enhanced diagnostic capability that can operate under the constraints of minimal size, weight, power requirements, and complexity will be desired for future generations of spaceflight medical kits. Near-infrared spectroscopy for noninvasive blood chemistry analysis[213] has been investigated for use in spaceflight. Enhancement of surgical care capabilities may entail the ultimate challenge in significantly expanding medical capabilities, because this involves balancing risk and capability against the specific aspects of a spaceflight scenario, such as remoteness, duration, and crew size. Clearly, the significant overhead associated with a terrestrial standard of surgical care, including extensive perioperative imaging, nursing care, delivery of anesthesia, sterility maintenance, and postoperative care,[11] makes application of a terrestrial surgical standard unobtainable in the spacecraft environment. Even so, ideas regarding future surgical capabilities currently drive an integrated assessment of spacecraft design, acceptable medical risk, and requirements for preflight preventive procedures. Evaluations of standard surgical preparation and technique in parabolic flight[23,25] have been performed, highlighting the challenges in maintaining hardware restraint and sterility in microgravity. Interventional radiologic techniques have likewise been considered as a means of further reducing the weight and complexity of hardware required to perform minimally invasive repair in space.[139] The need for a physician in long-duration deep-space missions seems likely, and this physician should have a skill set that includes percutaneous venous access for hydration and medication delivery, ultrasound for diagnosis and procedure guidance, proficiency in treating ophthalmologic and genitourinary problems, and sufficient surgical skills to treat dental problems and repair soft tissue trauma.

SUMMARY AND FUTURE CONSIDERATIONS

Much has been learned in the past 50 years to enable habitation of space, and humans have demonstrated a remarkable ability to adapt, both physiologically and behaviorally, to extended stays in low Earth orbit. Still, much work remains to be done to understand this new frontier, to reduce the risk that the inherent dangers of the spaceflight environment will affect human health, and to maximize human effectiveness. We are still only beginning to understand the space-normal human and time course of each body system as it adapts to microgravity, and how best to arrest processes that trend toward a state that is injurious on return to normal gravity. Although current countermeasures seem to protect the musculoskeletal and cardiovascular systems from the hypokinesis associated with microgravity, we are still unclear on the mechanical integrity of trabecular bone formed in space, and whether a permanent state of increased fracture exists. The pathologic mechanisms and long-term outcomes of VIIP syndrome, the most recently recognized microgravity-induced syndrome, are largely unknown, and knowledge gaps remain regarding individual susceptibility and effective mitigation strategies. The International Space Station, operating in a high-ionizing-radiation environment, offers a platform to better understand effects of a unique radiation population on cells and human physiology. As more astronauts fly, we will better understand the permanency of microgravity- and radiation-induced effects, as well as gender differences in these physiologic changes.

Regarding medical care delivery, knowledge is lacking about the pharmacodynamics, pharmacokinetics, and bioavailability of drugs in the setting of microgravity, particularly given the confounders of body fluid shifts, radiation effects on wound healing, immunosuppression, and lack of sterile capabilities in orbit.[49,55] There is still much to understand about minimizing risk through diagnostic evaluation of individuals during astronaut selection and improved preflight health maintenance programs, as well as determining which medical conditions should truly be considered risky or disqualifying for spaceflight. Commercial access to space will broaden the flying population and include a wider distribution of age and underlying medical conditions. This will improve our understanding of human limits to the stressors of launch, landing, and exposure to microgravity, expanding our collective knowledge of the human body in space.[1,106,107]

The health and function of humans in space will always be intimately tied to the engineering systems that propel and sustain crews, so that the medical risk for missions outside low Earth orbit will largely depend on the capabilities of future spacecraft. While crews will remain dependent on adequacy of rocket and spacecraft integrity to deliver them safely to and from space, advances in propulsion could have the greatest overall effect on medical risk by reducing transit times and therefore minimizing human exposure to ionizing radiation and microgravity. Improved shielding of spacecraft may be necessary to protect crews from the continual background radiation and intermittent solar particle events. As popular media have identified, exposure to the microgravity environment could be further reduced with artificial gravity generation through radial acceleration of rotating vehicles, although this approach carries significant engineering challenges.

As distance from Earth increases, so does the ability to resupply dwindling medical and life support equipment, such that more robust environmental control systems, exercise hardware, better food storage, and improved onboard medical capability will be necessary. Crewmembers will be less able to interact effectively with Earth-based experts and their nuclear social groups. Destinations for longer missions are likely to include the moon, an asteroid, or Mars, where extravehicular activities (spacewalks) are likely to be conducted, and next-generation spacesuits should incorporate ergonomic improvements that cause less injury, particularly to the shoulders, and better accommodate both genders and a wider variety of body types. Neurovestibular compromise on arrival into a new fractional-gravity gradient is still a concern, and methods for adapting crewmembers to gravity gradients will need to be considered. Spacecraft intended for deep-space missions, for the foreseeable future, will be small in size, and onboard medical systems will need to be designed to fit within these limits. The medical system will compete with other components essential to human survival, such as adequate water supply, power to run environmental control systems, and sufficient fuel to reach the destination with a margin for safety. Minimizing the weight and volume of medical hardware, reducing the number of consumable components, or perhaps including components that can be easily manufactured onboard through three-dimensional printing, will be desirable. There will be significant pressure to simplify medical delivery in space.

Our continued experience with long-duration flights will likely uncover other unexpected long-term effects of adaptation to microgravity and contribute further to the basic knowledge of human physiology. Spaceflight programs will continue to be both the recipients of and the contributors to advances in disease detection and health maintenance and to a better understanding of human disease as a whole.

ACKNOWLEDGMENTS

The authors wish to express thanks to the following individuals who provided expert guidance in the writing of this chapter: Dr. Michael Barratt, Dr. Peter Norsk, Dr. Yael Barr, Dr. Jean Sibonga, Dr. Mary Van Baalen, Dr. Gary Beven, Dr. Al Holland, Dr. Richard Scheuring, Dr. John Clark, Robert Pietrzyk, Dr. Brian Crucian, Dr. Richard Jennings, Dr. Serena Aunon, Dr. Richard Danielson, Dr. Robert Gibson, Malinda Hailey, Dr. Ashot Sargsyan, Dr. Mary Wear, Sara Mason, and Jennifer McDonald, as well as the Longitudinal Study of Astronaut Health.

REFERENCES

Complete references used in this text are available online at expertconsult.inkling.com.

REFERENCES

1. Aerospace Medical Association Commercial Spaceflight Working Group. Suborbital commercial spaceflight crewmember medical issues: Position paper concerning medical issues for commercial suborbital spaceflight crewmembers. Aviat Space Environ Med 2011;82:475–84.
2. Alexander DJ, Gibson CR, Hamilton DR, et al. Evidence report: Risk of spaceflight-induced intracranial hypertension and vision alterations. Washington, DC: National Aeronautics and Space Administration; 2012 <http://humanresearchroadmap.nasa.gov/evidence/reports/VIIP.pdf>.
3. Alford BR, Clark JB, Danielson RW. Ear, nose, and throat and auditory issues. In: Risin D, Stepaniak PC, editors. Biomedical results of the Space Shuttle Program, NASA/SP-2013-607. Houston: National Aeronautics and Space Administration; 2013. p. 59–72.
4. Alfrey CP, Driscoll TB, Haley WS, Leach-Huntoon CS. Blood volume and hematopoiesis. In: Space biology and medicine, vol. 3. Reston, Va: American Institute of Aeronautics and Astronautics; 1996. p. 105–15.
5. Alfrey CP, Udden MM, Leach-Huntoon C, et al. Control of red blood cell mass in spaceflight. J Appl Physiol 1996;81:98–104.
6. Alfrey CP, Rice L, Udden MM, Driscoll TB. Neocytolysis: Physiological down-regulator of red-cell mass. Lancet 1997;349:1389–90.
7. Al-Omran M, Mamdani M, McLeod RS. Epidemiologic features of acute appendicitis in Ontario, Canada. Can J Surg 2003;46:263–8.
8. Bacal K, Billica R, Bishop S. Neurovestibular symptoms following space flight. J Vestib Res 2003;13(2–3):93–102.
9. Baevsky RM, Baranov VM, Funtova II, et al. Autonomic cardiovascular and respiratory control during prolonged spaceflights aboard the International Space Station. J Appl Physiol 2007;103:156–61.
10. Baker JE, Moulder JE, Hopewell JW. Radiation as a risk factor for cardiovascular disease. Antioxid Redox Signal 2011;15:1945–56.
11. Ball CG, Kirkpatrick AW, Williams DR, et al. Prophylactic surgery prior to extended-duration space flight: Is the benefit worth the risk? Can J Surg 2012;55:125–31.
12. Baqai FP, Gridley DS, Slater JM, et al. Effects of spaceflight on innate immune function and antioxidant gene expression. J Appl Physiol 2009;106(6):1935–42.
13. Barr Y, Otto C, Brady R, et al. Differential VIIP incidence and severity among male and female astronauts. Aviat Space Environ Med 2014; 85:237.
14. Berry CA. Perspectives on Apollo. In: Johnson RS, Dietlein LF, Berry CA, editors. Biomedical results of Apollo, NASA SP-368. Washington, DC: US Government Printing Office; 1975. p. 581–92.
15. Beven G, Holland A, Sipes W. Psychological support for U.S. astronauts on the International Space Station. Aviat Space Environ Med 2008;79:1124.
16. Billica R. In-flight medical events for US astronauts during the space shuttle program STS-1 through STS-89, April 1981 to January 1998. In: Ball JR, Evans CH Jr, editors. Safe passage: Astronaut care for exploration missions. Washington, DC: National Academy Press; 2001 Presentation to the Institute of Medicine Committee on Creating a Vision for Space Medicine during Travel Beyond Earth Orbit, Johnson Space Center, Houston, 2000.
17. Blue RS, Hudson JC, Rieders MF, et al. Medical management of a potentially toxic accidental trialkylamine ingestion during spaceflight. Aviat Space Environ Med 2013;84:1100–4.
18. Brady JC, Hughes DF, Samonski FH, et al. Apollo command and service module and lunar module environmental control systems. In: Johnson RS, Dietlein LF, Berry CA, editors. Biomedical results of Apollo, NASA SP-368. Washington, DC: US Government Printing Office; 1975. p. 517–43.
19. Brown JW. Crew height measurement, the Apollo-Soyuz test project medical report. In: Report No NASA SP-411. Washington, DC: National Aeronautics and Space Administration; 1977.
20. Buckey JC Jr, Gaffney FA, Lane LD, et al. Central venous pressure in space. J Appl Physiol 1996;81:19–25.
21. Buckey JC, Lane LD, Levine BD, et al. Orthostatic intolerance after spaceflight. J Appl Physiol 1996;81:7–18.
22. Campbell MR, Billica RD, Johnston SL III. Surgical bleeding in microgravity. Surg Gynecol Obstet 1993;177:121–5.
23. Campbell MR, Billica RD. A review of microgravity surgical investigations. Aviat Space Environ Med 1992;63:524–8.
24. Campbell MR, Billica RD. Surgical care in spaceflight. In: Barratt MR, Pool SL, editors. Principles of clinical medicine for space flight. New York: Springer; 2008. p. 227–45.
25. Campbell MR, Dawson DL, Melton S, et al. Surgical instrument restraint in weightlessness. Aviat Space Environ Med 2001;72:871–6.
26. Campbell MR, Johnston SL III, Marshburn T, et al. Nonoperative treatment of suspected appendicitis in remote medical care environments: Implications for future spaceflight medical care. J Am Coll Surg 2004;198:822–30.
27. Carter JA, Buckey JC, Greenhalgh L, et al. An interactive media program for managing psychosocial problems on long-duration spaceflights. Aviat Space Environ Med 2005;76:B213–23.
28. Cervantes JL, Hong BY. Dysbiosis and immune dysregulation in outer space. Int Rev Immunol 2016;35(1):67–82.
29. Cevette MJ, Stepanek J, Cocco D, et al. Oculo-vestibular recoupling using galvanic vestibular stimulation to mitigate simulator sickness. Aviat Space Environ Med 2012;83(6):549–55.
30. Charles JB, Frey MA, Fritsch-Yelle JM, Fortner GW. Cardiovascular and cardiorespiratory function. In: Leach-Huntoon CS, Antipov VV, Girgoriev AI, editors. Humans in space flight, vol. 3. Reston, Va: American Institute of Aeronautics and Astronautics; 1996. p. 63–8.
31. Charles JB, Fritsch-Yelle JM, Whitson PA, et al., editors. Extended Duration Orbiter Medical Project: Final report 1989-1995. Houston: National Aeronautics and Space Administration; 1999. p. 1–19.
32. Charles JB, Lathers CM. Cardiovascular adaptation to spaceflight. J Clin Pharmacol 1991;31:1010–23.
33. Chiao L, Sharipov S, Sargsyan AE, et al. Ocular examination for trauma: Clinical ultrasound aboard the International Space Station. J Trauma 2005;58:885–9.
34. Chylack LT Jr, Peterson LE, Feiveson AH, et al. NASA study of cataract in astronauts. Report 1: Cross-sectional study of the relationship of exposure to space radiation and risk of lens opacity. Radiat Res 2009;172:10–20.
35. Clark J. A flight surgeon's perspective on crew behavior and performance. Presented at Workshop for Space Radiation Collaboration with BHP, CASS presentation, Houston, 2007.
36. Clark JB. Decompression-related disorders: Pressurization systems, barotrauma, and altitude sickness. In: Barratt MR, Pool SL, editors. Principles of clinical medicine for space flight. New York: Springer; 2008. p. 247–72.
37. Cohrs RJ, Mehta SK, Schmid DS, et al. Asymptomatic reactivation and shed of infectious varicella zoster virus in astronauts. J Med Virol 2008;80(6):1116–22.
38. Collins M. Lift-off. New York: Grove Press; 1988. p. 190.
39. Corson SL. Oral contraceptives for the prevention of osteoporosis. J Reprod Med 1993;38:1015–20.
40. Crucian B, Sams C. Immune system dysregulation during spaceflight: Clinical risk for exploration-class missions. J Leukoc Biol 2009;86(5): 1017–18.
41. Crucian BE, Cubbage ML, Sams CF. Altered cytokine production by specific human peripheral blood cell subsets immediately following space flight. J Interferon Cytokine Res 2000;20(6):547–56.
42. Crucian BE, Zwart SR, Mehta S, et al. Plasma cytokine concentrations indicate that in vivo hormonal regulation of immunity is altered during long-duration spaceflight. J Interferon Cytokine Res 2014; 34(10):778–86.
43. Crucian B, Stowe R, Quiriarte H, et al. Monocyte phenotype and cytokine production profiles are dysregulated by short-duration spaceflight. Aviat Space Environ Med 2011;82(9):857–62.
44. Cucinotta FA, Manuel FK, Jones J, et al. Space radiation and cataracts in astronauts. Radiat Res 2001;156(5 Pt 1):460–6.
45. Cucinotta FA, Kim MH, Chappell LJ, Huff JL. How safe is safe enough? Radiation risk for a human mission to Mars. PLoS ONE 2013;8(10):e74988.
46. Cucinotta FA, Kim MY, Chappell LJ. Space radiation cancer risk projections and uncertainties—2010. In: NASA/TP-2011-216155. Houston: NASA Johnson Space Center; 2011.
47. Davis JR, Jennings RT, Beck BG. Comparison of treatment strategies for space motion sickness. Acta Astronaut 1993;29:587–91.
48. Davis JR, Vanderploeg JM, Santy PA, et al. Space motion sickness during 24 flights of the space shuttle. Aviat Space Environ Med 1988;59:1185–9.
49. Dawson DL. On the practicality of emergency surgery during long-duration space missions. Aviat Space Environ Med 2008;79:12–13.
50. DeCherney A. Bone-sparing properties of oral contraceptives. Am J Obstet Gynecol 1996;174:15–20.
51. DeJournette RL. Rocket propellant inhalation in the Apollo-Soyuz astronauts. Radiology 1977;125:21–4.
52. Dijk DJ, Neri DF, Wyatt JK, et al. Sleep, performance, circadian rhythms, and light-dark cycles during two Space Shuttle flights. Am J Physiol Regul Integr Comp Physiol 2001;281:1647–64.
53. Doepp F, Valdueza JM, Schreiber SJ. Incompetence of internal jugular valve in patients with primary exertional headache: A risk factor? Cephalalgia 2008;28:182–5.
54. Draeger J, Schwartz R, Groenhoff S, Stern C. Self-tonometry under microgravity conditions. Clin Invest 1993;71(9):700–3.
55. Drudi L, Ball CG, Kirkpatrick AW, et al. Surgery in space: Where are we at now? Acta Astronaut 2012;79:61–6.
56. Durante M, Cucinotta FA. Physical basis of radiation protection in space travel. Rev Mod Phys 2011;83:1245.
57. Durante M. Space radiation protection: Destination Mars. Life Sci Space Res 2014;1:2–9.

58. Eckberg DL. Neurolab Autonomic Nervous System Team: Bursting into space: Alterations of sympathetic control by space travel. Acta Physiol Scand 2003;177:299–311.
59. Ertl AC, Diedrich A, Biaggioni I, et al. Human muscle sympathetic nerve activity and plasma noradrenaline kinetics in space. J Physiol 2002;538:321–9.
60. Fincke EM, Padalka G, Lee D, et al. Evaluation of shoulder integrity in space: First report of musculoskeletal US on the International Space Station. Radiology 2005;234:319–22.
61. Fitts RH, Trappe SW, Costill DL, et al. Prolonged spaceflight-induced alterations in the structure and function of human skeletal muscle fibres. J Physiol 2010;588:3567–92.
62. Flynn CF. Behavioral health and performance support. In: Barratt MR, Pool SL, editors. Principles of clinical medicine for space flight. New York: Springer; 2008. p. 391–412.
63. Foldager N, Anderson TAE, Jessen FB, et al. Central venous pressure in humans during microgravity. J Appl Physiol 1996;81(1):408–12.
64. Fortney SM, Mikhaylov V, Lee SM, et al. Body temperature and thermoregulation during submaximal exercise after 115-day spaceflight. Aviat Space Environ Med 1998;69(2):137–41.
65. Fournier C, Zahnreich S, Kraft D, et al. Fate of a normal human cell traversed by a single charged particle. Sci Rep 2012;2:643.
66. Fraser KS, Greaves DK, Shoemaker JK, et al. Heart rate and daily physical activity with long-duration habitation of the International Space Station. Aviat Space Environ Med 2012;83:577–84.
67. Friederici AD, Levelt WJ. Resolving perceptual conflicts: The cognitive mechanism of spatial orientation. Aviat Space Environ Med 1987;58(9 Pt 2):A164–9.
68. Fritsch-Yelle JM, Whitson PA, Bondar RL, Brown TE. Subnormal norepinephrine release relates to presyncope in astronauts after spaceflight. J Appl Physiol 1996;81:2134–41.
69. Fuchs BB, Medvedev AE. Countermeasures for ameliorating in-flight immune dysfunction. J Leukoc Biol 1993;54:245–52.
70. Fuglesang C, Narici L, Picozza P, Sannita WG. Phosphenes in low Earth orbit: Survey responses from 59 astronauts. Aviat Space Environ Med 2006;77(4):449–52.
71. Garrick-Bethell I, Jarchow T, Hecht H, Young LR. Vestibular adaptation to centrifugation does not transfer across planes of head rotation. J Vestib Res 2008;18(1):25–37.
72. George K, Rhone J, Beitman A, Cucinotta FA. Cytogenetic damage in the blood lymphocytes of astronauts: Effects of repeat long-duration space missions. Mutat Res 2013;756(1/2):165–9.
73. Gernhardt ML, Dervay JP, Waligora JM, et al. Multisystem issues and countermeasures-extravehicular activities. In: Risin D, Stepaniak PC, editors. Biomedical results of the Space Shuttle Program, NASA/SP-2013-607. Houston: National Aeronautics and Space Administration; 2013. p. 315–26.
74. Gibson CR, Mader TH, Chylack LT, Caputo MP. Ophthalmologic principles and concerns. In: Risin D, Stepaniak PC, editors. Biomedical results of the Space Shuttle Program, NASA/SP-2013-607. Houston: National Aeronautics and Space Administration; 2013. p. 47–58.
75. Ginter E, Simko V. Women live longer than men. Bratisl Lek Listy 2013;114:45–9.
76. Gopalakrishnan R, Genc KO, Rice AJ, et al. Muscle volume, strength, endurance, and exercise loads during 6-month missions in space. Aviat Space Environ Med 2010;81:91–102.
77. Gorgiladze GI, Brianov II. Space motion sickness. Kosm Biol Aviakosm Med 1989;23:4–14.
78. Gotoh TM, Fujiki N, Matsuda T, et al. Cerebral circulation during acute microgravity induced by free drop in anesthetized rats. Jpn J Physiol 2003;53:223–8.
79. Gradwell DP. Hypoxia and hyperventilation. In: Rainford DJ, Gradwell DP, editors. Ernsting's aviation medicine. 4th ed. New York: Oxford University Press; 2006. p. 44–50.
80. Grover S. Skin in aviation and space environment. Indian J Dermatol Venereol Leprol 2011;77(4):413–17.
81. Gueguinou N, Huin-Schohn C, Bascove M, et al. Could spaceflight-associated immune system weakening preclude the expansion of human presence beyond Earth's orbit? J Leukoc Biol 2009;86:1027–38.
82. Hamilton DR, Murray JD, Kapoor D, Kirkpatrick AW. Cardiac health for astronauts: Current selection standards and their limitations. Aviat Space Environ Med 2005;76:615–26.
83. Hamilton DR, Platts SH. Cardiovascular Issues. In: Risin D, Stepaniak PC, editors. Biomedical results of the Space Shuttle Program, NASA/SP-2013-607. Houston: National Aeronautics and Space Administration; 2013. p. 77–110.
84. Hamilton DR. Cardiovascular disorders. In: Barratt MR, Pool SL, editors. Principles of clinical medicine for space flight. New York: Springer; 2008. p. 324.
85. Harm DL, Parker DE. Perceived self-orientation and self-motion in microgravity, after landing and during preflight adaptation training. J Vestib Res 1993;3(3):297–305.
86. Harm DL, Sandoz GR, Stern RM. Changes in gastric myoelectric activity during space flight. Dig Dis Sci 2002;47:1737–45.
87. Harris BA Jr, Billica RD, Bishop SL, et al. Physical examination during space flight. Mayo Clin Proc 1997;72:301–8.
88. Hastings JL, Kowal J, Daniels J, et al. Cardiac arrhythmias during long duration spaceflight. Circulation 2013;128:A18904.
89. Hawkins WR, Zieglschmid JF. Clinical aspects of crew health. In: Johnston RS, Dietlein LF, Berry CA, editors. Biomedical results of Apollo, NASA SP-368. Washington, DC: US Government Printing Office; 1975. p. 43–81.
90. Heer M, Paloski WH. Space motion sickness: Incidence, etiology, and countermeasures. Auton Neurosci 2006;129(1/2):77–9.
91. Held KD. Effects of low fluences of radiations found in space on cellular systems. Int J Radiat Biol 2009;85:379–90.
92. Holland AW. NASA investigations of isolated and confined environments. Adv Space Biol Med 1993;3:15–21.
93. Hordinsky JR. Skylab crew health—Crew surgeons' reports. In: Johnston RS, Dietlein LF, editors. Biomedical results from Skylab, NASA SP-377. Washington, DC: National Aeronautics and Space Administration; 1977. p. 30–4.
94. Hu S, Kim MH, McClellan GE, Cucinotta FA. Modeling the acute health effects of astronauts from exposure to large solar particle events. Health Phys 2009;96:465–76.
95. Hughson RL, Shoemaker JH, Blaber AP, et al. Cardiovascular regulation during long-duration spaceflights to the International Space Station. J Appl Physiol 2012;112:719–27.
96. Ihle EC, Ritsher JB, Kanas N. Positive psychological outcomes of spaceflight: An empirical study. Aviat Space Environ Med 2006;77:93–102.
97. Ilcus LS, Johnston SL, Moynihan S, et al. Rashes and exanthems on long duration space flights [abstract]. Presented at 80th Annual Scientific Meeting of the Aerospace Medical Association, 2009. <http://naca.larc.nasa.gov/search.jsp?R=20090001314&qs=N%3D4294961259%2B4294955486%26No%3D50>.
98. Iwasaki K, Levine BD, Zhang R, et al. Human cerebral autoregulation before, during and after spaceflight. J Physiol 2007;579(3):799–810.
99. Jain V, Wotring VE. Menstrual cycle control in female astronauts and the associated risk of venous thromboembolism. Aerosp Med Hum Perform 2015;86:159.
100. James JT, Ott CM, Pierson DL. Environmental stressors: Toxicology and microbiology. In: Risin D, CStepaniak PC, editors. Biomedical results of the Space Shuttle Program, NASA/SP-2013-607. Houston: National Aeronautics and Space Administration; 2013. p. 327–46.
101. James JT. Health effects of atmospheric contamination. In: Barratt MR, Pool SL, editors. Principles of clinical medicine for space flight. New York: Springer; 2008. p. 427–44.
102. Javaheri S, Cobett WS, Simbartl LA, et al. Different effects of omeprazole and Sch 28080 on canine cerebrospinal fluid production. Brain Res 1997;754:321–4.
103. Jennings R. Managing space motion sickness. J Vestib Res 1998;8:67–70.
104. Jennings RT, Bagian JP. Musculoskeletal injury review in the U.S. space program. Aviat Space Environ Med 1996;67:762–6.
105. Jennings RT, Baker ES. Gynecologic and reproductive concerns. In: Barratt MR, Pool SL, editors. Principles of clinical medicine for space flight. New York: Springer; 2008. p. 381–90.
106. Jennings RT, Garriott OK, Bogomolov VV, et al. The ISS flight of Richard Garriott: A template for medicine and science investigation on future spaceflight participant missions. Aviat Space Environ Med 2010;81:133–5.
107. Jennings RT, Murphy DM, Ware DL, et al. Medical qualification of a commercial spaceflight participant: Not your average astronaut. Aviat Space Environ Med 2006;77:475–84.
108. Johansson RM, Malmvall BE, Andersson-Gäre B, et al. Guidelines for preventing urinary retention and bladder damage during hospital care. J Clin Nurs 2013;22(3/4):347–55.
109. Johnston SL, Blue RS, Jennings RT, et al. Astronaut medical selection during the shuttle era: 1981-2011. Aviat Space Environ Med 2014;85:823–7.
110. Johnston SL, Campbell MR, Scheuring R, Feiveson AH. Risk of herniated nucleus pulposus among U.S. astronauts. Aviat Space Environ Med 2010;81:566–74.
111. Jones JA, Karouia F. Radiation disorders. In: Barratt MR, Pool SL, editors. Principles of clinical medicine for space flight. New York: Springer; 2008. p. 475–520.
112. Jones JA, Pietrczyk RA, Whitson PA. Genitourinary disorders. In: Barratt MR, Pool SL, editors. Principles of clinical medicine for space flight. New York: Springer; 2008. p. 273–92.
113. Jones JA, Jennings RT, Baker ES. Renal, genitourinary, and gynecological health issues. In: Risin D, Stepaniak PC, editors. Biomedical results of the Space Shuttle Program, NASA/SP 2013-607. Houston: National Aeronautics and Space Administration; 2013. p. 141–56.

114. Kanas N. Psychiatric issues affecting long duration space missions. Aviat Space Environ Med 1998;69(12):1211–16.
115. Kass J, von Baumgarten R, Benson A, et al. The European vestibular experiments in Spacelab-1. Adv Space Res 1984;4(10):3–9.
116. Kass JR, Vogel H. Subjective vertical before and after space flight. Adv Space Res 1986;6(12):171–4.
117. Kerstman EL, Scheuring RA, Barnes MG, et al. Space adaptation back pain: a retrospective study. Aviat Space Environ Med 2012;83:2–7.
118. Kerstman EL. Rashes on space flights [abstract]. Presented at NASA Human Research Program Investigators' Workshop, Houston, Texas, 2010. <http://www.dsls.usra.edu/meetings/hrp2010/pdf/Thursday/Kerstman.pdf>.
119. Kirkpatrick AW, Nicolaou S, Campbell MR, et al. Percutaneous aspiration of fluid management of peritonitis in space. Aviat Space Environ Med 2002;73:925–30.
120. Kolesari GL, Kindwall EP. Survival following accidental decompression to an altitude greater than 74,000 feet (22,555 m). Aviat Space Environ Med 1982;53:1211–14.
121. Konstantinova IV, Rykova MP, Lesnyak AT, Antropova EA. Immune changes during long-duration missions. J Leukoc Biol 1993;54:189–201.
122. Kornilova LN, Naumov IA, Azarov KA, Sagalovitch VN. Gaze control and vestibular-cervical-ocular responses after prolonged exposure to microgravity. Aviat Space Environ Med 2012;83(12):1123–34.
123. Korth D, Leveton L, Dinges D. Slam shift figures summary. Presented at NASA Human Research Program Behavioral Health and Performance Element Programmatic Review, NASA Johnson Space Center, Houston, Texas, 2006.
124. Kramer LA, Hasan KM, Sargsyan AE, et al. MR-derived cerebral spinal fluid hydrodynamics as a marker and a risk factor for intracranial hypertension in astronauts exposed to microgravity. J Magn Reson Imaging 2015;42(6):1560–71. <http://onlinelibrary.wiley.com/doi/10.1002/jmri.24923/abstract;jsessionid=610A80F00890AB89D1C4CD39F533EB30.f04t04>.
125. Kramer LA, Sargsyan AE, Hasan KM, et al. Orbital and intracranial effects of microgravity: findings at 3-T MR imaging. Radiology 2012;263:819–27.
126. Krause KM, Pilmanis AA. The effectiveness of ground level oxygen treatment for altitude decompression sickness in human research subjects. Aviat Space Environ Med 2000;71(2):115–18.
127. Lam CW, Scully RR, Zhang Y, et al. Toxicity of lunar dust assessed in inhalation-exposed rats. Inhal Toxicol 2013;25:661–78.
128. Lane HW, Bourland C, Barrett A, et al. The role of nutritional research in the success of human space flight. Adv Nutr 2013;4:521–3.
129. Lang T, LeBlanc A, Evans H, et al. Cortical and trabecular bone mineral loss from the spine and hip in long-duration spaceflight. J Bone Miner Res 2004;9:1006–12.
130. Law J, Van Baalen M, Foy M, et al. Relationship between carbon dioxide levels and reported headaches on the International Space Station. J Occup Environ Med 2014;56(5):477–83.
131. Lawley JS, Williams MA, Petersen L, et al. ICP during daily life in healthy adults: What does microgravity add to the mix? Presented at NASA Human Research Program Investigators' Workshop, Galveston, Texas, 2015.
132. Leach CS, Alfrey CP, Suki WN, et al. Regulation of body fluid compartments during short-term spaceflight. J Appl Physiol 1996;81:105–16.
133. Lebedev V. Diary of a cosmonaut: 211 days in space. Moscow: Nauka I Zhizn; 1983. English translation c1988 by the GLOSS Co, New York, 1990, Bantam Books.
134. LeBlanc A, Lin C, Shackelford L, et al. Muscle volume, MRI relaxation times (T2), and body composition after spaceflight. J Appl Physiol 2000;89(6):2158–64.
135. Leblanc A, Matsumoto T, Jones J, et al. Bisphosphonates as a supplement to exercise to protect bone during long-duration spaceflight. Osteoporos Int 2013;24:2105–14.
136. LeBlanc AD, Spector ER, Evans HJ, Sibonga JD. Skeletal responses to space flight and the bed rest analog: A review. J Musculoskelet Neuronal Interact 2007;7:33–47.
137. Lee SM, Moore AD Jr, Fritsch-Yelle JM, et al. Inflight exercise affects stand test responses after space flight. Med Sci Sports Exerc 1999;31:1755–62.
138. Lee SMC, Guilliams ME, Siconolfi SF, et al. Concentric strength and endurance after long duration spaceflight. Med Sci Sports Exerc 2000;32:S363.
139. Lerner DJ, Parmet AJ. Interventional radiology: The future of surgery in microgravity. Aviat Space Environ Med 2013;84:1304–6.
140. Lesnyak AT, Sonnenfeld G, Rykova MP, et al. Immune changes in test animals during spaceflight. J Leukoc Biol 1993;54(3):214–26.
141. Levine BD, Lane LD, Watenpaugh DE, et al. Maximal exercise performance after adaptation to microgravity. J Appl Physiol 1996;81:686–94.
142. Li M, Gonon G, Buonanno M, et al. Health risks of space exploration: Targeted and nontargeted oxidative injury by high-charge and high-energy particles. Antioxid Redox Signal 2014;20:1501–23.
143. Lindvall-Axelsson M, Nilsson C, Owman C, Winbladh B. Inhibition of cerebrospinal fluid formation by omeprazole. Exp Neurol 1992;115:394–9.
144. Lobascio C, Briccarello M, Destefanis R, et al. Accelerator-based tests of radiation shielding properties of materials used in human space infrastructures. Health Phys 2008;94(3):242–7.
145. Mader TH, Gibson CR, Pass AF, et al. Optic disc edema in an astronaut after repeat long-duration space flight. J Neuroophthalmol 2013;33:249–55.
146. Mader TH, Gibson CR, Pass AF, et al. Optic disc edema, globe flattening, choroidal folds, and hyperopic shifts observed in astronauts after long-duration space flight. Ophthalmology 2011;118:2058–69.
147. Mader TH, Koch DD, Manuel K, et al. Stability of vision during space flight in an astronaut with bilateral intraocular lenses. Am J Ophthalmol 1999;127:342–3.
148. Mader TH, Taylor GR, Hunter N, et al. Intraocular pressure, retinal vascular, and visual acuity changes during 48 hours of 10 degrees head-down tilt. Aviat Space Environ Med 1990;61:810–13.
149. Mallis MM, DeRoshia CW. Circadian rhythms, sleep, and performance in space. Aviat Space Environ Med 2005;76:B94–107.
150. Manuel FK, Mader TH. Ophthalmologic concerns. In: Barratt MR, Pool SL, editors. Principles of clinical medicine for space flight. New York: Springer; 2008. p. 535–44.
151. Manzey D, Lorenz B, Schiewe A, et al. Dual-task performance in space: Results from a single-case study during a short-term space mission. Hum Factors 1995;37(4):667–81.
152. Manzey D, Lorenz B. Human performance during spaceflight. Hum Perf Extrem Environ 1999;4(1):8–13.
153. Mark S, Scott GB, Donoviel DB, et al. The impact of sex and gender on adaptation to space: Executive summary. J Womens Health (Larchmt) 2014;23:941–7.
154. Marshall-Bowman K. Increased intracranial pressure and visual impairment associated with long-duration spaceflight. Individual Project Report submitted to the International Space University in partial fulfillment of the requirements of the M.Sc. Degree in Space Studies, 2011. <http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20110014672.pdf>.
155. Marshburn T, Cole R, Pavela J, et al. Facial soft tissue measurement in microgravity-induced fluid shifts. Aviat Space Environ Med 2014;85:227.
156. Marshburn T. Phase I/Mir clinical experience. In: Ball JR, Evans CH Jr, editors. Safe passage: Astronaut care for exploration missions. Washington, DC: National Academy Press; 2001. Presentation to the Institute of Medicine Committee on Creating a Vision for Space Medicine during Travel Beyond Earth Orbit, Johnson Space Center, Houston, 2000.
157. Marshburn TH. Acute care. In: Barratt MR, Pool SL, editors. Principles of clinical medicine for space flight. New York: Springer; 2008. p. 101–22.
158. McCarthy EF. Perspective: Skeletal complications of space flight. Skeletal Radiol 2011;40:661–3.
159. McCuaig K. Surgical problems in space: An overview. J Clin Pharmacol 1994;34:513–17.
160. McMonigal KA, Braverman LE, Dunn JT, et al. Thyroid function changes related to use of iodinated water in the U.S. Space Program. Aviat Space Environ Med 2000;71:1120–5.
161. Meck JV, Reyes CJ, Perez SA, et al. Marked exacerbation of orthostatic intolerance after long- vs. short-duration spaceflight in veteran astronauts. Psychosom Med 2001;63:865–73.
162. Mehta SK, Stowe RP, Feiveson AH, et al. Reactivation and shedding of cytomegalovirus in astronauts during spaceflight. J Infect Dis 2000;182(6):1761–4.
163. Mermel LA. Infection prevention and control during prolonged human space travel. Clin Infect Dis 2013;56(1):123–30. erratum 56(11):1684.
164. Moore AD Jr, Downs ME, Lee SM, et al. Peak exercise oxygen uptake during and following long-duration spaceflight. J Appl Physiol 2014;117:231–8.
165. Murray DH, Pilmanis AA, Blue RS, et al. Pathophysiology, prevention, and treatment of ebullism. Aviat Space Environ Med 2013;84:89–96.
166. Myasnikov VI, Stepanova SI. Features of cerebral hemodynamics in cosmonauts before and after flight on the MIR Orbital Station. Moscow: Russian Federation: State Scientific Center—Institute of Biomedical Problems; 2012. p. 300–5.
167. Nagaraja MP, Risin D. The current state of bone loss research: Data from spaceflight and microgravity simulators. J Cell Biochem 2013;114(5):1001–8.

168. National Council on Radiation Protection and Measurements. Recommendations of dose limits for low Earth orbit, NCRP Report 132, Bethesda, Md, 2000.
169. Nedelmann M, Kaps M, Mueller-Forell W. Venous obstruction and jugular valve insufficiency in idiopathic intracranial hypertension. J Neurol 2009;256:964–9.
170. Newkirk D. Second-generation space stations. In: Almanac of Soviet manned space flight. Houston: Gulf Publishing; 1990. p. 158–296.
171. Newkirk D. Mars precursors. In: Almanac of Soviet manned space flight. Houston: Gulf Publishing; 1990. p. 297–372.
172. Nickerson C, Ott C, Mister S, et al. Microgravity as a novel environment signal affecting *Salmonella* enteric serovar *typhimurium* virulence. Infect Immun 2000;68(6):3147–52.
173. Norsk P, Asmar A, Damgaard M, Christensen NJ. Fluid shifts, vasodilatation and ambulatory blood pressure reduction during long duration spaceflight. J Physiol 2015;593:573–84.
174. Norsk P, Christensen NJ. The paradox of systemic vasodilatation and sympathetic nervous stimulation in space. Respir Physiol Neurobiol 2009;169S:S26–9.
175. Norsk P, Damgaard M, Petersen L, et al. Vasorelaxation in space. Hypertension 2006;47:69–73.
176. Norsk P. Cardiovascular and fluid volume control in humans in space. Curr Pharm Biotechnol 2005;6:325–30.
177. Oman CM, Lichtenberg BK, Money KE. Space motion sickness monitoring experiment: Spacelab 1. In: Crampton GH, editor. Motion and space sickness. Boca Raton, Fla: CRC Press; 1990. p. 217–46.
178. Ontiveros C, McCabe LR. Simulated microgravity suppresses osteoblast phenotype, Runx2 levels and AP-1 transactivation. J Cell Biochem 2003;88:427–37.
179. Orwoll ES, Adler RA, Amin S, et al. Skeletal health in long-duration astronauts: Nature, assessment, and management recommendations from the NASA Bone Summit. J Bone Miner Res 2013;28:1243–55.
180. Ott CM, Bruce RJ, Pierson DL. Microbial characterization of free floating condensate aboard the Mir spacestation. Microb Ecol 2004; 47:133–6.
181. Palinkas LA, Gunderson EK, Holland AW, et al. Predictors of behavior and performance in extreme environments: The Antarctic Space Analogue Program. Aviat Space Environ Med 2000;71:619–25.
182. Palinkas LA, Gunderson EK, Johnson JC, Holland AW. Behavior and performance on long-duration spaceflights: Evidence from analogue environments. Aviat Space Environ Med 2000;71:A29–36.
183. Paule MG, Chelonis JJ, Blake DJ, Dornhoffer JL. Effects of drug countermeasures for space motion sickness on working memory in humans. Neurotoxicol Teratol 2004;26:825–37.
184. Perez SA, Charles JB, Fortner GW, et al. Cardiovascular effects of anti-G suit and cooling garments during space shuttle re-entry and landing. Aviat Space Environ Med 2003;74:753–7.
185. Perhonen MA, Franco F, Lane LD, et al. Cardiac atrophy after bed rest and spaceflight. J Appl Physiol 2001;91:645–53.
186. Pierson DL, Chidambaram M, Heath JD, et al. Epidemiology of *Staphylococcus aureus* during space flight. FEMS Immunol Med Microbiol 1996;16(3/4):273–81.
187. Pierson DL, Mehta SK, Magee BB, Mishra SK. Person-to-person transfer of *Candida albicans* in the spacecraft environment. J Med Vet Mycol 1995;33(3):145–50.
188. Pierson DL, Viktorov AN. Microbiological investigations of the Mir space station and flight crew. In Shuttle-Mir Science Program Phase 1A Research Postflight Science Report, Unpublished NASA document, Houston, Texas, 1998, NASA-Johnson Space Center.
189. Pietsch J, Bauer J, Egli M, et al. The effects of weightlessness on the human organism and mammalian cells. Curr Mol Med 2011;11: 350–64.
190. Putcha L, Berens KL, Marshburn TH, et al. Pharmaceutical use by U.S. astronauts on space shuttle missions. Aviat Space Environ Med 1999;70:705–8.
191. Putcha L, Boyd JL, Vaksman Z, Dubois A. Gastrointestinal physiology and function: A pharmacotherapeutic perspective. In: Risin D, Stepaniak PC, editors. Biomedical results of the Space Shuttle Program, NASA/SP-2013-607. Houston: National Aeronautics and Space Administration; 2013. p. 131–40.
192. Reitz G. Characteristic of the radiation field in low Earth orbit and in deep space. Z Med Phys 2008;18:233–43.
193. Reschke MF, Wood SJ, Harm KL, et al. Neurovestibular/sensorimotor system. In: Risin D, Stepaniak PC, editors. Biomedical results of the Space Shuttle Program, NASA/SP-2013-607. Houston: National Aeronautics and Space Administration; 2013. p. 171–214.
194. Riggs BL, Melton LJ III, Robb RA, et al. A population-based assessment of rates of bone loss at multiple skeletal sites: Evidence for substantial trabecular bone loss in young adult women and men. J Bone Miner Res 2008;23:205Y14.
195. Ronca AE, Fritzsch B, Bruce LL, Alberts JR. Orbital spaceflight during pregnancy shapes function of mammalian vestibular system. Behav Neurosci 2008;122:224–32.
196. Rossum AC, Wood ML, Bishop SL, et al. Evaluation of cardiac rhythm disturbances during extravehicular activity. Am J Cardiol 1997;79: 1153–5.
197. Roth EM. Rapid (explosive) decompression emergencies in pressure-suited subjects, NASA CR-1223. NASA Contract Rep NASA CR 1968.
198. Sams CF, Pierson DL. Immunologic concerns. In: Barratt MR, Pool SL, editors. Principles of clinical medicine for space flight. New York: Springer; 2008. p. 307–16.
199. Sawin CF, Nicogossian AE, Rummel JA, Michel EL. Pulmonary function evaluation during Skylab and Apollo-Soyuz missions. Aviat Space Environ Med 1976;47(2):168–72.
200. Sayson JV, Hargens AR. Pathophysiology of low back pain during exposure to microgravity. Aviat Space Environ Med 2008;79: 365–73.
201. Reference deleted in proofs.
202. Scheuring RA, Mathers CH, Jones JA, Wear ML. Musculoskeletal injuries and minor trauma in space: Incidence and injury mechanisms in U.S. astronauts. Aviat Space Environ Med 2009;80:117–24.
203. Schultz JR, Fuhrmann K. DTO 635: Eyewash evaluation. In Results of life sciences DSOs conducted aboard the shuttle 1991-1993, Unpublished NASA document, Houston, 1994, NASA-Johnson Space Center, pp 121-2.
204. Shackelford LC. Musculoskeletal response to space flight. In: Barratt MR, Pool SL, editors. Principles of clinical medicine for space flight. New York: Springer; 2008. p. 293–306.
205. Shayler DJ. Disasters and accidents in manned spaceflight. Chichester: Praxis Publishing; 2000.
206. Shimoda T, Oikawa T, Miyake A. Sampling and analysis of human metabolites. In Society of Automotive Engineers Technical Paper No 981739, Warrendale. Presented at the 28th Conference on Environmental Systems, Danvers, 1998.
207. Sibonga JD. Spaceflight-induced bone loss: Is there an osteoporosis risk? Curr Osteoporos Rep 2013;11:92–8.
208. Sipes WE, Barger LK, Czeisler CA. Behavioral health and performance support. In: Risin D, Stepaniak PC, editors. Biomedical results of the Space Shuttle Program, NASA/SP-2013-607. Houston: National Aeronautics and Space Administration; 2013. p. 259–66.
209. Sipes WE, Vander Ark ST. Operational behavioral health and performance resources for International Space Station crews and families. Aviat Space Environ Med 2005;76:B36–41.
210. Smith SM, Abrams SA, Davis-Street JE, et al. Fifty years of human space travel: Implications for bone and calcium research. Annu Rev Nutr 2014;34:377–400.
211. Smith SM, Davis-Street JE, Fontenot TB, Lane HW. Assessment of a portable clinical blood analyzer during space flight. Clin Chem 1997; 43:1056–65.
212. Smith SM, Heer MA, Shackelford LC, et al. Benefits for bone from resistance exercise and nutrition in long-duration spaceflight: Evidence from biochemistry and densitometry. J Bone Miner Res 2012; 27:1896–906.
213. Soller BR, Cabrera M, Smith SM, Sutton JP. Smart medical systems with application to nutrition and fitness in space. Nutrition 2002; 18:930–6.
214. Sonnenfeld G. Effect of space flight on cytokine production. Acta Astronaut 1994;33:143–7.
215. Stein TP, Leskiw MJ, Schluter MD, et al. Energy expenditure and balance during spaceflight on the space shuttle. Am J Physiol 1999; 276:R1739–48.
216. Stenger MB, Lee SM, Westby CM, et al. Abdomen-high elastic gradient compression garments during post-spaceflight stand tests. Aviat Space Environ Med 2013;84:459–66.
217. Stepaniak PC, Ramchandani SR, Jones JA. Acute urinary retention among astronauts. Aviat Space Environ Med 2007;78(4):A5–8.
218. Stewart KT, Hayes BC, Eastman CI. Light treatment for NASA shiftworkers. Chronobiol Int 1995;12:141–51.
219. Strauss S, Krog RL, Feiveson AH. Extravehicular mobility unit training and astronaut injuries. Aviat Space Environ Med 2005;76(5): 469–74.
220. Styf JR, Kolebo P, Hargens AR. Lumbar intrathecal disc heights as measured by sonography. Aviat Space Environ Med 1994;65:450.
221. Suedfeld P. Invulnerability, coping, salutogenesis, integration: Four phases of space psychology. Aviat Space Environ Med 2005;76: B61–6.
222. Suman S, Rodriguez OC, Winters TA, et al. Therapeutic and space radiation exposure of mouse brain causes impaired DNA repair response and premature senescence by chronic oxidant production. Aging (Albany NY) 2013;5(8):607–22.
223. Summers RL, Martin DS, Meck JV, Coleman TG. Mechanism of spaceflight-induced changes in left ventricular mass. Am J Cardiol 2005;95:1128–30.
224. Sweet TB, Panda N, Hein AM, et al. Central nervous system effects of whole-body proton irradiation. Radiat Res 2014;182(1):18–34.

225. Tamma R, Colaianni G, Camerino C, et al. Microgravity during spaceflight directly affects in vitro osteoclastogenesis and bone resorption. FASEB J 2009;23:2549–54.
226. Taylor GR. Overview of spaceflight immunology studies. J Leukoc Biol 1993;54(3):179–88.
227. Thornton WE, Bonato F. Space motion sickness and motion sickness: Symptoms and etiology. Aviat Space Environ Med 2013;84:716–21.
228. Thornton WE, Hoffler GW, Rummel JA. Anthropometric changes and fluid shifts. In: Johnston RS, Dietlein LF, editors. Biomedical results from Skylab, NASA SP-377. Washington, DC: National Aeronautics and Space Administration; 1977. p. 330–8.
229. Trambovetskii EV, Krotov VP, Korol'kov VI. Intracranial pressure in monkeys during the flight of Cosmos-2229. Aviakosm Ekolog Med 1995;29:37–41.
230. Trappe S, Costill D, Gallagher P, et al. Exercise in space: Human skeletal muscle after 6 months aboard the International Space Station. J Appl Physiol 2009;106:1159–68.
231. Tronnier H, Wiebusch M, Heinrich U. Change in skin physiological parameters in space: Report on and results of the first study on man. Skin Pharmacol Physiol 2008;21(5):283–92.
232. Videbaek R, Norsk P. Atrial distension in humans during weightlessness induced by parabolic flights. J Gravit Physiol 1996;3(2):48–9.
233. Waters WW, Ziegler MG, Meck JV. Post-spaceflight orthostatic hypotension occurs mostly in women and is predicted by low vascular resistance. J Appl Physiol 2002;92:586–94.
234. Watt DG, Money KE, Bondar RL, et al. Canadian medical experiments on Shuttle flight 41-G. Can Aeronaut Space J 1985;31(3):215–26.
235. West JB, Prisk GK. Chest volume and shape and intrapleural pressure in microgravity. J Appl Physiol 1999;87:1240–1.
236. Whitson PA, Pietrzyk RA, Pak CY, Cintrón NM. Alterations in renal stone risk factors after space flight. J Urol 1993;150(3):803–7.
237. Whitson PA, Pietrzyk RA, Pak CY. Renal stone risk assessment during Space Shuttle flights. J Urol 1997;158(6):2305–10.
238. Whitson PA, Putcha L, Chen YM, Baker E. Melatonin and cortisol assessment of circadian shifts in astronauts before flight. J Pineal Res 1995;18:141–7.
239. Willey JS, Lloyd SA, Nelson GA, Bateman TA. Space radiation and bone loss. Gravit Space Biol Bull 2011;25:14–21.
240. Wilson MH, Imray CH, Hargens AR. The headache of high altitude and microgravity: Similarities with clinical syndromes of cerebral venous hypertension. High Alt Med Biol 2011;12(4):379–86.
241. Wong KL. Ammonia. In: Spacecraft maximum allowable concentrations for selected airborne contaminants, vol. 1. Washington, DC: National Academy Press; 1994. p. 39–59.
242. Wong KL. Carbon dioxide. In: Spacecraft maximum allowable concentrations for selected airborne contaminants, vol. 2. Washington, DC: National Academy Press; 1994. p. 105–87.
243. Wood SJ, Loehr JA, Guilliams ME. Sensorimotor reconditioning during and after spaceflight. Neuro Rehabil 2011;29(2):185–95.
244. Zuj KA, Argelle PH, Shoemaker JK, et al. Impaired cerebrovascular autoregulation and reduced CO_2 reactivity after long duration spaceflight. Am J Physiol Heart Circ Physiol 2012;302:H2592–8.
245. Zwart SR, Gibson CR, Mader TH, et al. Vision changes after spaceflight are related to alterations in folate- and vitamin B-12-dependent one-carbon metabolism. J Nutr 2012;142:427–31.

APPENDIX

Drug Stability in the Wilderness

BEAU A. BRIESE AND MILLICENT M. BRIESE

ENVIRONMENTAL FACTORS INFLUENCING DRUG STABILITY

The main environmental factors affecting drug stability are temperature, light, and humidity. In addition, additives included with a medication can preserve or, under certain conditions, diminish a drug's efficacy during long-term storage.

Extreme temperatures and light can cause medications to spontaneously decompose, reassemble, or react with air, contaminants, or a drug's otherwise inactive ingredients. Exposure to high humidity can decrease a drug's rate of dissolution, the bottleneck step in bioavailability of drugs taken by mouth. Common additives, such as bicarbonate or D5W, and dilution can drastically reduce robustness and durability of some medications in extreme environments. Knowing a medication's anticipated rate of breakdown indicates how to titrate the dosage during extended periods in the field and how often to resupply. The most accurate understanding of the robustness of a medication's efficacy in the wilderness stems from evidence-based research.

Manufacturers generally recommend little variation in storage temperatures. Brief excursions into temperatures below the minimum or beyond the maximum recommended temperatures for a given drug are often acceptable, so long as two conditions are met: (1) the drug is not exposed to a maximum temperature constituting excessive heat for a period longer than 24 hours and (2) the mean kinetic temperature (MKT), or average temperature, for the drug remains at or below the maximum temperature of its ideal range.[25] These less commonly known rules for excursions ease medical fieldwork related to wilderness emergency situations of short duration by minimizing the amount of artificial cooling absolutely necessary in the field, depending on the drugs involved.

Being mindful of the season for which one estimates the MKT enables one to anticipate conditions that risk degrading drug potency and composition. Monthly MKTs were recorded for an emergency medical services (EMS) vehicle in South Africa that was not equipped with an electrical cooling system.[21] For the 12-month study period, the MKT was 24.7° C (76.5° F), just under the 25° C (77° F) limit for drugs needing controlled room temperature storage. However, the 6-month period between the Southern Hemisphere's warmer months of October through March had an excessive MKT of 27° C (80° F).

Helm and colleagues evaluated how season affected storage temperatures in unpowered containers in prehospital EMS vehicles that included a helicopter, ambulance, and physician car in southern Germany.[12] Despite the temperate climate, storage temperature exceeded 25° C (77° F) 33% to 45% of time in summer, and was less than 0° C (32° F) 19% of the time in winter.

In another study, epinephrine kept primarily at room temperature but exposed to extreme heat for 8 hours per day had impaired drug durability following dilution. Exposure to 5° C (41° F) did not cause epinephrine (1 : 1000 or 1 : 10,000 concentration) degradation at 4, 8, and 12 weeks. At 70° C (158° F), no 1 : 1000 epinephrine was lost, but only 36% of 1 : 10,000 epinephrine remained at 12 weeks.[11] Similarly, 7 days of constant exposure to 65° C (149° F) caused no depletion of 1 : 1000 epinephrine but complete destruction of 1 : 10,000 epinephrine.[5] Degraded 1 : 10,000 epinephrine is less effective in clinical use. Compared with controls stored at room temperature, at least 30% more (by volume) 1 : 10000 epinephrine exposed cyclically to 70° C for 8 or 12 weeks was required to achieve the same physiologic effect.[11]

The effect of dry desert heat on drug stability has also been studied. Valenzuela and associates studied degradation of 23 prehospital drugs stored in an unventilated drug box in a metal storage shed in arid Tucson, Arizona during the dry desert heat of summer.[26] The drugs included aminophylline, atropine, bretylium tosylate, calcium chloride, dexamethasone, dextrose, diazepam, diphenhydramine, dopamine hydrochloride, epinephrine, furosemide, isoetharine, isoproterenol, lidocaine, metoprolol tartrate, morphine sulfate, naloxone, nifedipine, nitroglycerin tablets, phenobarbital, sodium bicarbonate, thiamine, and verapamil. This 4-week study exposed the drugs to temperatures that ranged from 28° to 39° C (82° to 102° F). By the end of the study, only isoproterenol, epinephrine, and nifedipine exhibited noteworthy chemical changes. Isoproterenol was 11% less potent. Although the epinephrine itself remained unaffected, the pH of its storage solution degraded, creating a more acidic solution. As with epinephrine, nifedipine remained stable, but its delivery vehicle was compromised. The nifedipine capsules melted.

In the temperate climate of Los Angeles County in southern California, Gill and colleagues evaluated how exceeding the recommended MKT in advanced life support paramedic vehicles affected drug concentrations.[10] The 45-day study measured temperature and drug concentrations of atropine, epinephrine, and lidocaine in 15 different geographic locations, including one laboratory control, an inland airport, a harbor, six high desert neighborhoods, and six inland suburbs. The manufacturers of all three drugs recommended storage at 20° to 25° C (68° to 77° F), with temperatures not to exceed 40° C (104° F) for periods of more than 24 hours. The authors' findings indicate that atropine, epinephrine, and lidocaine can withstand an MKT of up to 29° C (84° F) for up to 45 days without degradation. The study also concluded that those drugs could tolerate spikes up to 52° C (125° F) for a cumulative time of up to 13 hours (795 minutes) without degradation.

Drugs can withstand only so much heat and cold before they begin to lose potency. Gammon and colleagues tested stability of 23 prehospital drugs thermally cycled between 12-hour periods at −6° C (21° F) and 54° C (129° F) for 1 month.[9] The drugs included adenosine, albuterol, amiodarone, atropine, diltiazem, dopamine, epinephrine, etomidate, haloperidol, heparin, hydralazine, ipratropium, labetalol, lidocaine, naloxone, nitroglycerin, ondansetron, oxytocin, procainamide, succinylcholine, terbutaline, thiamine, and vasopressin. Eight drugs degraded to less than 90% potency: diltiazem, dopamine, haloperidol, ipratropium, lidocaine, naloxone, nitroglycerin, and succinylcholine.

Mathijssen and associates examined antibiotics stored at constant temperatures ranging from −80° C to 37° C (−112° to 99° F) for up to 1 year.[16] Antibacterial activity of linezolid and clindamycin pills and solution was unchanged. However, efficacy of oxacillin and cefazolin diminished when stored at 37° C (99° F) for 1 month or at 20° C for 6 months. Activity of vancomycin became substandard when stored at 37° C (99° F) for 6 months.

EXPIRATION DATES AND SHELF LIFE

The United States Pharmacopeia (USP) requires manufacturers to list a drug's shelf life as the time during which a drug's potency (or concentration of active product) is guaranteed to be 90% to 110% of its listed potency (or concentration).[25] Expiration dates are based on the shelf life of a drug under ideal, manufacturer-suggested conditions of temperature, humidity, light exposure, and packaging integrity. When stored in environments that do not correlate with those listed by the manufacturer, the printed expiration date no longer indicates whether a drug is potent. For this reason, multiple studies have evaluated the rate and extent of degradation and loss of potency for numerous drugs under various circumstances.

Tropical climates pose difficulty for drugs susceptible to heat and humidity. One concern is that drugs from one climate may not be optimized for use in another environment. Risha and colleagues evaluated two primary markers of potency, namely drug content and bioavailability, for ciprofloxacin and diclofenac tablets in Tanzania by exposing the drugs to the following conditions: temperature of 40° ± 2° C (104° ± 4° F) and 75% ± 5% relative humidity.[20] Drug content and bioavailability were tested at the beginning of the study and again at 3 and 6 months for eight formulations of ciprofloxacin from Belgium and India, and for four formulations of diclofenac sodium from Belgium, India, Malaysia, and Cyprus. All formulations of both drugs complied with the USP required level of 90% to 110% of labeled drug content during the entire span of the study. Oral bioavailability remained within required levels for all formulations of ciprofloxacin, because all formulations complied with ciprofloxacin's dissolution regulations, dissolving 80% or more of the drug within 30 minutes. However, dissolution levels were substandard for two of the four diclofenac formulations. Those from Camden (Malaysia) and Remedica (Cyprus) failed to dissolve during the full course of the dissolution test. This indicates that although a drug may be active, stable, and of proper concentration, it still may be inaccessible after ingestion if, because of environmental exposures, it can no longer dissolve. Drug potency and dissolution should be tested regularly, because the robustness of a drug in more extreme environments may vary between manufacturers, formulations, and batches.

Maintaining drug integrity requires quality control in all steps of the process, from manufacturing through storage to delivery. In preparing for an expedition, be mindful that drugs purchased in certain locations may be less potent and durable than those purchased elsewhere. Twagirumukiza and associates studied 16 formulations of the medications atenolol, captopril, hydrochlorothiazide, methyldopa, and propranolol, 10 purchased from Rwandan pharmacies and 6 reference formulations purchased in Belgium or France.[24] All drug formulations were labeled with expiration dates indicating that at least 2 years of shelf life remained. Of the 10 formulations purchased in Rwanda, 2 exhibited substandard percentages of content on initial receipt. After 6 months, 7 of the 10 medications purchased in Rwanda had less than 90% of their original content, and 6 had impaired dissolution profiles. This indicated both reduced content and diminished bioavailability of the remaining medication.

PACKAGING

Where drugs are stored significantly influences their stability and safety. Packaging can shield drugs from environmental assaults, but only when conditions optimize the packaging's performance. For example, glass or plastic syringes containing medication, such as epinephrine, for immediate use may develop hairline cracks when frozen, leading to leakage and compromising stability and sterility of the remaining drug.

Independent of environmental conditions, packaging can negatively influence a drug's stability by leaching chemicals into drugs, absorbing drugs, and reacting with medications. These effects may reduce efficacy of stored medications and increase potential for their toxicity. Polyvinylchloride (PVC) is known to contain toxic compounds that may seep into drugs in trace amounts; the most infamous is the carcinogen diethylhexyl phthalate, which represents 30% to 80% of the weight of medical bags and intravenous (IV) tubing that contain PVC.[22] Medications can also be absorbed by the PVC itself. Alternative packaging materials, such as polypropylene and polyethylene, have demonstrated lower risk for absorption than does PVC for medications such as nitroglycerin and diazepam.[23] Over extended periods, glass containers can deposit reactive alkali decomposition materials into drugs. Leaching, absorption, and reactivity of packaging made of glass, PVC, polypropylene, and polyethylene have not been studied for most medications.

Although packaging may protect a drug from environmental extremes and degradation, packaging might degrade the drug, reducing its bioavailability. For example, blister packs of atenolol from Alpharma maintain bioavailability after 28 days at temperatures of 40° C (104° F) and humidity levels of 75%, but blister packs of atenolol from CP Pharmaceuticals do not, even though both versions of atenolol, when not stored in blister packs, are equally robust in some similar environmental conditions.[7]

Hoye and colleagues evaluated efficacy of hydrofluoroalkane (HFA) inhalation aerosols commonly used in albuterol metered-dose inhalers (MDIs) for proficiency in drug delivery at high temperatures.[13] The study included (1) a 185-day evaluation of albuterol (Proventil) HFA and albuterol (Ventolin) HFA MDIs stored in extreme temperatures but tested at room temperature and (2) evaluation of the performance of the MDIs when actuated at 4°, 22°, 47° and 60° C (39°, 72°, 117°, and 140° F). In the first portion of the study, inhaler frames warped and canisters had a minor increase in rates of propellant leakage, but showed integrity for the size of emitted particles and the dose per actuation. Proper drug delivery was unaffected by storage at temperatures ranging from −3° to 88° C (26° to 190° F). The second portion of the study concluded that the amount of drug successfully delivered decreased as the temperature at the time of delivery increased. The dose per actuation was more drastically reduced for the Proventil HFA MDI, for which 15% less albuterol was dispersed at 60° C (140° F) than at 4° C (39° F); the Ventolin HFA MDI exhibited only an 8% decrease. It is therefore recommended that propellant-based drug delivery systems be sprayed in conditions as close as possible to the manufacturer's storage parameters.

It is not always possible to keep drugs in the original packaging. Drugs stored outside of the manufacturer's container might exhibit significantly altered shelf lives. Rawas-Qalaji and associates evaluated the effects of the duration of humidity and sunlight on the stability of 0.3 mg of 1 mg/mL epinephrine transferred to unsealed syringes stored at high temperatures.[19] Their study examined four standardized storage environments at a constant 38° C (100.4° F): darkness with low (15%) humidity, darkness with high (95%) humidity, sunlight with low (15%) humidity, and sunlight with high (95%) humidity. Results suggest that presence or absence of sunlight did not affect injectable epinephrine. On the other hand, low humidity accelerated decompensation. Syringes placed into low humidity decreased to 90% potency by the end of the second month, and drastically dropped to 60%, 55%, and 39% potency at the end of months three, four, and five, respectively. Syringes placed into high humidity statistically fared better. Their epinephrine reduced to 90% potency by the end of the third month and dropped below regulation limits, falling to 83% and 82% in months four and five, respectively. Humidity is protective of repackaged epinephrine solution.

Repackaged lidocaine solution is stable over a moderate range of temperatures. In one study, lidocaine placed into 2-mL Tubex cartridges (20 mg/mL) and lidocaine diluted with 5% dextrose in plastic infusion bags (4 mg/mL) remained potent and stable for 3 months at room temperature; the latter also remained stable for 3 months under refrigeration at 4° C (39° F).[14]

Storage of drug mixtures can cause loss of potency and bioavailability that worsens with time. One percent lidocaine buffered with bicarbonate at a pH range of 7.38 to 7.41 remained effective for up to 1 week, decreasing in potency by approximately 10%, the lowest acceptable loss of potency, after 1 week of storage.[2] Avoid storing pills of different drugs in the same container, because interactions in storage can reduce dissolution rates, as occurs when atenolol is costored with some formulations of generic aspirin.[7]

Other forms of drug costorage are extremely stable. Implantable infusion systems are helping previously impaired individuals return to the outdoors. Mixtures of bupivacaine and clonidine with either morphine or hydromorphone have been shown to remain stable in implantable infusion systems at 37° C (99° F) for 90 days.[3]

For small pill storage, be aware that some pill containers appear to maintain pill integrity more than do others. One study demonstrated that Medidose pill packs maintained atenolol's bioavailability more effectively than did either blister packs or refillable pill containers at 25° C (77° F), but the converse was true at 40° C (104° F) and 75% humidity.[7] Place such pill packs in a water-sealed, light-tight container not containing PVC. Brands such as SealLine, Seattle Sports, Dry Pak, and Sea to Summit offer bags that meet these requirements.

STERILITY

Some drugs must be mixed with buffer or saline solutions prior to use. When preparing these solutions, caution should be taken to ensure sterile formulation. Brief exposures of less than 4 hours to air do not appear to compromise sterility. Carrasco and colleagues evaluated stability and sterility of saline infusion solutions.[18] Solutions of 0.9% saline in their original containers were transferred to polyethylene bottles or PVC bags, each equipped with a 1.5-μm bacterial filter, air intake, and nonextendable three-way valve with protected caps. These solutions were placed in various mobile intensive care units in urban portions of western Andalusia, and tested for sterility after 24, 48, and 72 hours. Bacterial colonization was found within 1.7% of the 8028 cultures tested from 672 solution units. Only two cultures contained clinically relevant concentrations greater than 5 colony-forming units per milliliter. No significant difference existed between sterility of saline infusion solutions used immediately and those repackaged up to 72 hours prior to use.

STORAGE

The site selected to store medications affects drugs' stability either by shielding drugs from extremes of environment or by increasing the chance that medications will be exposed to those extremes.

Air conditioning and refrigeration, humidifiers and desiccants, and light control can create a stable environment for medications. Short of an ongoing energy source to power climate-controlled storage, all storage systems eventually fail in one or more ways to protect medication from ambient environmental conditions. Furthermore, storage systems can actively damage drugs. For example, storage containers can prolong exposure to high temperatures if they are overinsulated in heated environments, or can create an environment that is too arid when air conditioning is used, destabilizing drugs such as some formulations of epinephrine.

Vehicular storage is convenient in many wilderness and tactical settings. It offers a mobile source of medications and potential power supply for artificial cooling.

McMullan and associates evaluated the potency of midazolam, diazepam, and lorazepam, stored in the air-conditioned cabs of four EMS vehicles in two EMS systems in the southwestern United States.[17] After 120 days of being exposed to an MKT of 32° C (90° F), only the minimal acceptable potency of 90% (95% confidence interval [CI]: 85% to 95%) of lorazepam remained at 90 days. At 120 days, the concentration was below the acceptable level of potency at 86% (95% CI: 81% to 92%).

Air ambulances expose the drugs they carry to extremes that are similar to those of ground ambulances. Madden and associates evaluated ambient and internal temperatures of nylon drug bags carried on EMS helicopters in Texas.[15] Temperatures within the nylon bags failed to comply with the USP recommendations for room temperature of 15° to 30° C (59° to 86° F) on 49% of winter days, 62% of winter nights, 56% of summer days, and 27% of summer nights.

Other cooling and heating methods include ice packs and chemical heat or cold packs. Given the density of water, ice packs are heavy; chemical packs are expensive, given the vast quantities needed in most circumstances.

Nonelectric "coolers" sometimes heat their contents and generally should not be relied on for drug storage. Vehicle windows and the sides of some coolers passively transform solar into thermal energy, acting as solar cookers. Although these storage units might be effective in resisting rising temperatures, when overheated they maintain high temperatures well after the heat of the day. Nonelectric "cooler" insulation neither cooks nor heats, but provides a temporary buffer to temperature change. It lacks capacity for maintaining any temperature for more than a period of minutes to hours; the duration is determined by the specific insulator. An exception to this is the Cambodian Cooler Box, a 24-L chamber of thin galvanized iron covered by cotton sack cloth connected to a top dish holding 9 L of water.[4] At an MKT of 27° C (81° F) in Cambodian drug storerooms, the cooler box reduced the percentage of hours at more than 30° C (86° F) from 4.5% outside the box to 0.1% in the box.

Electrical systems cool more consistently do than nonpowered systems. The Koolatron P9 Traveler III Cooler runs on 12-volt car adapters, weighs 3 kg (7 lb), stores 7 L, and cools to 11° C to 22° C (20° F to 40° F) below the ambient temperature. Its durability is inconsistent. Dison and M-Cool manufacture portable insulin cases that provide 0.1 to 0.2 L of storage, weigh 0.5 to 1.5 kg (1 to 3 lb), last 8 to 24 hours with batteries, and can take AV or AC power. Insulin cases with portable solar systems weighing less than (7 kg) (15 lb) are available from Goal Zero and Instapark.

DuBois studied the temperature of nonelectrically and electrically cooled drug storage boxes on ground EMS vehicles in the Sonoran Desert of California when outside temperatures were 29° to 38° C (84° to 100° F).[8] Temperatures in nonelectrically cooled compartments were often as hot, and occasionally up to 6° C (10° F) hotter than, was ambient air temperature. While the vehicles were immobile, the temperature in electrically cooled compartments remained at approximately 27° C (81° F) when set at 25° C (77° F) to 38° C (100° F). However, when the vehicles were mobile, the temperature in those same compartments occasionally increased to up to 41° C (106° F), exceeding the ambient temperature and manufacturer's maximum suggested storage temperature by 12° C (22° F) for 16 of 17 drugs stored in these EMS vehicles. This indicates the vulnerability to high temperatures of even electrically cooled mobile systems.

All modalities risk uneven temperatures for multiple drugs stored in the same compartment.

DRUGS FOR A BASIC FIELD KIT

A basic field medical kit includes the following types of medications:

Analgesic
Antianaphylactic and antiallergy
Antibiotic
Antiemetic
Antiepileptic
Antipyretic
Sterile fluid (for IV use)

HOW TO READ THE DRUG LIST

The following list summarizes stable conditions for drugs most likely to be included in field or tactical medical kits.[1,6] The list offers options for similar types of drugs, depending on the particular requirements of the users.

Certain terms are used for brevity's sake. *Room temperature* is defined as 15° to 30° C (59° to 86° F). *Controlled room* *temperature* is defined as 20° to 25°C (68° to 77°F). *Excessive heat* is defined as a temperature exceeding 40°C (104°F).

In the United States, availability of medications is subject to regulations of the Food and Drug Administration and Drug Enforcement Agency (DEA). The following labels note drug availability in the United States: OTC (over-the-counter), Rx (prescription required), DEA Schedule (S II, S III, or S IV indicating drugs with abuse potential, with S II having the greatest abuse potential and S IV the least), or NA (not available).

Packaging and inert compounds used with a medication may vary, especially for generic drugs. In all cases, information from the manufacturer should supplement the guide below.

Deviation from the manufacturer's recommendations is the decision of the treating medical professional and not recommended by the authors of this Appendix. Medications are generally listed by their generic names. Mention of trade names does not imply endorsement.

DRUG LIST

ACETAMINOPHEN CAPSULES, TABLETS, ORAL SOLUTION, AND SUPPOSITORIES (OTC)

Store capsules, tablets, and the oral solution at a controlled room temperature. Most are fairly stable in light, moisture, and heat, but high humidity should be avoided for gel-coated capsules. High humidity and light should be avoided for oral-dissolving and chewable tablets. Excessive heat (≥ 40°C [104°F]) should be avoided for extended-release tablets. Solid forms of acetaminophen remain stable for 3 years and liquid forms remain stable for 2 years from the date of manufacture. Store suppositories at 8° to 25°C (46° to 77°F).

ACETAMINOPHEN WITH CODEINE TABLETS AND ORAL SOLUTION (S III)

Store tablets and the solution in light-resistant containers at a controlled room temperature.

ACETAMINOPHEN WITH HYDROCODONE TABLETS AND ORAL SOLUTION (S II)

Store tablets and the solution in light-resistant containers at a controlled room temperature.

ACETAZOLAMIDE TABLETS, EXTENDED-RELEASE CAPSULES, ORAL SOLUTION, AND INJECTION (RX)

Store tablets and extended-release capsules at a controlled room temperature. Brief excursions to 15° to 30°C (59° to 86°F) are permitted for tablets. Dry powder for the injection solution should be stored in an unopened vial at a controlled room temperature. Powder reconstituted with 5 mL sterile water is stable for 12 hours at room temperature, and is stable for 3 days if refrigerated at 2° to 8°C (36° to 46°F).

An extemporaneous formulation can be prepared in three ways:

To prepare a solution of acetazolamide 50 mg/mL, crush 20 acetazolamide 250-mg tablets in 25 mL glycerin or distilled water. Add flavored syrup or 2:1 simple syrup or flavored syrup to bring the total volume to 100 mL. Shake well before use. This solution should be stored under refrigeration and is stable for 1 week.

To prepare a solution of acetazolamide 5 mg/mL, crush two acetazolamide 250-mg tablets in 7 mL polyethylene glycol 400, 53 mL propylene glycol, 15 mL 70% sorbitol solution, 15 mL 85% sucrose solution, 1 mL sweet syrup, 0.5 mL ethanol, and 8 mL of 0.1M citrate to achieve a total volume of 100 mL.

The solution can be prepared in a concentration of acetazolamide 25 mg/mL by crushing 10 acetazolamide 250-mg tablets in 50 mL Ora-Sweet and 50 mL Ora-Plus. Store the solution in an opaque container at room temperature. This solution remains stable for 60 days.

ACETIC ACID OTIC SOLUTION (OTC)

Store the solution in an airtight, light-resistant container at room temperature. Protect from heat.

ALBUTEROL TABLETS, SYRUP, AND INHALED FORMULATION (RX)

Store tablets at 2° to 25°C (36° to 77°F). Store extended-release tablets at 15° to 30°C (59° to 86°F). Store syrup at 2° to 30°C (36° to 86°F). Store capsules for inhalation at room temperature.

For the inhalation route, be certain that albuterol is at room temperature prior to use. For the nebulization route, store albuterol solution for inhalation 0.083% (Proventil), 0.5% (Ventolin), and 0.42% or 0.21% (Accuneb) at 2° to 25°C (36° to 77°F). Accuneb nebulized solution must be used within 1 week after removal from the foil pouch. In the pouch, Ventolin Nebules inhalation solution can be stored at 2° to 8°C (36° to 46°F) for up to 6 months and remains stable at room temperature for 14 days.

Store albuterol aerosol inhalers containing chlorofluorocarbon propellants at room temperature. Store albuterol sulfate aerosol inhalers containing hydrofluoroalkane (HFA) propellants out of direct sunlight at 15° to 25°C (59° to 77°F). To avoid bursting, do not exceed 49°C (120°F). Do not puncture or incinerate. If infrequently used, Ventolin HFA is stable for 6 months from removal from the pouch. Store Ventolin HFA canisters with the mouthpiece down. If frequent nebulization is required, 200 mcg/mL of albuterol sulfate inhalation solution in normal saline remains stable for 7 days at room temperature or under refrigeration when placed in polyvinyl chloride or polyolefin bags, polypropylene syringes and tubes, or borosilicate glass tubes.

ALOE VERA GEL, OINTMENT, AND LAXATIVES (OTC)

Store gel, ointment, and laxatives away from excessive heat and prolonged strong direct light.

AMIODARONE TABLETS, ORAL SOLUTION, INHALANTS, AND INJECTIONS (RX)

Store tablets in a light-resistant container at a controlled room temperature. An extemporaneous 5 mcg/mL formulation can be created by crushing five amiodarone 200-mg tablets into a 200-mL solution of 1:1 of Ora-Plus to Ora-Sweet or Ora-Sweet SF. Solution stored in a glass or plastic bottle under refrigeration remains stable for 91 days. The solution remains stable at room temperature for 6 weeks. Shake before use. Store conventional amiodarone ampules at a controlled room temperature. Ampules may be briefly removed for use at temperatures of 15° to 30°C (59° to 86°F). Protect all injection solutions from light and excessive heat. Do not freeze.

ANTACIDS (OTC)

Store aluminum hydroxide and magnesium hydroxide (often called milk of magnesia) products in tightly sealed containers at a controlled room temperature. Store calcium carbonate conventional tablets at 15° to 30°C (59° to 86°F). Store calcium carbonate chewable tablets below 25°C (77°F). Protect all products from light, moisture, and excessive heat. Do not freeze.

ASPIRIN TABLETS, ORAL SOLUTION, AND SUPPOSITORIES (OTC)

Store tablets and solution in tightly sealed, light-resistant containers at room temperature. Protect from moisture. Store suppositories in the original sealed wrapper at 2° to 15°C (35° to 59°F). Do not freeze. Protect from light, moisture, and excessive heat. Discard aspirin if a strong vinegar odor is present, because potency may be significantly decreased.

ATENOLOL TABLETS (RX)

Store tablets in light-resistant containers at a controlled room temperature.

ATROPINE INJECTION AND OPHTHALMIC SOLUTION (RX)

Store ophthalmic and injection solutions in light-resistant containers at room temperature. In order to prevent contamination, do not touch the applicator tip directly to the eyes or skin. Atropine sulfate 1 mg/mL injection solutions in Tubex (0.5-mL and 1-mL) packaging have been shown to remain stable for 3 months. Atropine methyl nitrate 10 mg/mL solutions have been shown to remain stable for 6 months. Inspect the solution prior to administration for the presence of particulate matter, cloudiness, or discoloration, and discard if present. Do not freeze.

AZITHROMYCIN TABLETS, ORAL SOLUTION, INJECTION, AND OPHTHALMIC SOLUTIONS (RX)

Store tablets at room temperature. Store dry powder for reconstitution below 30° C (86° C). After reconstitution, store suspension at 5° to 30° C (41° to 86° F) and discard after use. After reconstitution, store extended-release solution at a controlled room temperature and use at room temperature. Do not refrigerate or freeze. The solution remains stable for 12 hours. Shake oral azithromycin suspension before use and do not take simultaneously with antacids containing aluminum or magnesium. The injection solution remains stable for 24 hours if stored at 30° C (86° F), or for 7 days if stored under refrigeration below 5° C (41′ F). Store ophthalmic solution in an unopened bottle under refrigeration at 2° to 8° C (36° to 46° F), and at 2° to 25° C (36° to 77° F) once opened. The solution remains stable for 14 days.

BACITRACIN TOPICAL FORMULATION (OTC)

Store the aqueous topical formulation at 2° to 8° C (36° to 46° F) for up to 1 week. Store the nonaqueous topical formulation at room temperature for 3 days, and for longer periods if stored in an anhydrous base, such as lanolin and paraffin.

BISMUTH SUBSALICYLATE TABLETS AND ORAL SOLUTION (OTC)

Store tablets and suspension in tightly sealed containers at room temperature. Protect from direct light and excessive heat. Do not freeze the suspension.

BRETYLIUM TOSYLATE (RX)

Store at a controlled room temperature.

BUPIVACAINE INJECTION (RX)

Store the injection solution at a controlled room temperature. Protect solutions containing epinephrine from light. Bupivacaine hydrochloride 1.25 mg/mL in 0.9% sodium chloride injection solution in disposable polypropylene syringes is stable for 32 days at 3° to 23° C (37° to 73° F).

BUTORPHANOL TARTRATE NASAL SPRAY AND IM AND IV INJECTIONS (S IV)

Store nasal spray at room temperature. Store the injection solution in the original container at 20° to 25° C (68° to 77° F). Protect from light. Discard if discoloration occurs or particulate matter forms in injection solution.

CALCIUM CHLORIDE, CALCIUM GLUCEPTATE, AND CALCIUM GLUCONATE INJECTION (RX)

Store injection solutions of calcium chloride, calcium gluceptate, and calcium gluconate at room temperature. Sterile solutions of calcium in water are indefinitely stable.

CALENDULA TOPICAL FORMULATION (OTC)

Protect from heat, moisture, and direct light.

CEFTRIAXONE INJECTION (RX)

Store dry powder for solution preparation in a light-resistant container at or below 25° C (77° F). Dry powder for injection solutions should not be combined with diluents containing calcium, such as Ringer's or Hartmann's solution, because there will be particulate formulation. After constitution, intramuscular (IM) solutions in water or normal saline remain stable for 2 days at 25° C (77° F) and for 10 days refrigerated at 4° C (39° F) in a concentration of 100 mg/mL; however, at a concentration of 250 mg/mL, such solutions remain stable for only 24 hours at 25° C (77° F) and 3 days refrigerated at 4° C (39° F). IV solutions at concentrations of 10, 20, and 40 mg/mL remain stable for 2 days at 25° C (77° F) and for 10 days refrigerated at 4° C (39° F). Do not refrigerate injection solutions that contain 5% dextrose and 0.9% or 0.45% sodium chloride diluent solutions. IV solutions of ceftriaxone that contain 5% dextrose and 0.9% sodium chloride solution can be frozen at −20° C (−4° F) in PVC or polyolefin containers and remain stable for 26 weeks. Thaw at room temperature before use, and discard any unused, thawed solution.

CEPHALEXIN CAPSULES, TABLETS, AND ORAL SOLUTION (RX)

Store capsules at room temperature. The suspension is stable for 14 days under refrigeration.

CHARCOAL, ACTIVATED (OTC)

Store activated charcoal in an airtight container. Sealed aqueous suspensions are stable for 1 year.

CIPROFLOXACIN TABLETS, CAPSULES, ORAL SOLUTION, INJECTION, OPHTHALMIC SOLUTION, AND OTIC SOLUTIONS (RX)

Store tablets below 30° C (86° F). Store extended-release tablets at 25° C (77° F). Brief excursions are permitted at room temperature. Store microcapsules and diluent for oral suspensions below 25° C (77° F). Do not freeze. After reconstitution, the solution should be stored below 30° C (86° F); it remains stable for 14 days. Store the ophthalmic solution in original vials at 2° to 25° C (36° to 77° F). Protect from light and excessive heat. Do not freeze tablets or oral and ophthalmic solutions. Store the otic solution in a light-resistant container at room temperature of 15° to 25° C (59° to 77° F).

CROTALIDAE ANTIVENOM (RX)

Store vials at 2° to 8° C (36° to 46° F). Do not freeze. Use within 4 hours of reconstitution.

CYCLOPENTOLATE HYDROCHLORIDE OPHTHALMIC SOLUTION (RX)

Store ophthalmic solution in the original container at room temperature. Use only if the sealing neckband on the container is intact.

DABIGATRAN TABLETS (RX)

Store in a tightly sealed container at 25° C (77° F). Brief excursions are permitted at room temperature. Protect from moisture. Once the container has been opened, use within 4 months.

DEET (*N,N*-DIETHYL-*META*-TOLUAMIDE, DIETHYLTOLUAMIDE)–CONTAINING INSECT REPELLENT (OTC)

Store the repellent below 49° C (120° F). Store away from heat and flame.

DERMABOND (2-OCTYL CYANOACRYLATE) TOPICAL SKIN ADHESIVE (RX)

Store the adhesive below 30°C (86°F). Discard if the package is open or has been tampered with. Discard the excess after use because the adhesive hardens on exposure to air. Protect from moisture and direct heat.

DEXAMETHASONE TABLETS AND ORAL, INJECTION, IMPLANTATION, INTRAVITREAL, AND OPHTHALMIC SOLUTIONS (RX)

Store tablets in a light-resistant container at a controlled room temperature. Protect from moisture. Store the oral solution in the original bottle and only dispense with the supplied calibrated dropper at a controlled room temperature. Once opened, the oral solution remains stable for 90 days. Discard if precipitation forms. Store the implantation, intravitreal, and ophthalmic solutions at room temperature. Extemporaneous formulations remain stable for 91 days.

DEXTROAMPHETAMINE TABLETS, CAPSULES, AND ORAL SOLUTION (S II)

Store non–extended-release capsules and tablets at room temperature. Store extended-release capsules and tablets at a controlled room temperature. Store the elixir in an airtight, light-resistant container at room temperature.

DEXTROSE ORAL SOLUTION (OTC) AND INJECTION (RX)

Store oral solution in a well-filled, airtight container. For injection, do not exceed 25°C (77°F). Do not freeze or expose to extreme heat. Discard if cloudy prior to use and discard any unused portions once open.

DIAZEPAM TABLETS, ORAL SOLUTION, SUPPOSITORIES, AND INJECTION (S IV)

Store tablets, oral solution, and suppositories at room temperature. Protect from light, heat, and moisture. Do not freeze the oral solution. Suppositories are stable for 8 months at 40°C (104°F) and can withstand at least three freeze-thaw cycles. Brief excursions are permitted to room temperature. Store the injection solution at a controlled room temperature. Do not refrigerate.

DIGOXIN TABLETS AND INJECTION (RX)

Store at a controlled room temperature. Brief excursions are permitted to room temperature. Protect from light. Protect tablets from moisture.

DILTIAZEM TABLETS, ORAL SOLUTION, AND INJECTION (RX)

Store tablets at 25°C (77°F). Brief excursions are permitted to 15° to 30°C (59° to 86°F). Avoid excess humidity. An extemporaneous formulation of a 1-mg/mL solution can be prepared using 250 mg diltiazem (2.5 mL of diltiazem hydrochloride stock solution) combined with dextrose, fructose, mannitol, sorbitol, or sucrose to a volume of 250 mL. A solution of 12-mg/mL diltiazem can be prepared by crushing 16 tablets of 90-mg diltiazem in 10 mL of 1:1 mixtures of Ora-Plus with either Ora-Sweet or Ora-Sweet SF or in 1:4 mixtures of flavored syrup with simple syrup and then bringing the solution to a total volume of 120 mL. Protect from light.

DIPHENHYDRAMINE TABLETS, ORAL SOLUTION (OTC), AND INJECTION (RX)

Store at a controlled room temperature in a light-resistant container. Do not freeze oral and injection solutions.

DOMEBORO (ACETIC ACID AND ALUMINUM ACETATE) OTIC SOLUTIONS (OTC)

Store otic solutions in a tightly sealed container at either room temperature or under refrigeration. Protect from direct light, heat, and moisture. Do not freeze.

DOPAMINE HYDROCHLORIDE INJECTION (RX)

Store the injection in a light-resistant container. Discard if the injection has yellow-brown discoloration or if pH outside of the 4.0 to 6.4 range is detected, because these are indications of decomposition. Dopamine 6.4 mg/mL in 5% dextrose injection is stable at a controlled room temperature for up to 24 hours in ambient humidity and in the presence of light.

DOXYCYCLINE CAPSULES, TABLETS, ORAL SOLUTION, AND INJECTION (RX)

Store capsules and tablets in light-resistant containers at room temperature. Store doxycycline hyclate delayed-release tablets in light-resistant containers at a controlled room temperature. Brief excursions are permitted at room temperature. Store lyophilized powder in a light-resistant container at a controlled room temperature. Refrigerate in a light-resistant container immediately after reconstitution, or dilute the injection solution to 0.1 to 1 mg/mL within 12 hours after reconstitution, where it will remain stable for up to 48 hours at 25°C (77°F) and 72 hours at 4°C (39°F). Avoid direct sunlight during storage and infusion. Infusions of doxycycline made with lactated Ringer's or 5% dextrose in lactated Ringer's diluents must be used within 6 hours of reconstitution to ensure stability. Solutions of 10 mg/mL doxycycline in sterile water can be frozen and stored at −20°C (−4°F) and remain stable for up to 8 weeks. Avoid excess heat after thawing and discard any unused thawed solution.

EDOXABAN TABLETS (RX)

Store at a controlled room temperature. Brief excursions are permitted at room temperature.

EMLA (LIDOCAINE/PRILOCAINE) TOPICAL FORMULATION (RX)

Store EMLA at room temperature. Do not freeze. Discoloration does not necessarily indicate lack of stability. Precipitate indicates that the solution is not stable.

EPINEPHRINE INJECTION AND TOPICAL, INHALED, AND INTRANASAL FORMULATIONS (RX)

Store injection ampules at 5° to 25°C (41° to 77°F). Do not freeze. Injection ampules stored at 38°C (100°F) will last less than 3 months at low humidity (15%) and less than 4 months at high humidity (85%). An extemporaneous formulation of a topical anesthetic solution can be prepared with 2.25 mg/mL of racemic epinephrine hydrochloride, 40 mg/mL of lidocaine hydrochloride, 5 mg/mL of tetracaine hydrochloride, and 0.63 mg/mL of sodium metabisulfite. Store this topical solution in a light-resistant container at 18°C (64°F) for no more than 4 weeks, and at 4°C (39.2°F) for up to 26 weeks. Store the epinephrine inhaler at a controlled room temperature. Do not exceed 49°C (120°F). Do not puncture or incinerate the inhaler. Store the intranasal solution in a light-resistant container at 15° to 25°C (59° to 77°F). Do not freeze.

ERYTHROMYCIN TABLETS, ORAL SOLUTION, AND TOPICAL OINTMENT (RX)

Store tablets and oral solution at less than 30°C (86°F). Reconstituted granules must be used within 10 days. Reconstituted erythromycin ethyl succinate solution must be used within 14 days if kept at room temperature. Reconstituted EryPed solution should be stored at less than 25°C (77°F) and used within 35 days. Refrigeration of the suspension is encouraged for the best taste. Optimal stability is maintained at pH above 6.0, with significant decomposition at or below pH of 4.0. Store the topical ointment at less than 27° C (81° F).

FAMOTIDINE TABLETS (OTC) AND INJECTION (RX)

Store regular and chewable tablets at a controlled room temperature. Brief excursions for chewable tablets to room temperature are permitted. Protect from moisture. Store injection vials in a light-resistant container at 2° to 8° C (36° to 46° F).

FENTANYL ORAL LOZENGES, SUBLINGUAL TABLETS, SUBLINGUAL SPRAY, BUCCAL FILM, INJECTION, AND INTRANASAL FORMULATION (RX)

Store oral lozenges, sublingual tablets, sublingual spray, and buccal film at a controlled room temperature. Brief excursions to room temperature are permitted. Protect from moisture. Do not freeze. Store the injection solution in a light-resistant container at a controlled room temperature. Store the intranasal canister in a light-resistant container at 2° C to 25° C (36° F to 77° F).

FLUOCINOLONE ACETONIDE TOPICAL OINTMENT, OTIC SOLUTION, AND SHAMPOO (RX)

Store topical cream, ointment, and shampoo at room temperature. Do not freeze. Store the otic solution at a controlled room temperature.

FURAZOLIDONE TABLETS AND SOLUTION (NA)

Store tablets and liquid in light-resistant containers. Tablets can be crushed and administered with a spoonful of corn syrup. Exposure to strong light may cause darkening.

FUROSEMIDE TABLETS, SOLUTION, AND INJECTION (RX)

Store tablets and solution in light-resistant containers at 25° C (77° F). Brief excursions are permitted to 15° to 30° C (59° to 86° F). Protect from moisture. Store the injection solution in a light-resistant container at room temperature. Discard all types of furosemide if discoloration occurs.

GLUCAGON INJECTION (RX)

Store dry powder in a light-resistant container at a controlled room temperature. Do not freeze. Powder remains stable for 24 months. Use the injection solution immediately after reconstitution and discard unused portions.

HALOPERIDOL TABLETS AND INJECTION (RX)

Store tablets in a tightly closed, light-resistant container at a controlled room temperature. Store the injection solution in a light-resistant container at room temperature. Do not freeze.

HYDROCORTISONE TABLETS, SOLUTION, INJECTION, AND TOPICAL CREAM (RX)

Store tablets, oral solution, injection, and topical cream at room temperature in the original container. Protect from light, moisture, and heat. Do not freeze the oral solution or injections.

HYDROMORPHONE TABLETS, SOLUTION, SUPPOSITORIES, AND INJECTION (S II)

Store tablets, solution, suppositories, and injectables in light-resistant containers at a controlled room temperature. Excursions are permitted to 15° to 30° C (59° to 86° F). Slight yellow discoloration of the injection liquid does not affect potency.

IBUPROFEN TABLETS AND SOLUTION (OTC)

Store the tablets at a controlled room temperature and the solution at room temperature.

INSULIN (REGULAR) INJECTION AND INHALED FORMULATION (RX)

Store the subcutaneous and IV injections in a light-resistant container refrigerated at 2° to 8° C (36° to 46° F). Do not freeze. Store the open vials at room temperature for up to 31 days. Store inhalers refrigerated at 2° to 8° C (36° to 46° F). Store at room temperature for up to 10 days. Discard unused cartridges from an open blister pack strip after 3 days.

INTRAVENOUS SOLUTIONS (D_5W, NS, LR, D_5NS, AND OTHER ADMIXTURES)

Store solutions below 90° C (194° F) and preferably at room temperature for ease of use. Pure sodium chloride and lactated Ringer's solutions at concentrations used in medicine are unlikely to show precipitation at 0° C (32° F), or if frozen for 3 months.

ISOPROTERENOL HYDROCHLORIDE INHALANT AND INJECTION (RX)

Store the inhalation solution and injection in light-resistant containers at room temperature. The injection is stable indefinitely in normal saline. Avoid excessive heat. Discard inhalation or injection solution if pink or brown discoloration or precipitation occurs. Store isoproterenol (5 mg/L) in 5% dextrose in water at room temperature. This solution remains stable for 24 hours.

IVERMECTIN TABLETS (RX)

Store below 30° C (86° F).

KALETRA (LOPINAVIR/RITONAVIR) TABLETS (RX)

Store tablets at a controlled room temperature. Brief excursions are permitted at room temperature. Once the tablet container is opened or tablets are exposed to high humidity, tablets remain stable for up to 2 weeks.

KETOCONAZOLE TABLETS, SHAMPOO, FOAM, AND GEL (RX)

Store tablets in light-resistant containers at a controlled room temperature. Protect from heat and moisture. Store the shampoo in a light-resistant container below 25° C (77° F). Store the foam in a light-resistant container at a controlled room temperature. Do not refrigerate. Avoid direct light. Store the foam at a controlled room temperature with excursions permitted to room temperature.

Store drops in a tightly sealed container below 30° C (86° F).

LACOSAMIDE TABLETS, ORAL SOLUTION, AND INJECTION (RX)

Store at a controlled room temperature. Brief excursions are permitted at room temperature. The oral solution remains stable for up to 7 weeks after the bottle has been opened. Do not freeze. Once the injection is diluted, store at room temperature for up to 4 hours.

LACRISERT (HYDROXYPROPYL METHYLCELLULOSE) OPHTHALMIC SOLUTION (RX)

Store drops in a tightly sealed container below 30° C (86° F).

LEMON GRASS *(CYMBOGOGON)* CITRONELLA OIL TOPICAL FORMULATION (OTC)

Store at room temperature. Protect from heat, moisture, and direct light.

LEVETIRACETAM TABLETS, ORAL SOLUTION, AND INJECTION (RX)

Store immediate-release tablets, extended-release tablets, and oral solution at 25° C (77° F). Brief excursions are permitted at room temperature. The injection diluted in solution in a polyvinyl chloride bag is stable for at least 24 hours. Discard the unused portion of the vial after opening.

LEVOFLOXACIN TABLETS, SOLUTION, INJECTION, AND OPHTHALMIC FORMULATION (RX)

Store tablets at 15° to 30° C (59° to 86° F). Store the oral solution at 25° C (77° F). Brief excursions are permitted at 15° to 30° C (59° to 86° F). The injection solution remains stable for 72 hours if stored at or below 25° C (77° F). Injection solution can be diluted in plastic or glass containers, and then frozen at −20° C (−4° F), where it remains stable for up to 6 months. Thaw slowly (no hot water baths or microwaves) at 25° C (77° F) or under refrigeration at 8° C (46° F). Use immediately after thawing. Do not refreeze. Store flexible containers of premixed solutions in a light-resistant container at or below 25° C (77° F). Avoid excessive heat and do not freeze. Store levofloxacin 0.5% and 1.5% ophthalmic solutions at 15° to 25° C (59° to 77° F).

LIDOCAINE INJECTION AND TOPICAL, INTRADERMAL, AND OPHTHALMIC SOLUTIONS (RX)

Store injection solution in a light-resistant container at room temperature. Do not freeze. Do not reuse “one-time-use” injection bottles, because they lack methylparaben preservative. Store the topical gel and jelly at a controlled room temperature. Store the viscous topical preparation, topical patches, and intradermal powder in sealed original packaging at room temperature at 15° to 30° C (59° to 86° F).

LIDOCAINE/EPINEPHRINE/TETRACAINE (LET) TOPICAL SOLUTION (RX)

Store solution in a light-resistant container. The solution remains stable at 18° C (64° F) for 4 weeks, and at 4° C (39° F) for 26 weeks.

LINDANE (GAMMA-HEXACHLOROCYCLOHEXANE) LOTION AND SHAMPOO (RX)

Store lotion and shampoo at a controlled room temperature.

LOPERAMIDE HYDROCHLORIDE CAPSULES (OTC)

Store capsules at 15° to 25° C (59° to 77° F). Placing the contents of 10 of the 2-mg capsules in hard fat, such as suet, leaf lard, or fatback lard, and rolling into shape can also create rectal suppositories of 20 mg loperamide.

LORAZEPAM TABLETS, ORAL SOLUTION, AND INJECTION (S IV)

Store tablets in a tightly sealed container at a controlled room temperature. Store oral solution at 2° to 8° C (36° to 46° F). Discard an opened bottle after 90 days. Store IM and IV solutions in light-resistant containers at 2° to 8° C (36° to 46° F).

MALARONE (ATOVAQUONE/PROGUANIL) TABLETS (RX)

Store in a light-resistant container at a controlled room temperature. Brief excursions to room temperature are permitted.

MANNITOL INJECTION (RX)

Store vials of mannitol solution and powder for reconstitution at a controlled room temperature. Discard the unused portion of the solution. Concentrations of 15% or more may crystallize when exposed to lower temperatures. To resolubilize crystals, place the vial in a heated water bath at 60° to 80° C (140° to 176° F) and shake occasionally. Using a microwave is not recommended, because the vial is likely to explode. Cool to room temperature before use. Do not heat the solution if a white flocculent precipitate forms after contact with PVC, because crystals will re-form rapidly.

MEBENDAZOLE TABLETS (RX)

Store at 15° to 25° C (59° to 77° F).

MEPERIDINE HYDROCHLORIDE TABLETS, ORAL SOLUTION, AND INJECTION (S II)

Store tablets and the oral solution at a controlled room temperature. Brief excursions are permitted to 15° to 30° C (59° to 86° F). Store the injection solution in a light-resistant container at a controlled room temperature.

METOPROLOL TABLETS, ORAL SOLUTION, AND INJECTION (RX)

Store tablets at a controlled room temperature. Brief excursions are permitted to room temperature. An extemporaneous oral suspension solution can be created by combining 12 crushed 100-mg metoprolol tablets with a small amount of Ora-Sweet, Ora-Sweet SF, or Ora-Plus and bringing the volume to 120 mL with water. The suspension remains stable for 60 days under refrigeration. Shake well before use. Store the injection ampules in tight, light-resistant, moisture-free containers at a controlled room temperature.

METRONIDAZOLE CAPSULES, TABLETS, AND INJECTION (RX)

Store capsules at 15° to 25° C (59° to 77° F). Store extended-release tablets at a controlled room temperature. Brief excursions are permitted to room temperature. Store the injection solution in a light-resistant container at room temperature.

MIDAZOLAM ORAL SOLUTION AND INJECTION (S IV)

Store oral solution at a controlled room temperature. Brief excursions are permitted to 15° to 30° C (59° to 86° F). Store injection solution at a controlled room temperature. The injection solution may be stored for at least 28 days at 3° to 25° C (37° to 77° F).

MODAFINIL TABLETS (S IV)

Store tablets at a controlled room temperature.

MORPHINE SULFATE TABLETS, EPIDURAL SUSPENSION, AND INJECTION (S II)

Store tablets in light-resistant containers at a controlled room temperature. Excursions are permitted to room temperature.

Store the epidural extended-release suspension under refrigeration at 2° to 8° C (36° to 46° F). Do not freeze. Unopened vials remain stable for 30 days at a controlled room temperature. Do not return vials to the refrigerator once they have been stored at room temperature. Solution withdrawn from the vial can be stored at room temperature for up to 4 hours prior to administration. After that, all withdrawn solution should be discarded.

Store injection solution in the original carton at a controlled room temperature. Brief excursions are permitted to 15° to 30° C (59° to 86° F). Do not freeze. Discard any unused solution.

Pain cocktails containing preservatives without alcohol or chloroform water will remain stable for 3 weeks after compounding.

MOXIFLOXACIN TABLETS, ORAL SOLUTION, INJECTION, AND OPHTHALMIC ROUTE (RX)

Store tablets and injection solution at a controlled room temperature. Brief excursions are permitted to 15° to 30°C (59° to 86°F). Do not refrigerate injection solution because precipitate forms. Extemporaneous oral suspension can be formed to create 60 mL of 20 mg/mL moxifloxacin hydrochloride by combining three crushed 400-mg tablets with 30 mL of Ora-Plus, Ora-Sweet, or Ora-Sweet SF. When stored in a light-resistant amber plastic bottle, oral suspension remains stable for 90 days if stored at 23° to 25°C (73° to 77°F). Store 0.5% moxifloxacin ophthalmic solution at 2° to 25°C (36° to 77′ F).

MUPIROCIN TOPICAL FORMULATION (RX)

Store cream and ointment at a controlled room temperature. Do not freeze cream.

NALBUPHINE HYDROCHLORIDE INJECTION (RX)

Store injection solution in a light-resistant container at a controlled room temperature.

NALOXONE HYDROCHLORIDE INJECTION (RX)

Store injection solution ampules and vials in original containers at a controlled room temperature. Use infusion solutions within 24 hours of opening. For Evzio, store between 15° and 25°C (59° and 77°F). Brief excursions are permitted to 4° to 40°C (39° to 104°F).

NEOSPORIN OINTMENT (OTC)

Store ointment in the original container with the cap tightly sealed at room temperature. Protect from light, moisture, and heat.

NIFEDIPINE CAPSULES, TABLETS, ORAL SOLUTION, AND INJECTION (RX)

Store capsules in a light-resistant container at 15° to 22° (59° to 77°F). Store tablets in a light-resistant container below 30°C (86°F). An extemporaneous formulation of an oral solution can be made by combining five nifedipine 10-mg tablets and soaking them in a small amount of 1% hypromellose for 5 minutes and then bringing the total volume to 50 mL with 1% hypromellose. Package the 1-mg/mL extemporaneous suspension in single-dose syringes stored in opaque black plastic bags. Solution remains stable for 28 days at 6° or 22°C (43° or 72°F).

Store injection solution in a light-resistant container below 25°C (77°F). Because the infusion is extremely light sensitive, the solution retains its potency for 1 hour in daylight and 6 hours in artificial light. Do not remove the vial from the container until immediately before use.

NITROGLYCERIN CAPSULES, SUBLINGUAL TABLETS AND SPRAYS, INJECTION, PATCHES, AND TOPICAL FORMULATION (RX)

Store capsules at room temperature. Store sublingual tablets and sprays at a controlled room temperature. Protect tablets from moisture. Sprays may have brief excursions to room temperature. Store concentrated nitroglycerin for injection solution in a light-resistant container at room temperature. Injection solutions in polyolefin containers can be stored at room temperature for at least 24 hours. Premixed nitroglycerin in either normal saline or 5% dextrose can be stored for 48 hours at room temperature and 7 days under refrigeration. The extemporaneous formulation of solutions with a concentration of 0.035 to 1 mg/mL in glass containers remains stable for 70 days at room temperature and 6 months under refrigeration. Store transdermal patches at room temperature. Store topical ointment at a controlled room temperature.

NORFLOXACIN TABLETS, ORAL SOLUTION, AND OPHTHALMIC SOLUTION (RX)

Store tablets at a controlled room temperature in tightly sealed containers. Brief excursions to room temperature are permitted. Extemporaneous oral solution can be created by crushing three 400-mg tablets into a small amount of Ora-Plus and flavored syrup to taste and bringing the total volume to 60 mL to create a 20-mg/mL solution. Under experimental conditions, the suspension remains stable (containing ≥ 93% norfloxacin) for at least 56 days at a temperature of 23° to 25°C (73.4° to 77°F) or under refrigeration at 3° to 5°C (37.4° to 41°F). Store the ophthalmic solution at room temperature.

OFLOXACIN TABLETS, INJECTION, OPHTHALMIC SOLUTION, AND OTIC SOLUTION (RX)

Store tablets in a tightly sealed container below 30°C (86°F). Store single-use vials and premixed bottles of injection solution in light-resistant containers at room temperature. Brief exposure to temperatures up to 40°C (104°F) are permitted. Do not freeze. In diluted concentrations between 0.4 and 4 mg/mL and stored in a glass or plastic container, solution remains stable for 14 days under refrigeration at 5°C (41°F), or for 6 months frozen at −20°C (−4°F). Solution will remain stable for up to 14 days under refrigeration at 2° to 8°C (36° to 46°F) after thawing. Do not use hot water or a microwave oven for rapid thawing. Store ophthalmic and otic solutions at 15° to 25°C (59° to 77°F).

PENICILLIN G PROCAINE INJECTION (RX)

Store at 2° to 8°C (36° to 46°F). Avoid freezing. Injection is stable for 7 days at 25°C (77°F) and 1 day at 40°C (104°F). Wycillin remains stable for 6 months if stored at room temperature.

PENICILLIN GK AND G SODIUM INJECTION (RX)

Store penicillin GK vials at a controlled room temperature. Once they have been diluted, refrigerate for up to 7 days. Once prepared, penicillin G solutions remain stable and free from allergenic components for 24 hours at room temperature or under refrigeration. At a concentration of 40 million units/L, more than 90% potency was retained for 1 month for penicillin GK, and for 39 days for penicillin G when stored in PVC containers at −20°C (−4°F), and for 70 days for penicillin G under refrigeration.

PHENOBARBITAL TABLETS, SOLUTION, AND IM AND IV INJECTIONS (S IV)

Store tablets, oral solution, and IM and IV injection solutions in tightly sealed light-resistant containers at a controlled room temperature. Protect oral solution and tablets from moisture. Slight discoloration is allowable. Discard the solution if there is more discoloration or any precipitation.

PHENYLEPHRINE INJECTION AND OPHTHALMIC SOLUTION (RX) AND NASAL SPRAY (OTC)

Store injection solution in a light-resistant container at a controlled room temperature. Brief excursions to room temperature are permitted. Once it has been diluted, the solution is stable for 4 hours at room temperature and 24 hours if refrigerated. Store nasal spray in light-resistant containers at room temperature. Refrigerate ophthalmic solution. Discard all forms of phenylephrine if brown discoloration occurs or a precipitate forms.

PHENYTOIN CAPSULES, TABLETS, ORAL SOLUTION, AND INJECTION (RX)

Store capsules, tablets, and oral solution at a controlled room temperature. Keep extended-release tablets and oral solution in a light-resistant container. Do not freeze oral solution. Store phenytoin sodium injection solution at a controlled room temperature. The solution is usable while clear or faintly yellow. Discard if the solution becomes hazy or if a precipitate forms and persists at room temperature. Because phenytoin is more stable in saline than in dextrose, use or discard phenytoin in 5% dextrose solution within 2 hours of mixing.

POLYSPORIN OINTMENT (RX)

Store at room temperature. Do not freeze.

POTASSIUM PERMANGANATE ASTRINGENT SOLUTION (OTC)

Store solution in a tightly sealed container at 15° to 30° C (59° to 86° F).

POVIDONE-IODINE SOLUTION (OTC)

Store solution at a controlled room temperature. Brief excursions are permitted to 15° to 30° C (59° to 86° F).

PREDNISONE TABLETS AND ORAL SOLUTION (RX)

Store tablets and oral solution at a controlled room temperature. Brief excursions to room temperature are permitted. Extemporaneous formulations should be stored at room temperature or under refrigeration, and will remain stable for 1 to 2 months.

PROCHLORPERAZINE CAPSULES, TABLETS, ORAL SOLUTION, AND INJECTION (RX)

Store capsules, tablets, and solution in tightly closed light-resistant containers at room temperature. Slight yellow discoloration is acceptable. Discard if more discoloration develops. If preparing an IV admixture, use it immediately or dissolve the prochlorperazine in a dextrose solution and store under refrigeration in a light-resistant container. Prochlorperazine 5 mg/mL or 10 mg/2 mL retained 100% potency when it was stored at room temperature in Tubex containers for 3 months.

PROMETHAZINE CAPSULES, TABLETS, SOLUTION, INJECTION, AND SUPPOSITORIES (RX)

Store capsules, tablets, and oral and injection solutions in a light-resistant container at a controlled room temperature. Light pink discoloration of white promethazine tablets does not indicate a significant loss of potency. Discard the solution if color or precipitate develops. Refrigerate suppositories. Suppositories remain stable at room temperature for 2 weeks, and under refrigeration for weeks.

PSEUDOEPHEDRINE AND PSEUDOEPHEDRINE/ TRIPROLIDINE CAPSULES AND TABLETS (OTC)

Store capsules and tablets in light-resistant containers at 15° to 25° C (59° to 77° F). Protect them from moisture.

RIVAROXABAN TABLET (RX)

Store at 25° C (77° F). Brief excursions are permitted to room temperature.

ROCURONIUM INJECTION (RX)

Store at 2° to 8° C (36° to 46° F). Do not freeze. Injection solutions can be stored at a controlled room temperature for 60 days. Open vials should be used within 30 days.

SILDENAFIL TABLETS (RX)

Store tablets at a controlled room temperature. Brief excursions are permitted to room temperature.

SIMETHICONE CAPSULES, TABLETS, DROPS, AND ULTRASOUND SUSPENSION (OTC)

Store capsules, tablets, and drops in a light-resistant container below 40° C (104° F), and preferably at room temperature. Do not freeze.

SODIUM BICARBONATE TABLETS, INJECTION, AND SUPPOSITORIES (RX)

Store tablets at room temperature. Do not refrigerate. Store injection solution at a controlled room temperature in an airtight container to stop the solution from changing to sodium carbonate. Brief exposure to 40° C (104° F) does not affect stability or potency.

SODIUM SULFACETAMIDE TABLETS, CREAM, LOTION, OINTMENT, AND OPHTHALMIC ROUTE (RX)

Store tablets and cream in light-resistant containers at 15° to 30° C (59° to 86° F). Do not freeze vaginal cream. Store 10% sulfacetamide topical lotion and ointment at room temperature. The lotion will remain stable for 4 months. Do not freeze. Store ophthalmic solution in a light-resistant container at 8° to 15° C (46° to 59° F). Discard if it becomes darkened.

SUCCINYLCHOLINE INJECTION (RX)

Store at 2° to 8° C (36° to 46° F). The injection solution is stable at a controlled room temperature for 14 days. Once it has been diluted, discard within 24 hours.

TEMAZEPAM CAPSULES (S IV)

Store capsules in light-resistant containers below 30° C (86° F). Protect from moisture.

TETANUS TOXOID, TETANUS TOXOID/ DIPHTHERIA/ACELLULAR PERTUSSIS, AND HYPERIMMUNE TETANUS GLOBULIN VACCINE SOLUTIONS (RX)

Store vaccine solutions at 2° to 8° C (36° to 46° F). Do not freeze. The solutions are stable for 72 hours at a controlled room temperature.

TETRACAINE HYDROCHLORIDE OPHTHALMIC SOLUTION (RX)

Store ampules in light-resistant containers at 2° to 8° C (35.6° to 46.4° F) to prevent oxidation and crystallization. Tetracaine hydrochloride remains stable for 3 days at room temperature, and retains the original manufacturer's expiration date if returned to refrigeration. For topical "LET" solution information, see the Lidocaine/Epinephrine/Tetracaine entry.

TETRACYCLINE CAPSULES, TABLETS, ORAL SOLUTION, INJECTION, AND TOPICAL OINTMENT (RX)

Store capsules, tablets, oral solution, and topical ointment in light-resistant containers at room temperature. Reconstituted solutions are stable for 12 hours, and tetracycline hydrochloride is stable in 5% dextrose and water for 6 hours. Do not use outdated products, because they may cause proximal renal tubular acidosis and Fanconi's syndrome.

TOLNAFTATE TOPICAL ANTIFUNGAL SOLUTION (OTC)

Store topical solution at room temperature. Solidification may occur at lower temperatures, but the solution reliquefies easily when warmed.

TRIAZOLAM TABLETS (S IV)

Store tablets at a controlled room temperature.

TRIMETHOPRIM/SULFAMETHOXAZOLE (80 MG/400 MG) TABLETS, ORAL SOLUTION, AND INJECTION (RX)

Store tablets, oral solution, and unopened injection vials at a controlled room temperature. Protect tablets from moisture. Store the oral solution in a light-resistant container. Injection solution, including 80 mg trimethoprim in 100 mL D_5W, is stable for 4 hours, but will last longer if it is more dilute. Vials drawn into a polypropylene syringe will remain stable for 60 hours. Do not refrigerate. Do not inject intramuscularly. Discard if cloudiness or precipitation develops.

TRUVADA (EMTRICITABINE/TENOFOVIR) TABLETS (RX)

Store in a tightly closed container at 25° C (77° F). Brief excursions are permitted at room temperature.

VERAPAMIL HYDROCHLORIDE CAPSULES, TABLETS, AND INJECTION SOLUTION (RX)

Store verapamil sustained-release and all immediate-release tablets in a light-resistant container at 15° to 25° C (59° to 77° F). Immediate-release tablets remain stable for 3 years. Protect all tablets and capsules from moisture. An extemporaneous oral suspension of 50 mg/mL verapamil can be created from 20 of the 80-mg verapamil tablets in a 1 : 1 mixture of Ora-Plus with Ora-Sweet, Ora-Sweet SF, or flavored syrup mixture (1 : 4 concentrated flavoring to simple syrup). When stored in light-resistant amber polyethylene terephthalate bottles, the solution retains 91% potency for 60 days at 25° C (77° F) or under refrigeration at 5° C (41° F).

Store verapamil hydrochloride powder and premixed vials in a light-resistant container at room temperature. Protect from moisture. Discard unused portions of the injection solution.

WARFARIN TABLETS (RX)

Store in a light-resistant container at room temperature. Protect from moisture.

ZINC SALTS (OTC)

Store zinc salts in an airtight, nonmetallic container. An extemporaneous formulation of an oral solution can be made up for zinc sulfate by combining 22 g of zinc sulfate powder with 250 mL of flavored syrup and bringing the total volume to 500 mL with purified water. The solution of 10 mg/mL zinc remains stable for 60 days under refrigeration, or for 12 months after addition of a paraben concentrate for a final zinc concentration of 0.5%.

ZOLPIDEM TABLETS, SUBLINGUAL TABLETS, AND SPRAY (S IV)

Store sublingual, immediate-release, and extended-release tablets and oral spray in a light-resistant container at a controlled room temperature. Protect from light and moisture. Brief excursions are permitted to temperatures of 15° to 30° C (59° to 86° F). Do not freeze.

REFERENCES

> ***Complete references used in this text are available online at expertconsult.inkling.com.***

REFERENCES

1. DRUGDEX System (Micromedex 2.0). Greenwood Village, Colorado: Truven Health Analytics; 1974–2015.
2. Bartfield JM, Homer PJ, Ford DT, et al. Buffered lidocaine as a local anesthetic: an investigation of shelf life. Ann Emerg Med 1992;21:16.
3. Bianchi F, Ginggen A, Tardy Y. Stability and compatibility of drug mixtures in an implantable infusion system. Anaesthesia 2008;63:972.
4. Chanthap L, Ariey F, Socheat D, et al. Low-technology cooling box for storage of malaria RDTs and other medical supplies in remote areas. Malar J 2010;9:31.
5. Church WH, Hu SS, Henry AJ. Thermal degradation of injectable epinephrine. Am J Emerg Med 1994;12:306.
6. Cohen V, Jellinek SP, Teperikidis L, et al. Room-temperature storage of medications labeled for refrigeration. Am J Health Syst Pharm 2007; 64:1711.
7. Donyai P. Quality of medicines stored together in multi-compartment compliance aids. J Clin Pharm Ther 2010;35:533.
8. Dubois WC. Drug storage temperatures in rescue vehicles. J Emerg Med 2000;18:345.
9. Gammon DL, Su S, Jordan J, et al. Alteration in prehospital drug concentration after thermal exposure. Am J Emerg Med 2008;26:566.
10. Gill MA, Kislik AZ, Gore L, Chanda A. Stability of advanced life support drugs in the field. Am J Health Syst Pharm 2004;16:597.
11. Grant TA, Carroll RG, Church WH, et al. Environmental temperature variations cause degradations in epinephrine concentration and biological activity. Am J Emerg Med 1994;12:319.
12. Helm M, Castner TH, Lampl L. Environmental temperature stress on drugs in prehospital emergency medical service. Acta Anaesthesiol Scand 2003;47:425.
13. Hoye WL, Mogalian EM, Myrdal PB. Effects of extreme temperatures on drug delivery of albuterol sulfate hydrofluroalkane inhalation aerosols. Am J Health Syst Pharm 2005;62:2271.
14. Levin HJ, Fieber RA, Levi RS. Stability data for Tubex filled by hospital pharmacists. Hosp Pharm 1973;8:310.
15. Madden JF, O'Connor RE, Evans J. The range of medication storage temperatures in aeromedical emergency medical services. Prehosp Emerg Care 1999;1:27.
16. Mathijssen NMC, Petit PLC, Pilot P, et al. Impregnation of bone chips with antibiotic and storage of antibiotics at different temperatures: an *in vitro* study. BMC Musculoskelet Disord 2010;11:96.
17. McMullan JT, Jones E, Barnhart B, et al. Degradation of benzodiazepines after 120 days of EMS deployment. Prehosp Emerg Care 2014; 18:368.
18. Peculo Carrasco JA, Cosano Prieto I, Gomez Escorza J, et al. Physicochemical stability and sterility of previously prepared saline infusion solutions for use in out-of-hospital emergencies. Resuscitation 2004; 62:199.
19. Rawas-Qalaji M, Simons ER, Collins D, Simons KJ. Long-term stability of epinephrine dispensed in unsealed syringes for the first-aid treatment of anaphylaxis. Ann Allergy Asthma Immunol 2009;102:500.
20. Risha PG, Vervaet C, Vergote G, et al. Drug formulations intended for the global market should be tested for stability under tropical climatic conditions. Eur J Clin Pharmacol 2003;59:135.
21. Stein C. Medication storage temperatures in primary response vehicles. S Afr Med J 2008;98:535.
22. Tickner JA, Schettler T, Guidotti T, et al. Health risks posed by use of Di-2-ethylhexyl phthalate (DEHP) in PVC medical devices: a critical review. Am J Ind Med 2001;39:100.
23. Treleano A, Wolz G, Brandsch R, Welle F. Investigation into the sorption of nitroglycerin and diazepam into PVC tubes and alternative tube materials during application. Int J Pharm 2009;369:30.
24. Twagirumukiza M, Cosijns A, Pringels E, et al. Influence of tropical climate conditions on the quality of antihypertensive drugs from Rwandan pharmacies. Am J Trop Med Hyg 2009;81:776.
25. United States Pharmacopeial Convention: United States Pharmacopeia, National Formulary (USP 38-NP 33). Rockville, Maryland: United States Pharmacopeia; 2015.
26. Valenzuela TD, Criss EA, Hammargren WM, et al. Thermal stability of prehospital medications. Ann Emerg Med 1989;18:173.

Index

Page numbers followed by "*f*" indicate figures, "*t*" indicate tables, and "*b*" indicate boxes.