AFMC Primer on Population Health
An AFMC Public Health Educators' Network resource

Copyright © The Association of Faculties of Medicine of Canada

Content of the AFMC Primer on Population Health by The Association of Faculties of Medicine of Canada, Public Health Educators' Network is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ or send a letter to Creative Commons, 444 Castro Street, Suite 900, Mountain View, California, 94041, USA.

Permissions beyond the scope of this license may be available at www.afmc-phprimer.ca/termsofuse.

The AFMC Primer on Population Health **content** is freely available under the Creative Commons Attribution-NonCommerical-ShareAlike 3.0 Unported license. Please review the license at http://creativecommons.org/licenses/by-nc-sa/3.0/. A brief interpretation of the license is that it allows others to remix, tweak, and build upon this work non-commercially, as long as you credit us and license any new creations under identical terms. Suggestions or appropriate wording to use in the credit can be found at: http://www.afmc-phprimer.ca/termsofuse

We also ask that if you use the Primer content in your teaching and then improve it, please send the improvements to publichealth@afmc.ca so we can consider incorporating them into the AFMC Primer.

IMPORTANT: Parts of the Primer include **third party content** that is licensed under different terms than those listed above. The license terms of that content are either indicated at the instance or the content is acknowledged to be "Used with permission." All permissions must be sought directly from the owners of the third party content and cannot be assumed from this AFMC product.

Please see the full Copyright and Terms of Use for the AFMC Primer on Population Health: http://www.afmc-phprimer.ca/termsofuse

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

AFMC The Association of Faculties of Medicine of Canada
L'Association des facultés de médecine du Canada

A virtual textbook on Public Health concepts for clinicians.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Chapter Objectives

After completing this chapter, the reader will be able to:

- Define and discuss the concepts of health, wellness, illness, disease, and sickness;
- Discuss alternative definitions of health;
- Discuss the concepts of life course and natural history of disease, particularly with respect to possible public health and clinical interventions;
- Describe how the medical, social and spiritual determinants of health and well-being for First Nations, Inuit, Métis peoples impact their health;
- Demonstrate the role that physicians can play in promoting health and preventing diseases at the individual and community level.

Case Study

Paul Richards consults Dr Rao about his state of health. He hasn't really been well since he lost his job 5 years ago. He is now complaining of tiredness and chest pain, which Dr Rao thinks is probably angina. Dr Rao is reflecting on the case. He will certainly investigate the chest pain, but he believes that Paul's job loss is the main cause of his ill-health and that the best therapy for Paul would be another job. However, Dr Rao can't just prescribe one. He could refer Paul for social or psychological counselling, but is not sure if it would be of any use. He would like to see Paul take up a hobby or voluntary work, which might take his mind off his symptoms and renew his interest in life.

Introduction: the widening scope of medicine

Over the past half-century, the Canadian population has seen unprecedented gains in longevity, health, and well-being. Improvements in the environment and in health policies, changing lifestyles and therapeutic advances have all contributed to enhancing the length and quality of life. We can now expect to live into our 80s, and ours is an era of rising expectations. The public confidently expects that new treatments will be developed to cure previously untreatable conditions. Patients expect faster access to health care to deliver these treatments, and more complete information about them.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Illustrative Materials

Improving Health in Canada

Mortality rates have been declining for many years. When adjustments are made to remove the effect of the rising average age of the population, Canadian mortality from all causes fell from 8.2 deaths per thousand population in 1981 to 6.9 per thousand in 1996, and to 5.4 in 2006.

There has been a corresponding rise in life expectancy, the number of years a newborn Canadian can expect to live. Life expectancy at birth increased from approximately 75 years in 1981 to almost 79 years in 1996, and, in 2006 was over 81 years.

The infant mortality rate has been declining for 50 years. In 1982 there were 9.1 infant deaths for every thousand live births. The figure fell to 5.6 deaths in 1996 and in 2006 was around 5.0 infant deaths per 1,000 live births. This is among the lower rates in the world, but considerably higher than in Japan, which has the lowest IMR at 3.8 deaths per 1,000.

Between 1979 and 1996, Canada saw the following changes in death rates:

- Acute rheumatic fever fell by 40%, although there remain pockets of high incidence
- Acute respiratory infections fell by 40%
- Ischaemic heart disease fell by 26%; between 1956 and 2002, cardiovascular disease showed an overall decline of 70%
- Diseases of the esophagus, stomach, duodenum fell by 24%
- Emphysema fell by 22%
- Hypertensive disease fell by 20%
- Cerebrovascular disease fell by 16%
- Homicides fell from 2.3 per 100,000 population in 1981 to 1.8 in 1996 and have remained roughly at that level since then
- Suicides fluctuate between 13 and 15 per 100,000 and have not shown a clear trend over time
- Despite the ever-increasing number of vehicles on the roads, the annual number of deaths fell 52% from 5,933 in 1979 to 2,875 in 2004.

Canada lies roughly third among industrial countries in terms of life expectancy, mortality rates and people's judgments of their own health.

(Sources: various reports from Statistics Canada)

Perhaps surprisingly, the improvements in health have not reduced the demands on doctors. Instead, doctors are called on to broaden the scope of what they treat. Conditions, previously not regarded as medical problems, such as hyperactivity in children, infertility in young couples, weight gain in middle-aged adults, or the various natural effects of aging, now commonly lead patients to consult their doctor; the list is likely to expand.

Dramatic medical advances are exciting, but they bring challenges and have raised concerns. First, there are concerns over equity: not everyone has benefitted equally from improvements in health and identifiable sections of society consistently have poorer health than the average. This has led to calls for action to reduce health inequalities, discussed in Chapter 2. Second, therapeutic innovations force us to consider the cost implications for a publicly funded, universal health care system. Third, as well as financial concerns, there are philosophical implications of the broadening scope of care. Applying medical treatments to palliate avoidable problems (such as obesity or type II diabetes) that arise in large measure from lifestyles focuses attention on social accountability. In this complex arena of debate a practical question is raised: what conditions should doctors be expected (and paid) to treat; and hence, how should they be trained?

In partial response, the Royal College of Physicians and Surgeons of Canada published the CanMEDS framework in the 1990s to define competencies that physicians would need, and roles they should master, in providing the best quality of care in the new millennium[1]. The physician roles acknowledge that, in addition to being medical experts, graduates are also expected to be competent as communicators, collaborators, managers, health advocates, scholars, and professionals. A far cry from the traditional solo practitioner engaged in long-term caring relationships with his patients, the physician has become an agent of health whose work involves treating patients, advocating for better policy, gate keeping health resources and attending conferences and engaging in research.

Discussions over setting appropriate boundaries for medicine led to the surprising insight that there is no agreed criterion for defining what constitutes a disease. The push of supply (whether from companies or from doctors) and the pull of demand (from patients and society) have led us to classify more and more common conditions as diseases. The development of Viagra transformed impotence (which had presumably existed for millennia) from a matter of personal embarrassment into a widely publicized problem for which treatment is routinely prescribed. Broadening the definition of disease has benefits but may also have disadvantages. Potential hazards were recognized in the 1970s by Ivan Illich, ex-priest, social critic and polemicist, in his discussion of ?the morbid society.'

The Morbid Society

"Each civilization defines its own diseases. What is sickness in one might be chromosomal abnormality, crime, holiness, or sin in another . . . In a morbid society the belief prevails that defined and diagnosed ill-health is infinitely preferable to any other form of negative label or to no label at all. It is better than criminal or political deviance, better than laziness, better than self-chosen absence from work. More and more people subconsciously know that they are sick and tired of their jobs and of their leisure passivities, but they want to hear the lie that physical illness relieves them of social and political responsibilities. They want their doctor to act as lawyer and priest?"[2]

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

?Non-Disease'

In 2002, the British Medical Journal stimulated a debate over the appropriate expectations to place on doctors and on how to define the limits of medicine. Richard Smith, editor of the Journal, surveyed readers to collect examples of non-diseases, and found almost two hundred. He defined non-disease in terms of "a human process or problem that some have defined as a medical condition but where people may have better outcomes if the problem or process was not defined in that way." Examples include burnout, chemical sensitivity, genetic deficiencies, senility, loneliness, bags under the eyes, work problems, baldness, freckles, and jet lag.

Smith's purpose was to emphasize that disease is a fluid concept with no clear boundaries. He noted various dangers in being over-inclusive in defining disease: when people are diagnosed with a disease and become patients they could be denied insurance, lose their job, have their body invaded in the name of therapy, or be otherwise stigmatised.[3]

The debate is covered in the British Medical Journal, April 13, 2002; vol. 324: pages 859-866 and 883-907.

Illness, Sickness, and Disease

Discussing the complexities of what constitutes a disease requires careful distinction among related, but distinct concepts. In 1973, Susser, an epidemiologist, proposed some definitions that remain useful.[4] He used ?**illness**' to refer to the subjective sense of feeling unwell; illness does not define a specific pathology, but refers to a person's subjective experience of it, such as discomfort, tiredness, or general malaise. The way a patient reports symptoms is influenced by his or her cultural background, and Susser applied the term ?**sickness**' to refer to socially and culturally held conceptions of health conditions (e.g., the dread of cancer or the stigma of mental illness), which in turn influence how the patient reacts (See **Culture** in Glossary). The social perceptions of disease that Illich described modify the ways a patient perceives and presents his symptoms.[2] Cultural conventions likewise affect where the boundary between disease and non-disease is placed: menopause may be considered a health issue in North America, but symptoms are far less commonly reported in Japan.[5,6]?Disease' implies a focus on pathological processes that may or may not produce symptoms and that result in a patient's illness. For example, a patient complains of tiredness and malaise?his illness as he experiences it. He consults a doctor about it?because he believes that he might have a sickness. The doctor might attribute the patient's symptoms to a thyroid condition?a disease.

The ?biomedical model' of disease has dominated medical thinking since the time of Louis Pasteur (1822?1895) and the microbiological revolution. This model focuses on pathological processes, and on understanding, diagnosing, and treating the physical and biological aspects of disease. The goal of treatment is to restore the patient's physiological integrity and function. Diagnosis involves recognizing and applying a label to a pattern of signs and symptoms that is at least partly understood in terms of abnormal structure or function of cells, organs, and systems. This offers a rational basis for the investigation of effective treatments. For instance, a certain pattern of chest pain known as angina pectoris is understood biologically as a disorder of the coronary arteries that causes cardiac ischeamia, and the treatments it are geared to the specific causes of restoring cardiac blood flow and reducing cardiac effort.

Early biomedical conceptions supposed that a disease is either present or absent: a bacterium has invaded the body or it has not. However, as medicine increasingly tackled conditions, such as hypertension, which represent deviations from normal values, which themselves have a range and can be debated, it became apparent that there may be no set threshold for defining disease. Thus, instead of being seen as a state that is qualitatively distinct from health, many diseases have to be approached as a quantitative threshold on a continuum of biological variability (see Nerd's Corner box on alternative definitions of disease). Organizations such as the World Health Organization (WHO) and the National Institutes of Health have proposed different classifications of hypertension and have changed how they constitute hypertension over time. Hypertension can be mild, moderate or severe, or defined as pre-hypertension or hypertension stage 1 or stage 2. Ways of establishing what is a ?normal' value is further discussed in Chapter 6.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

Alternative Approaches to Defining Disease

Basis for the definition of ?disease'	Comments
	It can be difficult to define what ?pathological' means, especially in cases of psychological or behavioural disorders. What is the threshold between ?normal' and ?pathological'? This often leads to defining threshold in terms of impact on function: see below.
Abnormal conditions that lead to impaired or altered function	Implies that disease exists only when it impairs function. This further implies that, if the person does not complain, then it's not worth bothering with. This may leave conditions such as slowly deteriorating eyesight in elderly people, untreated.
Illness or sickness that produces unpleasant symptoms	Use of alternative terms (?illness' or ?sickness') does not really help; again implies that disease must produce symptoms.

Questions to ponder:
If no symptoms are produced, is it a disease?
Are health and disease different entities or merely different points along a continuum?
If so, should we abandon the notion of disease and think only of different levels of health, changing from a categorical to a dimensional model?

Suggestion:
Perhaps disease should be defined in terms of a pathological process (physical or mental) that, if left untreated, would naturally progress to producing symptoms and damaging function.

Disease or Syndrome?

As we learn more about the biological basis for a patient's illness, it may be reclassified as a disease. For example, constant feelings of tiredness became accepted as the medical condition of chronic fatigue syndrome. Sometimes when a doctor formally labels (diagnoses) a patient's complaint, the complaint is legitimized and this may reassure the patient.[7] Often, however, a set of signs and symptoms eludes biomedical understanding. If the set is frequent enough to be a recognised pattern, it is termed a syndrome instead of a disease. A syndrome refers to a complex of symptoms that occur together more often than would be expected by chance alone. Whereas diseases often receive explanatory labels (such as hemorrhagic stroke), syndromes are often given purely descriptive labels (e.g., Restless Leg Syndrome). Confusingly, the label ?syndrome' often persists long after the cause is discovered, as with Down syndrome, AIDS (Acquired Immunodeficiency Syndrome) or SARS (Severe Acute Respiratory Syndrome). Meanwhile, Chronic Fatigue Syndrome, Fibromyalgia, Irritable Bowel Syndrome, and Restless Leg Syndrome remain syndromic conditions which, so far, are not well explained by conventional biomedical models.

Disease as a Process: Natural History and Clinical Course

The nineteenth-century revolution in thinking brought about by Koch and Pasteur led to the recognition of distinct stages in the development of a disease. If left untreated, a disease would evolve through a series of stages that characterize its natural history. But if an intervention is applied, the natural history is modified, producing a typical clinical course for the condition. Figure 1.1 represents the concept of health and disease as processes (rather than states) that unfold over time in a series of steps. The dashed line in the centre of the diagram indicates that the disease progression may be interrupted at any stage. It's important to note that not all cases progress across all the stages.

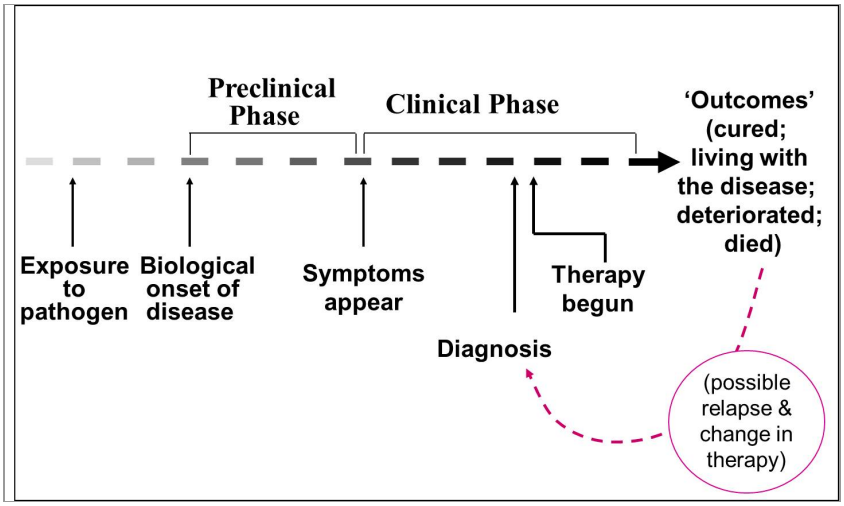


Figure 1.1: The clinical course of a disease, from the biomedical perspective

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

After contact with an infectious agent (or following some other physiopathologic event) there is a theoretical point at which the disease process may begin. Symptoms may appear after a delay that can vary from seconds (as with anaphylaxis) to years (as with some cancers). The patient may interpret their symptoms as indicating an illness and may seek professional care. Shortly after a medical diagnosis, therapy is normally begun and short- and longer-term outcomes can be recorded. If the outcome is unsatisfactory or unexpected there may be a loop back to re-diagnose the condition and alter the therapy, as suggested by the pink curved line. The move towards thinking of disease as a process, rather than a state, required new concepts to describe the stages in this process.

The Sequence of Disease Outcomes

In 1980 the WHO published the International Classification of Impairment, Disability and Handicap (ICIDH), which proposed standard terms for the stages in the clinical course of a disease (see Figure 1.2).[8] In this conception, pathology produces some form of disease and results in ?**impairment**', a deviation from normal function in an organ or system. For example, when atherosclerotic plaque narrows coronary arteries, the patient may experience angina and normal cardiac function is impaired. Impairments are not always perceived by the patient, and screening tests are used to identify impairments of which the person is not aware.

In turn, an impairment can, although does not necessarily, lead to a **disability**. A disability is defined as "any restriction or lack (resulting from an impairment) of ability to perform an activity in the manner or within the range considered normal for a human being." For instance if Paul Richards's narrowed arteries cause him chest pain and if this limits his ability to walk, he has a disability due to heart disease and angina. However, an impairment can often be corrected (medically, surgically or by a prosthesis), so there may be no resulting disability.

In its turn, disability may or may not limit the patient in performing his normal social roles. For instance, severe angina may prevent a patient from working, producing social, psychological, and economic hardships in terms of lost income, self-esteem, and social position. ?**Handicap**' is defined as "a disadvantage for a given individual, resulting from an impairment or a disability, that limits or prevents the fulfillment of a role that is normal (depending on age, sex, and social and cultural factors) for that individual." Handicap relates the impact of a disease to the social roles of the person with it. Practical interventions can prevent a disability from becoming a handicap: finding a desk job for a person with angina of effort, or making buildings wheel-chair accessible for people with mobility problems. But some consequences are harder to compensate: perhaps a 49 year old patient with intermittent claudication no longer feels like "one of the guys" since he can no longer kick the soccer ball with them on Sunday or feels he is letting his son down at the annual father-son two-on-two basketball tournament. Because they are hard to correct, social and psychological sequelae are often overlooked.

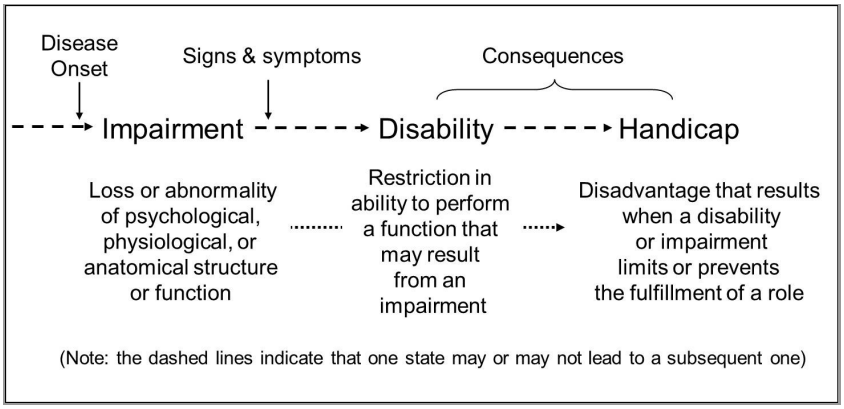


Figure 1.2: The WHO impairment, disability, and handicap triad

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

Mapping the Hierarchy of Disease Outcomes

Another sequence in the outcomes of a health problem is called the five Ds. The idea was originally proposed by Fries, who listed several possible outcomes of the clinical course of an illness: discomfort, drug toxicity, disability, death, and the dollar cost of care.[9] Subsequent authors have modified the list, often referring to discomfort, dysfunction, disability, distress, and death. Whatever terms are used, the idea captures the broad range of impacts of a disease on different patients, and has encouraged physicians to think more broadly than just about a patient's physical function.

It may be possible to link these various descriptions of the impact of a disease, along the axes of scale (running from cellular to social function), and time, as illustrated in Figure 1.3.

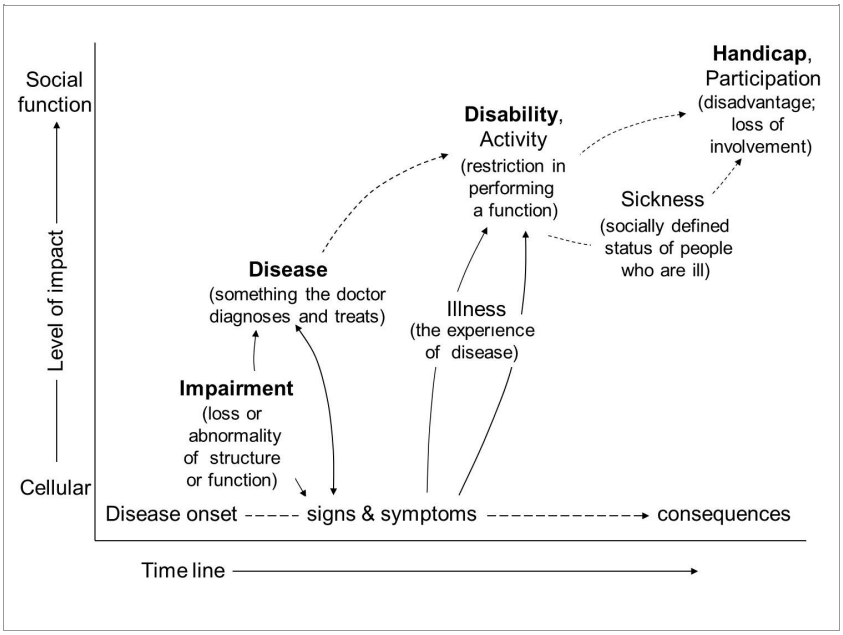


Figure 1.3: Connections among disease, illness, disability, sickness, and handicap

The International Classification of Function

Concepts such as disability and handicap focus on the negative consequences of disease, and may mask the reality that many people cope very successfully with their condition. In 2001, the WHO proposed more positive phrasing in terms of activities and abilities, resulting in the International Classification of Function, or ICF.[10] In this revised classification, ?activity' and ?participation' replace ?disability' and ?handicap', which further blurs the distinction between health and disease. However, the ICF goes beyond merely proposing new terms; it forms a complete classification system for health states, covering body structures and functions, impairments, activities, and the person's participation in society. It also considers contextual factors such as housing, transportation and work that can affect activity levels; these are among the social determinants of health that will be described in Chapter 2. In other words, function is viewed as an interaction between a person's health condition (such as a disease or injury) and the context in which he or she lives, including physical environment and cultural norms relevant to the disease. As an international project, the ICF establishes a common language for describing states of health for comparative analyses across diseases and countries.

Nerd's Corner

Some of the ICF Concepts

Body structures are anatomical parts of the body, such as organs, limbs, and their components.

Body functions are the physiological functions of body systems, including psychological functions.

Impairments are problems in body structure or function, such as a significant deviation or loss.

Activity is the execution of a task or action by an individual.

Activity limitations are difficulties an individual may have in executing tasks.

Participation is involvement in a life situation.

Participation restrictions are problems an individual may have in involvement in life situations.

Environmental factors make up the physical, social, and attitudinal environment in which people live and conduct their lives. Environmental factors can serve as barriers or as facilitators in a person's adjustment to impairments.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Case Study

Dr Rao arranges for investigation and initial treatment of Paul's chest pain. He then sees Mrs Rebikov, an 86 year old lady who was diagnosed with colon cancer 6 years ago after presenting with bowel obstruction. Mrs Rebikov was always a vigorous woman; she recovered quickly from her bowel resection and always said she enjoyed going to Weenigo for her chemotherapy because it gave her a chance to get into a really good bookshop. Mrs Rebikov's cancer has progressed to the point where she now receives palliative care only. She is currently writing about her experience of cancer for the Goosefoot Times. Dr Rao can't help feeling that Mrs Rebikov is much healthier that Paul Richards.

Definitions of Health

If there are complexities in defining disease, there are even more in defining health. Definitions have evolved over time. In keeping with the biomedical perspective, early definitions of health focused on the theme of the body's ability to function; health was seen as a state of normal function that could be disrupted from time to time by disease. An example of such a definition of health is: "a state characterized by anatomic, physiologic, and psychological integrity; ability to perform personally valued family, work, and community roles; ability to deal with physical, biologic, psychological, and social stress".[11] Then, in 1948, in a radical departure from previous definitions, the WHO proposed a definition that aimed higher, linking health to well-being, in terms of "physical, mental, and social well-being, and not merely the absence of disease and infirmity".[12] Although this definition was welcomed by some as being innovative and exciting, it was also criticized as being vague, excessively broad, and unmeasurable. For a long time it was set aside as an impractical ideal and most discussions of health returned to the practicality of the biomedical model.

Health as a Resource

Just as there was a shift from viewing disease as a state to thinking of it as a process, the same shift happened in definitions of health. Again, the WHO played a leading role when it fostered the development of the **health promotion** movement in the 1980s. This brought in a new conception of health, not as a state, but in dynamic terms of resiliency, in other words, as "a resource for living".[13] The 1984 WHO revised definition of health defined it as "the extent to which an individual or group is able to realize aspirations and satisfy needs, and to change or cope with the environment. Health is a resource for everyday life, not the objective of living; it is a positive concept, emphasizing social and personal resources, as well as physical capacities".[14]Thus, health referred to the ability to maintain homeostasis and recover from insults. Mental, intellectual, emotional, and social health referred to a person's ability to handle stress, to acquire skills, to maintain relationships, all of which form resources for resiliency and independent living.

Nerd's Corner

More Detail on Evolving Conceptions of Health

In 1977, psychiatrist George Engel had presented the idea of health capacity under the less than catchy title of the biopsychosocial model of health.[15] He portrayed disease as the body's physiological responses to demanding environmental or social triggers; like Susser, Engel saw illness in terms of the patient's interpretation of symptoms. This interpretation was influenced by the patient's beliefs and personal relationships, both of which are components of health, or resiliency. He thereby presented health and disease as not merely a continuum along which an individual was situated at any moment, but as a form of interaction between the opposing forces of stimulus or challenge and coping response.

Much earlier, Freud had spoken of health as the capacity to love and to work (a view that appeals to some medical students). A person who has the capacity to do what he chooses is healthy; this represented an early shift from an absolute definition to a relative one, and brought health closer to notions of freedom and the subjective quality of life.

A corollary is that a patient may have a serious disease yet perceive himself as reasonably healthy if he can still do what he wishes to do: health can thus be viewed in terms of adjustment to reality. Similarly, the notion of healthy aging may be defined as a person's adjustment to the natural decline in their faculties, as living in balance with their capacities. The process whereby people adjust their expectations downward to match their declining health, and so maintain their satisfaction, is termed response shift.[16]

Wellness

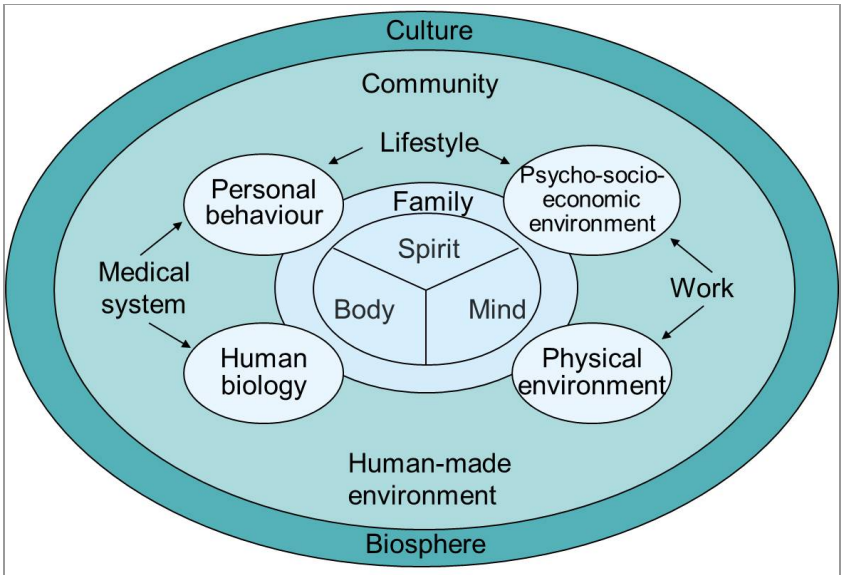
This chapter opened with the theme of rising aspirations and the resulting reconceptualization of disease and health. In response, many practitioners have expanded their focus to include wellness at the positive end of the health continuum. Some distinguish two interacting dimensions: disease versus non-disease and well-being versus ill-being; others expand the number of dimensions to include spiritual, emotional, social, and mental. Last commented that wellness is "a word used by behavioural scientists to describe a state of dynamic physical, mental, social, and spiritual well-being that enables a person to achieve full potential and an enjoyable life".[17]

But with so much disease to treat, should physicians concern themselves with wellness? Is it appropriate for medicine to seek ways to promote positive health states? Some academics distinguish between a medical care system and a health care system, arguing that, to constrain costs, public funding should be limited to treating illness and restoring the patient's functional capacity. Others note that activities such as counselling and educating healthy individuals on diet and exercise promote wellness and resiliency, and so fall within the scope of normal practice as a part of preventive medicine. Some go further and argue that physicians should advocate for improved work and environmental conditions, such as promoting walking and cycling rather than driving, and should advocate for policies that redistribute income, limit access to unhealthy foods, and support children's programs. As concepts of health and disease continue to broaden, there will no doubt be pressure for physicians to expand their role to include the promotion of positive health states in their patients. Reflecting this trend, clinical trials evaluating new pharmaceuticals must now

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

include improved quality of life as an outcome, which obviously extends beyond simply improving biomedical indicators of pathology.

Discussions of wellness eroded the hold of the biomedical model. In its place, ecological models of health appeared; these recognize the complex interactions among people, their personal characteristics and the environment, and how these influence health. An example of this thinking was provided by Trevor Hancock in his ?mandala of health'.[18] This is a model of health and the community ecosystem that represents health determinants as concentric nested influences, beginning with the person at the centre (distinguishing body, mind, and spirit), then moving outwards to the social and physical environment, and then moving further out to culture, economic, and societal influences. The mandala is intended to draw attention to the wide range of health determinants, and to the need to address many levels in developing strategies for improving health.



Redrawn from: Hancock T, Perkins F. The mandala of health: a conceptual model and teaching tool. *Health Education* 1985;Summer, pp. 8-10.

Figure 1.4 The Mandala of Health.

More recently, a global health perspective has added further rings to the concentric circles to represent the health influences of global climate, economic processes, wars, culture, and the impact of travel in quickly spreading disease. The theme of social determinants of health is discussed in Chapter 2. An extension of the mandala that includes a time dimension is the integrated **life course** and social determinants model of Aboriginal health proposed by Reading and Wein.[19, p26] This model represents health influences as a set of concentric spheres (there is a cut-away so that the inner spheres can be seen). The child is at the centre, youth and adulthood in surrounding spheres, and successive layers of determinants outside of these. The inner spheres are segmented to represent the mental, physical, emotional, and spiritual aspects of health.[19, p26] This model of aboriginal health is an example of this more integrated complex idea of health. It suggests the time dimension by showing the child growing through youth into adulthood, within the influence of culture and society and the other health determinants.

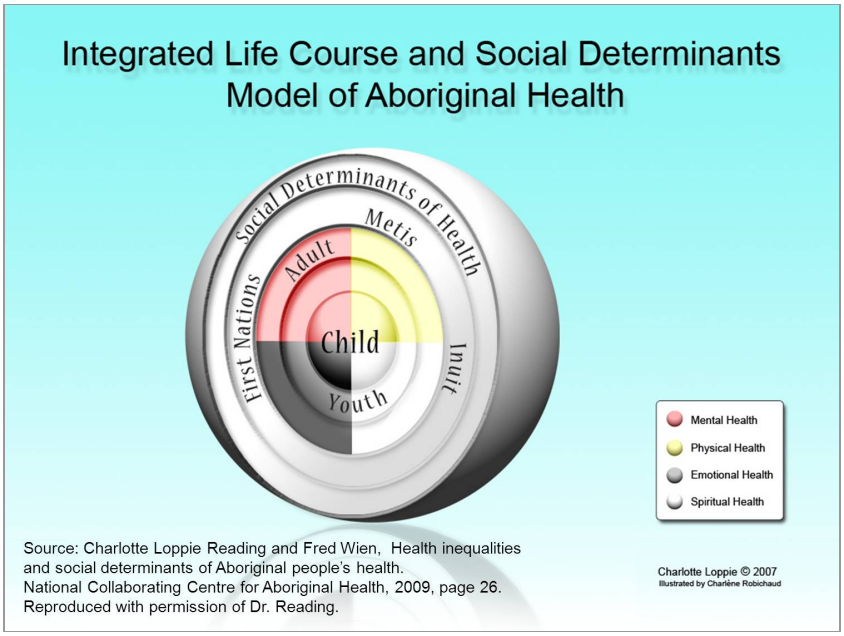


Figure 1.5. Integrated life course and social determinants model of Aboriginal health[19]

Spiritual health and Canadian indigenous healing traditions

As conceptions of health broadened, they challenged the traditional western Cartesian tradition of separating mind and body (see Nerd's Corner box

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

titled "Descartes"). The biomedical model hugely advanced our ability to treat the body but is somewhat less successful in treating the mind. One response has been to adopt medical traditions from other cultures, which commonly emphasize healing rather than treating, and focus attention on the patient's spirit as well as their body. Spiritual health may be indicated by a sense of peace, hope, purpose, commitment, or worth. For some people spirituality is found through religious practice, while others find it in connection with their values, or in nature, art, sport, or music. A physician who recognizes the relevance of spiritual health may be better able to support her patients, bringing a sense of peace, comfort, strength, love, or connection in order to improve the patient's sense of well-being.[20] For example, chemical dependence is increasingly being considered as having biological, psychological, social and spiritual dimensions.

Nerd's Corner

Descartes

Before the French mathematician and philosopher René Descartes (1596?1650), explanations for the observed world were based on a mix of religious beliefs and interpretations of frequently biased observations. Descartes pioneered the open-minded scepticism that characterizes modern science and began by rejecting all forms of preconceptions and beliefs. Science had to consider only what was observable, and the mind was not. Descartes argued that our senses could not be relied on as accurate sources of information, hence proposed a rigid division between mind and body: the Cartesian dualism.

As an example of a challenge to the Cartesian way of thinking, First Nations, Inuit and Métis indigenous teachings have for a long time recognized other, non-physical dimensions of health.[21,22] These various traditions take a broad ecological approach, and consider health in terms of a balance of the spiritual, emotional, physical and intellectual dimensions of a person acting as an individual, but also as a member of a family, community and nation in a variety of cultural, social, economic and political environments. In contrast to Western medicine, which focuses on treating illness, indigenous medical thinking emphasizes healing, which is achieved by restoring balance in the four realms of spiritual, emotional, mental and physical health. Indigenous healing relationships are based on respect, humility, compassion, truth, sharing, hospitality, and divine love. They recognize more routes to healing than biomedical science, and the contrast highlights the dilemma facing modern physicians who aspire to provide health care when they often only have time to treat disease.[23] These ideas are symbolized in various ways by different indigenous groups. For the First Nations peoples in Canada, the ideas are often represented by quadrants on a medicine wheel (see the Illustrative Materials box).[24] For Inuit peoples, the Learning Blanket has been used as a symbol to explain health and wellness, while there are references to the Métis view of holistic health within the Métis tree model of Holistic Lifelong Learning (see Illustrative Materials box).

In recognition of the relevance of Indigenous concepts of health and healing, the Indigenous Physicians Association of Canada (IPAC) and the Association of Faculties of Medicine of Canada (AFMC) created a set of competency standards for medical students working with First Nations, Inuit, and Métis patients. The competencies are structured around the CanMeds physician roles; they are designed to train physicians to recognize the complexities of the historical relationship between indigenous and non-indigenous peoples, and how this complexity continues to affect the health care provided to indigenous people. The competencies are intended to guide the development of curricula in order to support medical students in developing culturally safe practice when engaging with First Nations, Inuit, and Métis patients. The First Nations, Inuit, Métis Health Core Competencies are available on the AFMC web site at: http://www.afmc.ca/social-Aboriginal-health-e.php

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Illustrative Materials

The Medicine Wheel

The medicine wheel symbolizes the interconnection of all life, the various cycles of nature, and how life represents a circular journey. It has four quadrants that are often coloured red, yellow, black, and white or green (an illustration is available at http://www.webpanda.com/There/uot_directions-colors.htm).

The number four is sacred to many Aboriginal peoples of North America as representing many things:

1) The four cardinal points, each with a guiding spirit, symbolize stages in the life journey. The East, the daily birthplace of the sun, represents a person's beginning and early years. The South relates to childhood and intellectual growth. The West symbolizes adulthood and introspection, while the North represents the old age, wisdom, and the spiritual aspects of life. The centre of the wheel is the axle of Mother Earth and the Creator, representing their role in the beginning and continuation of life.

2) The four points can also represent the balance between four aspects of health: spiritual (East), mental (North), physical (West), and emotional (South).

3) The wheel can represent the process of making decisions. Values (represented by the East, where the sun rises) guide decisions taken in the mental realm (placed in the North, at the top). Then, decisions are implemented in the physical realm (West), and actions produce reactions in the emotional realm (South). These reactions feed back into the value system, completing the circle of value, decision, action, and evaluation. As an example, the loss of traditional values resulting from experiences in residential school could affect health decisions (e.g., consumption of alcohol or loss of traditional diet) that then create an imbalance in the emotional state of a person.

4) The wheel can represent the Four Sacred Medicines:

Sweetgrass (the North) is used by almost all Indigenous peoples in North America for ritual cleansing. When sweetgrass is walked on, it bends but does not break, hence its link to virtue: an injustice can be returned by a kindness, by bending, not breaking.

Tobacco (the East) is revered as a scared plant: it connects people to the spirit world; it absorbs prayers and carries them to the spirit world and thanks the Creator for his gifts. Traditionally, tobacco was not smoked, except on special ceremonial occasions.

Cedar (the South) is used for purification and (taken as a tea) to attract positive energy and emotions, as well as for balance. Its vitamin C content helped prevent scurvy when fruits and vegetables were unavailable during the winter months.

Sage (the West) is a medicine for women, conferring strength, wisdom, and clarity of purpose. It is a powerful purifying medicine that drives away negative energies. Sage can be found braided and hung in people's homes, perhaps tied with a ribbon in one of the colours of the medicine wheel. The threefold braid represents body, mind, and spirit.

Illustrative Materials

The Learning Blanket

A colourful, circular blanket portrays scenes from Inuit life that symbolize the sources of learning and domains of knowledge?culture, people, and *sila* (life force or essential energy, which derives from the Inuit people's longstanding relationship with their environment). The experiential learning that occurs on the land affirms the Inuit sense of identity. There are also other sources of knowledge, including languages, traditions, family, community, Elders, land and the environment. All of these are illustrated on the blanket by images drawn from Inuit life. The blanket's circular shape represents the interconnectedness of all life forms, and the continuous cycle of life, death and regeneration that connects past, present and future. An image of the blanket model is shown at: http://cli.ccl-cca.ca/Inuit/index.php?q=model (accessed September 2010).

(Source: the Canadian Council on Learning web site on Inuit Holistic Lifelong Learning: http://cli.ccl-cca.ca/Inuit/index.php?q=home)

Illustrative Materials

The Métis Holistic Lifelong Learning Model

There may not be a broadly accepted Métis conception of health and wellness; various models exist among the different groups. One, for example, describes Métis people as viewing life in terms of a learning process, part of a regenerative, living system, a natural order that governs the passage of seasons and encompasses a community of learners. Within this organic system, relationships interconnect so that balance and harmony are maintained. This may be symbolized by a forest of trees. The tree roots represent the individual's health and well-being (social, physical, economic, spiritual), providing the conditions that nurture lifelong learning. The trunk contains the growth rings with spiritual health at its core, spanning outwards to the emotional, physical, and mental dimensions of the Métis identity. These evolve over the life course, as more rings are added each year. Knowledge is added like leaves; different branches represent different sources: self, other people, the land, and traditions. A diagram of the tree is shown at: http://www.ccl-cca.ca/pdfs/RedefiningSuccess/CCL_Learning_Model_MET.pdf

(Source: the Canadian Council on Learning web site: http://www.ccl-cca.ca/CCL/Reports/RedefiningSuccessInAboriginalLearning/RedefiningSuccessModelsMetis.html)

Integrative medicine

Contemporary Western medicine is increasingly being challenged to consider how to respond to perspectives and treatments other than those of conventional allopathic medicine (see Nerd's Corner box "Allopathic Medicine"). One response has been to propose ?integrative medicine'[25] as a collaboration between biomedical approaches and other healing traditions, including herbal remedies, manual interventions such as massage therapy or chiropractic, and mind-body practices such as hypnosis. Similarly, the Canadian College of Naturopathic Medicine trains naturopathic

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

doctors who employ natural therapies as well as using the more standard medical diagnostics of allopathic medicine.

Definition

"Integrative medicine is about changing the focus in medicine to one of healing rather than disease. This involves an understanding of the influences of mind, spirit, and community as well as of the body. It entails developing insight into the patient's culture, beliefs, and lifestyle that will help the provider understand how best to trigger the necessary changes in behaviour that will result in improved health".[25]

Nerd's Corner

Allopathic Medicine

In pre-scientific societies, medicine was founded on a mixture of magic, religion, and empirically tested folk remedies. As magic and superstition declined during the Middle Ages in Europe, medicine sought a new basis; it had to be something active that would impress patients, a substitute for the incantations and incense that sceptics no longer respected. The answer that was adopted, and which was much later named allopathy, had been developed long before in the Middle East around the first century B.C. The idea was simple: when the body's workings deviate from the normal, the doctor should try to restore normal function. If a man was feverish, he should be cooled down; if constipated, given a laxative. Illnesses were seen as caused by toxins that should be eliminated. This led to therapies such as bleeding, leeches, enemas, and purgatives. Interventions were often harsh, as depicted cynically in Molière's plays, and a war metaphor took hold: patients ?battle disease;' doctors ?wage war' on pathogens and both ?join forces to fight cancer'. Note that in French, alternative (i.e. non-allopathic) therapies are called ? médecine douce', or gentle medicine. The term allopathy was coined by Samuel Hahnemann, who founded homeopathy. Whereas allopathy implies opposing the symptoms of disease, homoepathy implies working with the disease by stimulating the body to produce its natural defensive (e.g., immune) responses.

For a time during the mid-nineteenth century, homeopathy (treating like with like) was a serious rival to the allopathic approach, but the development of the germ theory gave allopathy a scientific foundation for many of its remedies. However, by the mid twentieth century disillusionment began when, despite advances in ?the conquest of infectious disease' hospitals remained full and waiting lists stayed long. This may have reflected a rising demand for care induced by the perception of its success, but the very success of allopathic medicine (along with improved social conditions) enabled people to live long enough to suffer degenerative diseases for which the allopathic approach is less effective. Moreover, the allopathic approach has some undesired consequences including the rapid increases in costs and the large numbers of people with iatrogenic disorders.[2] While allopathic remedies are often highly effective, practitioners are also aware that the best cure may be for the patient to simply restore balance in their life and get adequate sleep, exercise, and good nutrition.

Public and Population Health

While **public health** is a familiar term, it can be difficult to give it a single precise definition. Its general focus is on preventing disease and protecting health (See **Health Protection** in Glossary): "Public health is defined as the organized efforts of society to keep people healthy and prevent injury, illness, and premature death. It is a combination of programs, services, and policies that protect and promote the health of all Canadians".[26] But this definition does not give us a clear picture of what is, and what is not, included. In part the difficulty arises because public health does not concern a specific organ system, type of disease or therapeutic approach, but employs a variety of approaches to address whatever health issues are most pressing in each place and time. The discipline has seen a succession of names as it wrestled with whether environmental factors, or individual behaviour, or societal policies should form the main focus of interventions (see Nerd's corner "Evolution of thinking").

Nerd's Corner

Evolution of thinking about public health

As new patterns of disease unfolded over the twentieth century, the nature of public health efforts changed to keep pace, and this was reflected in changing names for the discipline. The earliest approach, termed ?public hygiene', focused on environmental sanitation. Then, between 1920 and 1940, in a society concerned with diseases transmitted by immigration and increasing travel, the term and focus shifted to ?health protection'. As diseases of lifestyle became more common in the 1950s and 1960s, the emphasis shifted towards behaviour so ?health education' was born. Subsequent recognition that education alone was insufficient broadened the field to ?health promotion'. The social context was formally recognized when university departments of ?community health' or ?community medicine' began to appear. The health promotion movement also changed the perspective on the role of individual people, who moved from being recipients of health education to active participants who should take active responsibility for their health, mainly through improving lifestyles. National programs promoted exercise, good nutrition, and smoking cessation. These often implied environmental change (e.g., building sports facilities, walking paths, pedestrian malls), and in the 1980s the healthy cities movement focused attention on the built environment. Debates arose over whether public health should be broadened to include such activities, or whether it should retain a narrower focus.

Meanwhile, critiques of health education argued that many people, especially those who were poorer, had a limited capacity to improve their lifestyles because they were deeply constrained by their socio-economic and perhaps cultural milieux. A common phrase was that health education risked ?blaming the victims' of their circumstances (see Nerd's Corner box Blaming the Victim) and that, by focusing on lifestyle changes, underlying health determinants would not be addressed. Behaviours were the symptoms of underlying social determinants, rather than the main causes of poor health. This gave rise to the population health perspective, which highlights underlying social determinants, as described in Chapter 2.

This debate gave rise to the **population health** perspective, which holds that, while fostering individual responsibility for health, we must also address underlying social determinants, such as poverty or lack of access to care, that constrain people's ability to achieve real gains in health.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

Blaming the Victim

?Blaming the victim' was a phrase coined by William Ryan, an American psychologist, in 1976. Ryan criticized an earlier report on black families in the U.S. that inappropriately attributed their enduring poverty to their cultural patterns and behaviour rather than to the structural conditions of society that constrained their choices.

The concept is relevant in thinking about the challenge of how clinicians can deal empathetically with patients whose illness can be attributed to their lifestyle (e.g., smoking, lack of exercise, alcohol abuse). In factual terms, it can be almost impossible to determine how far a patient's lifestyle is a matter of choice versus social pressures and constraints. Nonetheless, when a patient does not respond to advice to change, the clinician can be tempted inappropriately to ?blame the victim'. For example, in a superficial analysis, obesity arises in part from poor dietary choices, but these choices may be constrained by food availability, affordability, and by the person's social environment in ways likely to be unknown to the clinician. The challenge is to help the patient to find the resources to overcome the problem while maintaining a just perspective on the patient's capacity to control his behaviour and incapacity to control his environment.

In part, a tendency to blame the victim comes from a belief in a just world: if we believe that the world rewards good behaviour, it is difficult to believe that bad things happen to good people. From this perspective, it becomes logical to attribute at least some blame to the person who is suffering.

As population health is a relatively new concept, uncertainties remain over details of how, precisely, it differs from public health. Both are concerned with patterns of health and illness in groups of people rather than in individuals; both monitor health trends, examine their determinants, propose interventions at the population level to protect and promote health, and discuss options for delivering these interventions. The distinction is subtle, but population health is seen as broader, as offering a unifying paradigm that links disciplines from the biological to the sociological. It provides a rational basis for allocating health resources that balances health protection and promotion against illness prevention and treatment, while also making a significant contribution to basic science.[27] When public health tackles a health issue, its interventions are focused on maintaining health or preventing disease. For example, the public health approach to childhood obesity might advocate education for parents and children, subsidized healthy school lunch programs, banning soft drinks in school vending machines, tougher regulations on marketing of junk food to children, promoting physical activity, etc. A population health approach would tackle childhood obesity in a broader context. A population health approach might be to consider the food system itself: how do agricultural subsidies affect the price of food? Can city planning policies prevent the problem of urban food deserts where significant areas of the population lack access to a grocery store? Public health focuses on prevention and health protection services, whereas the population health approach is somewhat broader. It still values "health" as a key outcome, but views issues from a broader perspective and tends to include additional considerations, such as economics, environmental sustainability, social justice, etc.

Definition

Definitions of Population Health

Some define population health descriptively, along simple lines such as "the health of the population, measured by health status indicators.".[17] However, most definitions present population health as a way of thinking about health patterns in society:

John Frank (founding director of the CIHR Institute for Population Health), 1995:

"Population health is a conceptual framework for thinking about why some people, and some peoples, are healthier than others - the determinants of health at individual and population levels. The major determinants of human health status, particularly in countries at an advanced stage of socio-economic development, are not medical care inputs and utilization, but cultural, social and economic factors - at both the population and individual levels."[28]

Kue Young (author of a leading textbook on population health), 1998:

"A conceptual framework for thinking about why some people are healthier than others, as well as the policy development, research agenda, and resource allocation that flow from it (?.)

Population health studies serve the objectives of describing the health status of a population, explaining the causes of diseases, predicting health risks in individuals and communities, and offering solutions to prevent and control health problems. To achieve these aims, population health requires collaboration between the core science of epidemiology, several social sciences which are also concerned with population phenomena, the humanities, and laboratory-based biomedical sciences."[27, p4]

These ways of thinking give rise to ways of acting to improve health, as described by Health Canada, in 1994:

"Population health strategies address the entire range of individual and collective factors that determine health. Traditional health care focuses on risks and clinical factors related to particular diseases. Population health strategies are designed to affect whole groups of populations of people. Clinical health care deals with individuals one at a time, usually individuals who already have a health problem or are at significant risk of developing one."[29]

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

An Organic View of Population Health

Although, in general we think of a population in geographic terms, for instance the population of a country, city, or neighbourhood, a population can also be defined by any factor a group of people might have in common, such as age, socio-economic status, language, or lifestyle. Further, a population can be seen simply as an aggregate of people, or it can be seen as something more than the sum of its members, as a collective organism, as people acting as a group. We can describe Canada as a country that worships hockey, enjoys increasingly drinkable beer and believes in publicly funded health care.

The view of a population as an aggregate of individuals focuses on health *in* the population. By contrast, when the population is seen in emergent terms, as an interacting whole, population health can analyse the health *of* the population.[30] In this view, a healthy community or population is one that works as a group to promote its welfare and address challenges. For example, a healthy population would rally to a natural disaster to establish an effective collective response. Hence, in this view, a healthy population supports and promotes the health of people within it, thereby contributing to individual health; examples include social equity legislation and the development of healthy public policies that characterize ?the caring society.'

Discussions of this type have expanded the health promotion approach into a hybrid sometimes termed population health promotion. Many public health units now take a dynamic view and develop programs that foster the ability of community groups to work together for the improvement of their own health.

Although, in general, physicians treat individual patients and not populations, for several reasons it is important that they are aware of patterns of illness in the population being served. First, although a medical condition may result from the patient's lifestyle, the driving forces for this lifestyle lie in the social environment. Second, the old chestnut of medical teaching "common things are common" is true: in effect, the patient's condition is a symptom of population-wide health patterns. The underlying population prevalence will therefore affect the hierarchy of a doctor's differential diagnoses. In Canada, chest pain in a 50-year-old male is more likely to be of cardiac origin than an identical complaint in a 15-year-old female. Third, because of the social context in which a patient lives, efforts to help him alter health behaviours may be frustrated by the social pressures he experiences. Hence, it may prove more efficient to tackle a disease at the population level (e.g., by lobbying for taxation on high-fat foods, then using the proceeds to subsidize exercise programs) than by treating large numbers of individuals (see Chapter 8).

Reflecting a population health approach, the College of Physicians and Surgeons of Canada described the role of physicians as health advocates: "As Health Advocates, physicians responsibly use their expertise and influence to advance the health and well-being of individual patients, communities, and populations".

Case Study

When Paul Richards came back for his follow-up appointment, Dr Rao, on a professional level, was pleased that he had correctly diagnosed Paul's angina although he had to admit that it Paul had presented with a pattern of symptoms that were typical of a common disease, so the diagnosis was fairly obvious. Moreover, Dr Rao has been seeing more stress associated illness since Goosefoot's economic troubles hit. In effect, Paul's angina is a symptom of population-wide health patterns.

Dr Rao realizes that his efforts to help Paul and other patients change their lifestyles may be frustrated by social pressures in the community. He is well aware of the health advocate role and tries to be a role model for his patients, exercising regularly and keeping an eye on his diet. He has also lent his voice to certain lobby groups in their efforts to get the mayor's office to establish community heart-health awareness weeks and to promote access to methods of transport that promote physical activity.

Self-test questions

Reflection Questions

- 1. How would you define health?
- 2. How would you know if someone is healthy or not?
- 3. If a pathogen lies dormant in a person's body (think of a herpes virus, for example) does the person have a disease? Or, does it only become a disease when symptoms appear (shingles, for example)?
- 4. How do you distinguish between disease and injury?
- 5. How do you establish a balance between preserving freedom to choose how we live, versus taking responsibility for health? We may uphold a person's right to act in a way that can damage his health, but all taxpayers pay for the medical bills that this behaviour entails, and this forms an opportunity cost to everyone: those taxes could have been spent on something that might have benefited more people.
- 6. Addiction is coming to be considered a disease, so should we excuse the anti-social behaviour of a drug addict because they have a medical problem?
- 7. How would you achieve a balanced approach to helping a patient whose lifestyle has caused their medical condition, between holding them responsible for their fate versus rejecting this as an example of ?blaming the victim'?
- 8. Give examples of the way in which social environments impact someone's health. Then, do the same for urban environments.
- 9. Back in the 1950s, Talcott Parsons described ?the sick role' as including various behaviours that were acceptable for someone who is sick, but not for someone who is well. They included things such as lying in bed instead of going to work, having someone cook for you, etc. What behaviours would you consider appropriate for a person who is ill in the modern era? How may this vary across socio-economic strata and across cultural groups?
- 10. If a drug for rheumatoid arthritis causes a change in sedimentation rate or rheumatoid factor but has no detectable effects on the quality of life of those who take it, is it an effective drug?

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Self-Testing

1. How has the World Health Organization's definition of health evolved over time?

The main change was in moving from a ?state of well-being' which seemed a bit static, to seeing health as a capacity to respond to challenges: a ?resource for everyday life'

2. What are key dimensions of health?

There is little dispute over physical, mental and social dimensions. The social is harder to describe than the others; it often refers to the capacity to establish and maintain social bonds and to interact successfully with others. Some authors include dimensions referring to spirituality, in the sense of being at ease with one's destiny or at peace with oneself.

3. Name one difference between a traditional biomedical definition of health and a traditional Aboriginal conception chosen from an Indigenous group with which you are familiar.

Breadth: the Aboriginal conception includes the spiritual dimension; focus on pathology and disease versus on the positive aspects of well-being; focus on treating illness versus on healing the person; eliminating pathogens versus establishing balance, etc.

4. Distinguish between illness and disease.

Perspective: what the patient experiences versus the physician's interpretation of symptoms and signs.

5. List levels at which social environments impact health according to an ecological model.

Often illustrated via concentric circles radiating outward from the person: their family and close friends who influence their feelings and behaviours; their work or school context that may create stresses; their community, province and region that influence the quality of care they receive; the nation that decides on major issues such as going to war or tackling climate change, and global issues that affect health via patterns of climate or travel that can affect health via transmission of infectious diseases, for example

6. What is illness behaviour?

The pattern of reactions a patient has to their perception of being sick, including whether or not to seek care, and how far to comply with the doctor's recommendations if care is sought. These reactions vary from culture to culture and also according to personality: no two patients respond the same way!

7. Illustrate ways in which a person's culture may affect their approach to accessing health care.

First, via differing conceptions of what is considered a ?disease'; second, in terms of how one should respond to symptoms (take medication or not; follow a doctor's advice or grandmother's?); third, in terms of the emotional reaction to a diagnosis; fourth, in terms of acceptance of different forms of treatment, including whole systems of medicine: acupuncture or Ayurveda versus western medicine

References

1. Royal College of Physicians and Surgeons of Canada. The CanMEDS physician competency framework. 2009.

2. Illich I. Limits to medicine. London: Marion Boyars; 1976.

3. Smith R. Non-disease. Br Med J 2002;324:883-885.

4. Susser M. Causal thinking in the health sciences. New York: Oxford University Press; 1973.

5. Lock M, Kaufer P, Gilbert P. Cultural construction of the menopausal syndrome: the Japanese case. Maturitas 1988;10:317-332.

6. Melby M, Lock M, Kaufert P. Culture and symptom reporting at menopause. Hum Reprod Update 2005;11(5):495-512.

7. Wessely S. Pros and cons of medicalisation. Br Med J 2002;324:912.

8. World Health Organization. International classification of impairments, disabilities and handicaps; 1980. 1980.

9. Fries J. The hierarchy of outcome assessment. J Rheum 1993;20:546-547.

10. World Health Organization. Internal classification of functioning, disability, and health. 2001.

11. Stokes J, Noren J, Shindell S. Definition of terms and concepts applicable to clinical preventive medicine. J Community Health 1982;8:33-41.

12. World Health Organization. The first ten years of the World Health Organization. 1948.

13. World Health Organization. Ottawa charter for health promotion. 1986.

14. World Health Organization. Health promotion: a discussion document; 1984. 1984.

15. Engel G. The need for a new medical model: a challenge for biomedicine. Science 1977;196:129-136.

16. Schwartz C, et al. The clinical significance of adaptation to changing health: a meta-analysis of response shift. Qual Life Res 2006;15:1533-1550.

17. Last JM. A dictionary of public health. New York: Oxford University Press; 2007.

18. Hancock T. Health, human development and the community ecosystem: three ecological models. Health Promotion Int 1993;8(1):41-47.

19. Reading C, Wein F. Health inequalities and social determinants of Aboriginal people's health. 2009; Available at: http://www.ahrnets.ca/NCCAH-report-LoppieWein-download(1)%5B1%5D.pdf.

20. FamilyDoctor.org[homepage on the Internet]. Spirituality and health. 2000; Available at:

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

http://familydoctor.org/online/famdocen/home/articles/650.html.

21. Turton C. Ways of knowing about health: an Aboriginal perspective. Adv Nurs Sci 1997;19(3):23-36.

22. Maher P. A review of 'traditional' Aboriginal health beliefs. Aust J Rural Health 1999;7(4):229-236.

23. Hunter L, Logan J, Goulet J, Barton S. Aboriginal healing: regaining balance and culture. J Transcult Nurs 2006;17(1):13-22.

24. Warne D. Traditional perspectives on child and family health. Paediatr Child Health 2005;10:542.

25. Rakel D, Weil A. Philosophy of integrative medicine. In: Rakel D, editor. Integrative medicine Philadelphia: WB Saunders; 2003. p. 6.

26. Public Health Agency of Canada Report on the state of public health in Canada, 2008. Available at: http://www.phac-aspc.gc.ca/publicat/2008/cpho-aspc/index-eng.php. Accessed November, 2009.

27. Young T. Population health: concepts and methods. 1998.

28. Frank J. Why 'population health'? Can J Public Health 1995;86:162-164.

29. Health and Welfare Canada. Strategies for population health: investing in the health of Canadians. 1994.

30. McDowell I, Spasoff R, Kristjansson B. On the classification of population health measurements. Am J Public Health 2004;94:388-393.

AFMC Primer on Population Health

Chapter 2 Determinants of Health and Health Inequities

Chapter Objectives

After completing this chapter, the reader will be able to:

- Define the basic concepts used in discussing causes of health and illness: causes, risk factors, and determinants.
- List the major determinants of health and describe possible mechanisms by which each determinant influences health status:

a. Income and social status

b. Social support networks

c. Education and literacy

d. Early child development

e. Social and cultural environments

f. Physical environments

g. Employment/working conditions

h. Health services

i. Gender

j. Health behaviours

- Describe one or more models of behaviour change, including predisposing, enabling, and reinforcing factors.

Case Study

Ms. Sulawesi is consulting Dr Rao yet again for Peter, her three-year-old son. They live in a damp and poorly heated apartment, and Peter has had 3 emergency room visits in the past month for poorly controlled asthma. Ms Sulawesi is a single mother recently immigrated to Canada with no family and few friends here. She is working as a cleaning lady in the local school, working mainly in the evenings and on week-ends. Dr. Rao tried involving social services, but they are overwhelmed with more pressing issues of domestic violence and abuse. The landlord refuses to make repairs to the apartment and winter is rapidly approaching.

Health inequalities

In every society there are variations in health between people, and extensive medical and social science research addresses the question of why some people are healthy while others are not.[1] Epidemiology focuses principally on variations between people that are systematic, meaning that identifiable groups of people (whether defined in terms of age, sex, social status, ethnic group or in some other way) experience different levels of health. Other variations occur at the individual level: one member of a family gets a cancer when others do not. Explanations for individual variations in health are generally formulated in terms of risk factors, which are characteristics of the person, of their behaviour or environment that affect their chances of contracting a given disease. The probabilistic language of ?risk factors' is used because very few causes inevitably produce health consequences: additional factors always modify their effect. Explanations for broad patterns of disease in whole groups of people refer to ?health

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

determinants'. Determinants often exert their effect via individual risk factors, and so may be viewed as the causes of the causes of disease.

Most diseases and health states occur in predictable patterns. To give an overall illustration, Figure 2.1 shows life expectancy in Canada plotted by sex and income.[2] The data were taken from a study of 2,735,000 Canadians who were followed for 10 years after the 1991 census. Deaths during that time were linked back to the income information recorded on the 1991 census. Income has been adjusted according to the size of the family that depended on it, in order to give a better indicator of income adequacy. The result is shown grouped into five categories, or quintiles, running from poor (coded as 1) to rich (coded as 5). The vertical bars show the years of life remaining, on average, to a 25-year-old person in each sex and income category.

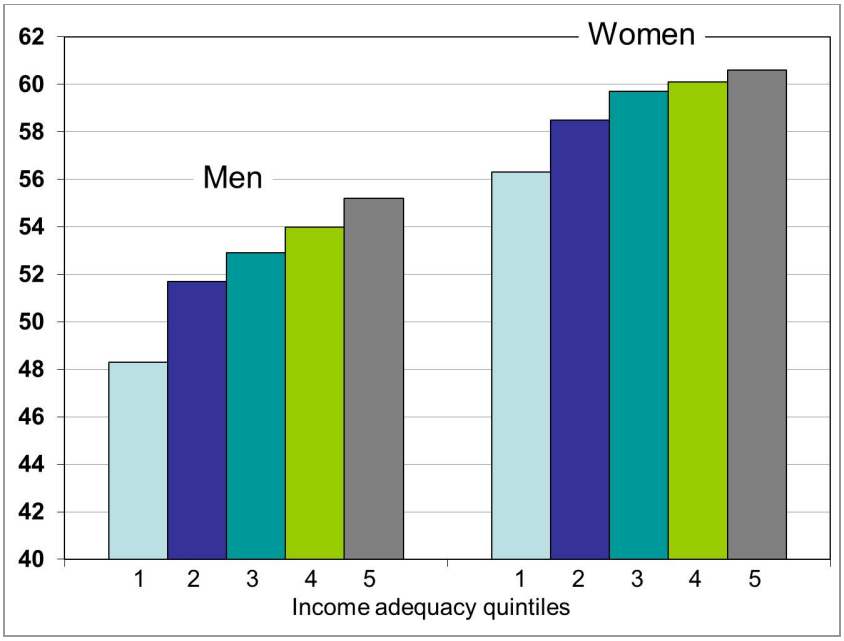


Figure 2.1: Remaining life expectancy at age 25 in Canada by sex and income quintile, non-institutionalized population, 1991 to 2001

The results show that men in the poorest category can expect to live about 7 years less than men in the richest category (48 versus 55 years beyond age 25); for women the contrast is just over 4 years. However, it is not just those with the lowest income who suffer: a key point is that there is a steady rise or gradient in longevity across income levels and this is called the ?social gradient in health'. Similar health gradients occur across levels of education, occupation and residential area, all of which reflect socio-economic status. There are also consistent contrasts in health between certain racial or ethnic groups, as well as a clear contrast between the sexes?the figure above shows that even poor women live longer on average than rich men.[2] Note that this chart only shows life expectancy; there are also gradients in other aspects of health such that rates of disability and ill-health also rise as socioeconomic status declines. Poorer people experience the double deficit of a shorter life and a less healthy one.

An obvious question concerns whether this gradient could simply result from chance factors or whether it represents systematic variations attributable to some cause. There is natural biological variation among people's health and susceptibility to influences on health. Traditional cultures often attributed these individual differences to fate, season of birth, *karma,* or even to witchcraft; we now refer to random variation. As science has increasingly provided explanations for disease, the fraction of variation in health that is attributed to random or chance variations has steadily declined (see Nerd's Corner box "The role of chance").

Nerd's Corner

The role of chance in disease

Is chance an inherent factor in explaining disease? Does God really play dice? Or is chance merely a term we use to account for the cases of disease that we cannot yet explain given our current scientific understanding? We suspect that unknown factors will be discovered and will explain more, but how far can this go? With a fuller understanding of genetics, environmental influences and so forth, will we be able ultimately to predict the time and manner of a person's death?

The position you take on this issue will influence how you react to patients who ask you what they can do to avoid falling sick. If chance plays an important role, what will you tell them?

Similar thinking about the origins of variability can be applied to patterns of disease in whole populations. One reason for presenting results such as those in Figure 2.1 is to search for systematic variations between groups of people. Variations that are not random should in principle be correctable: if richer people can live longer, there is presumably no inherent reason why poorer people could not live equally as long. A health disparity is a contrast between groups of people that confers a consistent disadvantage on one group (such as the difference in life expectancy between women and men). Where this disadvantage is correctable or could have been avoided, it is termed an *inequity,* carrying the idea of unfairness and unacceptability. Inequities form a moral incentive for action. The most prominent sources of **health inequity** in Canada are socio-economic status and Aboriginal identity.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Definition

Health disparities or **health inequalities** refer to systematic differences in health status that occur among population groups. As these are contrasts in health between socially defined groups, the assumption is that they are potentially correctable, and so the reduction of social inequalities in health has become a central goal of population health policy.[3]

The terms ?disparities' and ?inequalities' are broadly equivalent although, as Marmot has noted, disparities is more widely used in the United States while inequalities is used in Britain.[4] Whichever term is used, it is unrealistic to think that we can remove all forms of inequality, so priorities need to be set, and a useful concept here is health inequities.

Health inequity refers to those inequalities in health that are deemed unfair or that stem from some form of injustice. "Reducing health inequities is an ethical imperative. Social injustice is killing people on a grand scale . . . The conditions in which people live and die are shaped by political, social, and economic forces."[3]

Nerd's Corner

Health Inequities among Indigenous Peoples in Canada

Most health indicators for the Aboriginal peoples of Canada have shown clear improvements over the past 20 years, and yet they lag behind equivalent statistics for the population as a whole.

Health Canada produced a 2003 Statistical Profile on the Health of First Nations in Canada (http://www.hc-sc.gc.ca/fniah-spnia/pubs/index-eng.php). Here are some highlights:

- Life expectancy at birth has improved among the First Nations population. In 2000, it rose to 68.9 years for males and 76.6 years for females, an increase of 13.1% and 12.6%, respectively, since 1980. But this is still about 9 years shorter than the Canadian average for men and 5 years shorter for women.
    - Life expectancy for Inuit in Nunavut for 1999 was 67.7 years for males and 70.2 years for females
    - In 1979, the First Nations infant mortality rate was 27.6 deaths per 1,000 live births. By 1999, this had declined to 8.0 per 1,000, 50% higher than the Canadian infant mortality rate of 5.5.
    - In 2003, the most common cause of death for ages 1 to 44 was injury and poisoning. Among children under 10, deaths were primarily due to injuries, but suicide and self-injury were the leading causes of death for youth and adults up to age 44. For people aged 45 and older, circulatory diseases were the leading cause of death. These trends parallel the Canadian population as a whole.
    - With respect to suicide, all First Nations age groups up to age 65 are at higher risk than the Canadian population. The greatest contrast with the overall Canadian rates is for females aged 15 to 24, and males aged 25 to 39 (approximately eight and five times the Canadian rates, respectively).
    - First Nations people experienced a disproportionate burden of infectious diseases. These include pertussis (3 times higher than the national average), chlamydia (7 times higher), hepatitis A (5 times higher) and shigellosis (almost 20 times higher). [Shigellosis is a common bacterial infection in developing countries and results from poor water quality and inadequate sewage disposal. These in turn reflect poverty and poor quality infrastructure].
    - The proportion of Canada's total AIDS cases contracted by Aboriginal people climbed from 1.0% in 1990 to 7.2% in 2001.
    - Over that same period, the tuberculosis rate among First Nations people remained 8 to 10 times that seen in the Canadian population as a whole.
    - Dental decay rates for Aboriginal children in Ontario are two to five times higher than rates among non-Aboriginal children.

Links

Establishing Equity

Equity assumes three elements:

1. Equality: for example, the movement for women's equality seeks to gain equal access to educational or employment opportunities;

2. Fairness: that a person of African descent should get a job, not because he is black, but because he has the right qualifications;

3. Avoidability: there is recognition that something can and should be done to redress situations in which inequities arise.

Social inequities are often redressed via affirmative action programmes that preferentially provide services for disadvantaged groups. Although this might appear to contradict the principle of equal opportunity for all, the purpose is to achieve equity. The Canadian Charter of Rights and Freedoms acknowledges this paradox and Section 15(2) of the Charter notes that equality should not preclude "any law, program or activity that has as its object the amelioration of conditions of disadvantaged individuals or groups..." See the Department of Justice website at: http://laws.justice.gc.ca/en/charter/1.html#anchorbo-ga:l_I-gb:s_15

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Illustrative Materials

The big picture: international disparities in health

Applying the death rates for 2005, Canadians could expect to live 80.4 years on average, tying with Spain. Citizens of Sweden, Australia, Iceland, Switzerland, and Japan slightly outlive us. The British had a life expectancy of 79.1 years, while Americans had a life expectancy of 77.8 years on average, placing them in 23rd place in the international ranking.

There are huge international inequities in life expectancy. Women in Japan live over 85 years on average, more than twice the expectancy for a woman living in Afghanistan (42 years), as illustrated in the figure below. Longevity is related to the wealth of a country, although there are exceptions: the U.S. is very wealthy but has a shorter average life expectancy than several other countries.

In general, life expectancy rises rapidly as very poor countries gain in wealth, but only up to a Gross National Income of around $10,000 US per year?see the inflection in the curve between Jordan and Malaysia in the figure. Thereafter, gains in life expectancy flatten and appear to be more related to patterns of income distribution and expenditures *within* the country than to overall wealth.

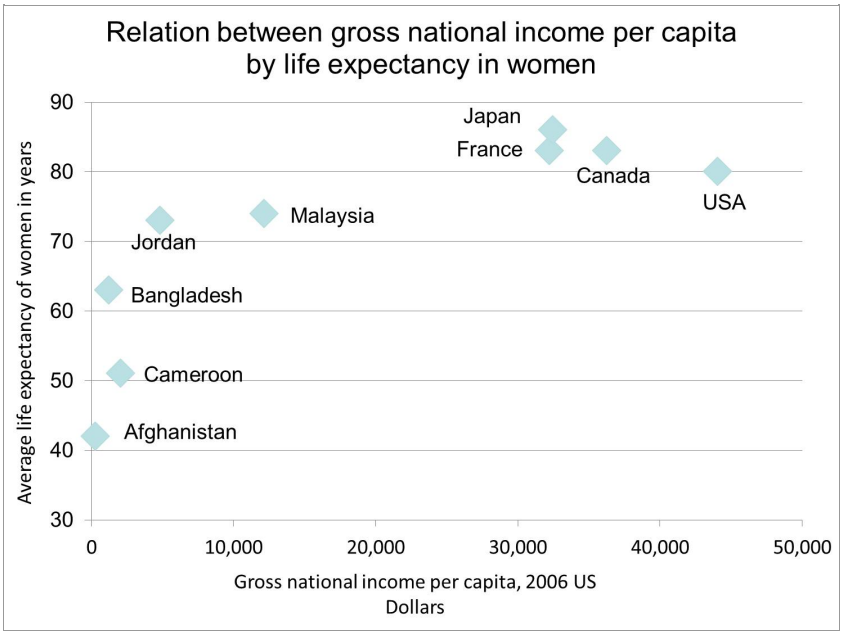


Figure 2.2: Women's life expectancy, and gross national income per capita, in selected countries, 2006 (Statistics from the WHO and World Bank reports).
Discussion Point: How could you explain the patterns shown in the graph?

Basic concepts: causes, risk factors and determinants

Case Study

Paul Richards has been having some chest pain. He first noticed it when he was collecting wood for the family stove. He doesn't think it is anything serious; he doesn't feel old and he quit smoking, although recently he has been having the occasional cigarette when he is with his friends: they all smoke, so it is can be difficult to refuse. Julie reminds him that his father died aged 62 of a heart attack and insists that he consult Dr Rao about it.

With Paul's family history, lifestyle and pattern of pain, Dr. Rao suspects he has mild effort-induced angina. They discuss the risks and benefits of exercise, particular those of the intense exertion involved in hauling lumber coupled with his overall inactivity and lack of fitness. Julie comments that money is tight; Paul's disability pension is inadequate. The food bank, which they must use regularly, does not supply fresh fruit and vegetables, but mainly processed foods that are high in carbohydrates or fat and salt, so healthy eating is a problem. They rely on the wood stove in winter because wood is free as long as Paul collects it himself. It is unlikely that Paul will ever find another job because most of the jobs on offer require people with more education than he has.

Effective disease management requires management of the immediate problem, then of the patient's risk factors and lastly of the underlying determinants. For example, a patient presenting with angina requires treatment of the immediate problem: the pain and the risk of complications, then the risk factors: smoking and lack of exercise, and the underlying determinants: poverty, environment, social norms. In the classic representation often applied to infectious disease, a disease arises from a complex interaction between the person, or host, the disease agent (virus, bacterium or ingested substance, such as tobacco smoke), and the environment: see Figure 2.3. In non-infectious disease, the causal factor, the atherosclerotic plaque in a coronary artery in the case of a cardiac infarct, can be substituted for the infective agent.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

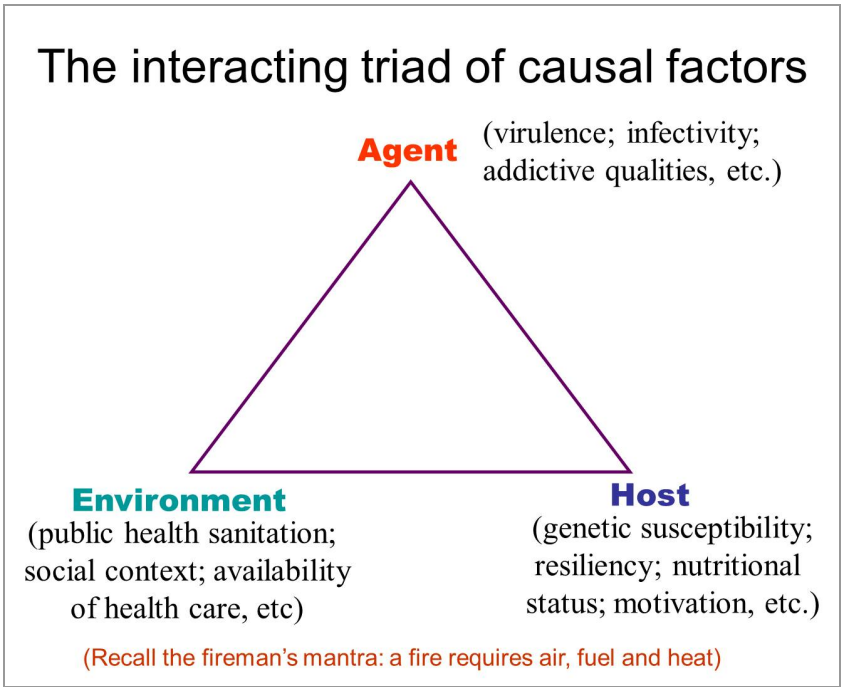


Figure 2.3: The Epidemiological Triad of agent, host, and environmental factors

Risk factors

This model is useful for explaining cases of disease after they arise, but a central problem is that we are only just beginning to understand how to assess a person's susceptibility ahead of time, so (at least until genetic analyses become much more sophisticated) we cannot accurately tell who will fall sick when exposed to a disease agent. We deal with this uncertainty probabilistically, and speak of ?risk factors' that increase the statistical probability that a person will fall sick. Some risk factors have a direct and probabilistic effect on the likelihood of an adverse outcome, as with not wearing a seatbelt in the event of a motor vehicle collision. Others form part of a complex causal pattern, as with the effect of diet on coronary disease, in which diet interacts with numerous other factors so that it is convenient to summarize the complexity in a statement of probability. While the notion of risk factors implies a causal relation, the statistical associations that identify them can also include variables that form markers for the true risk factor. Age may be an example in some instances, in which it is not age per se, but other changes that arise with age, that form the risk factor. In such instances it may be more appropriate to use the term ?risk marker' or ?risk indicator'.

But risk factors themselves have causes, termed **?**determinants'. To illustrate the distinction, Figure 2.4 extends the natural history diagram from Chapter 1 back in time to include the antecedents of the health condition. This represents an ecological model: we are now situating the patient within his or her social context, which is relevant both before the disease develops and during the course of the illness while care is being provided (for example, poverty will continue to exert an influence even though the patient is being treated).

Nerd's Corner

Ecology

Ecology is the branch of biology dealing with relations between organisms and their environment, including other organisms. *Oikos* (transliterated as ?eco') is a Greek word meaning habitation, house or dwelling place.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

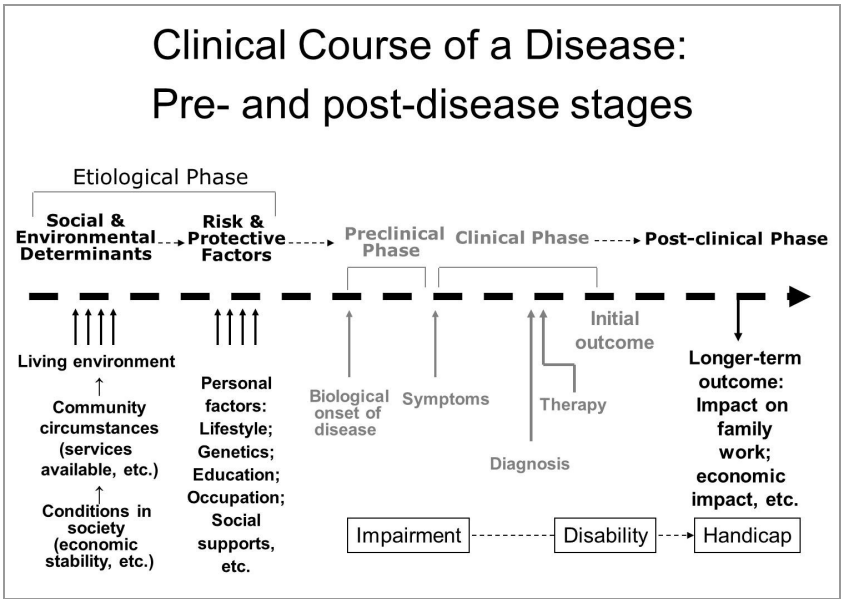


Figure 2.4: Disease precursors and clinical course

Health determinants refer to underlying characteristics of society that ultimately shape the health of individuals and communities. They can be thought of as the causes of the causes of ill health, or as ?upstream factors'. Be careful: the term ?determinant' is not intended to imply inevitability or determinism; it comes from the Latin *de termine,* or "from the end" (i.e., the origin) of the causal chain.

Links

Evidence on health determinants

For a summary of the evidence on the impact that selected determinants have on health, see the federal Public Health Agency of Canada website at:
http://www.phac-aspc.gc.ca/ph-sp/determinants/determinants-eng.php

Illustrative Materials

The Upstream-Downstream Metaphor

The term ?upstream' comes from a metaphor for the public health approach to prevention. The metaphor described a dangerous river in which several people have drowned. Rescue workers have struggled to pull them out, but come to realize that no matter how hard they work they cannot resuscitate all the victims. They discuss how to raise funds to hire more staff, and suggest that warning signs be installed. By contrast, the public health approach recommends looking upstream to figure out why people are falling or jumping into the river in the first place. Perhaps that can be fixed, for example by building a bridge or installing a fence.

Another common metaphor distinguishes between proximal factors, such as whether an individual obtains an immunization, intermediate factors such as accessibility of local health care facilities, and distal factors, such as the governmental policy that covers the provision of immunization services.

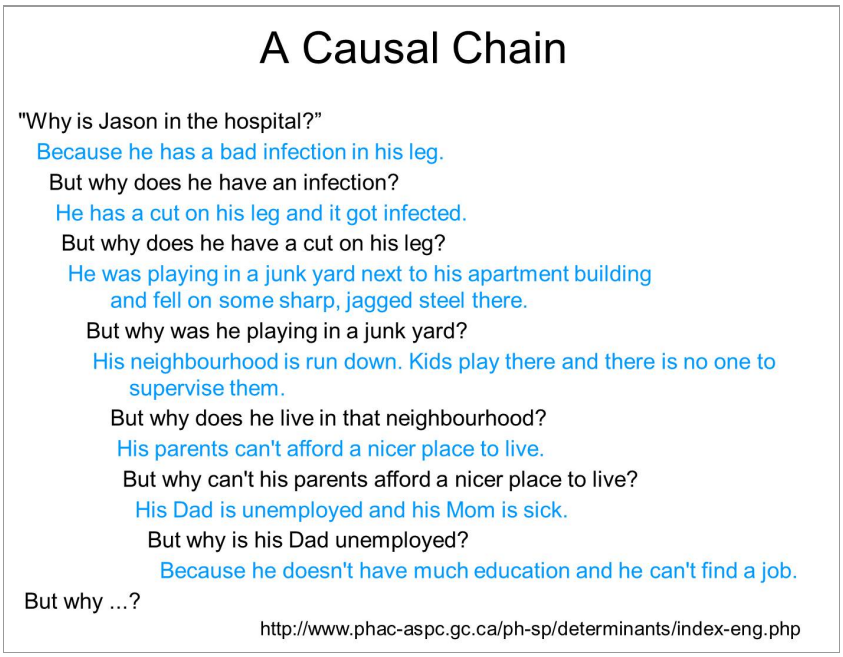
Determinants include non-specific factors (a recession, poverty, lack of education) and particular policies (such as alcohol or tobacco laws) aimed at improving health behaviours or health in general rather than particular diseases. Their influence is transmitted via a chain of intermediate processes that lead, ultimately, to a specific case of a disease. (See Illustrative Materials box).

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Illustrative Materials

A chain of causal influences

Health Canada offers the imaginary example of Jason, a boy who is in hospital because of an infected cut on his leg. This clearly presents the idea of a chain of causation (see **Cause** in Glossary), beginning from the individual case and panning outward towards the underlying social determinants, the causes of the causes. (cf The 5 Why's of ?Investigating quality' Chapter 13)



Patterns of health, such as those in Figure 2.1, may be consistent and predictable, but this does not necessarily mean explainable. Showing that, on average, poorer men live shorter lives than richer ones does not tell us *why*, and nor does it tell us *which* poor men will die young. The social gradient suggests that there is some systematic process at work, but this only takes us part of the way to understanding what it may be, and therefore what to do about it. To achieve this we need to balance analyses at the individual level with those at the group level.

Further Reading

Models of health determinants: particular or general?

An interesting debate concerns whether there can be a generally applicable model to explain the action of social determinants or whether models have to be specific to particular groups; this was briefly discussed by Reading and Wein for Aboriginal populations[5], and was also discussed by Dyck for the Métis population (she provided a Métis version of the story of Jason, see illustrative materials box above). M. Dyck's "Social determinants of Métis health" can be viewed at the National Aboriginal Health Organization (NAHO) website:

http://www.naho.ca/metiscentre/english/documents/Research_SocialDeterminantsofHealth.pdf

The factors that influence health can be considered at the individual level (What made this patient sick?) or at the population level (Why do people in British Columbia live longer on average than those in the Yukon?) British epidemiologist Sir Geoffrey Rose drew a distinction between the causes of individual cases and the causes of patterns of incidence in a population.[6,7] The distinction is subtle but valuable in thinking about the logic of making a diagnosis. For example, southern Uganda and Jamaica have similar climate and vegetation, but long ago Jamaica had the socio-economic resources and political will necessary to implement effective malaria eradication activities, so that malaria is now quite rare in Jamaica whereas it remains common in southern Uganda. Underlying determinants determine which diseases are common in a population. This difference in prevalence will, of course, guide your differential diagnosis when you see a child with a fever: in Uganda malaria will be at the top of your list, while in Jamaica, it will be low down. But determinants do not indicate which child will get malaria; for that we need to refer to individual-level factors such as whether or not the child sleeps under a bed net (i.e., the clinical history). Those who consistently use bed nets may reduce their risk below the population norm, while those who do not will have an incidence above the population average. Thus, explaining an individual case requires that we consider both underlying causes of disease incidence in the population and individual circumstances that modify the base-rate for the individual.

Details of the distinction between determinants and risk factors can be debated: many authors consider the population distribution of a risk factor as forming a determinant for individual cases. Thus, as more individuals choose to use bed nets, demand rises and the price falls, and using them become the norm, and hence a social determinant of health.

Finally, diseases themselves fall into a spectrum according to their pattern of causation, running from exclusively genetic conditions (such as cystic fibrosis) through conditions such as breast cancer, autism, or Type 2 diabetes that reflect both genetic and environmental influences, to others such as environmental mercury poisoning or post-traumatic stress disorder that are principally due to population-level determinants.[8]

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

The ecological model of health introduced in Chapter 1 is often represented by concentric circles of determinants around the individual, as illustrated in Figure 2.5. The south-west part of the figure illustrates the idea that environmental determinants exert their influence from the outer circles inwards: national policies influence communities, neighbourhoods and then individuals. At the same time, the SE part suggests that behaviours of individuals can also influence their environment, so that risk factors from inner levels aggregate up to form outer level influences. For example, patterns of health behaviours in families aggregate to produce broader patterns at the neighbourhood and ultimately community levels. The upper half of the diagram illustrates two ways of thinking about explanation: the "How?" and "Why?" questions. Explanations of why a disease is prevalent probably need to seek answers in the outer circles in the diagram (NW corner), whereas questions concerning mechanisms focus more on the inner circles (NE side). A physician treating an individual patient directs management towards the inner circles: the patient and their immediate environment such as their family and perhaps their work setting. By contrast, public health physicians address broader determinants (government policies and environment), and thus work on the middle circles. Meanwhile, the Minister of Health will need to consider the broadest determinants portrayed on the outer circles of the diagram. Together these differing approaches are intended to form a comprehensive and complementary set of approaches to improve the health of individuals and of the population at large.

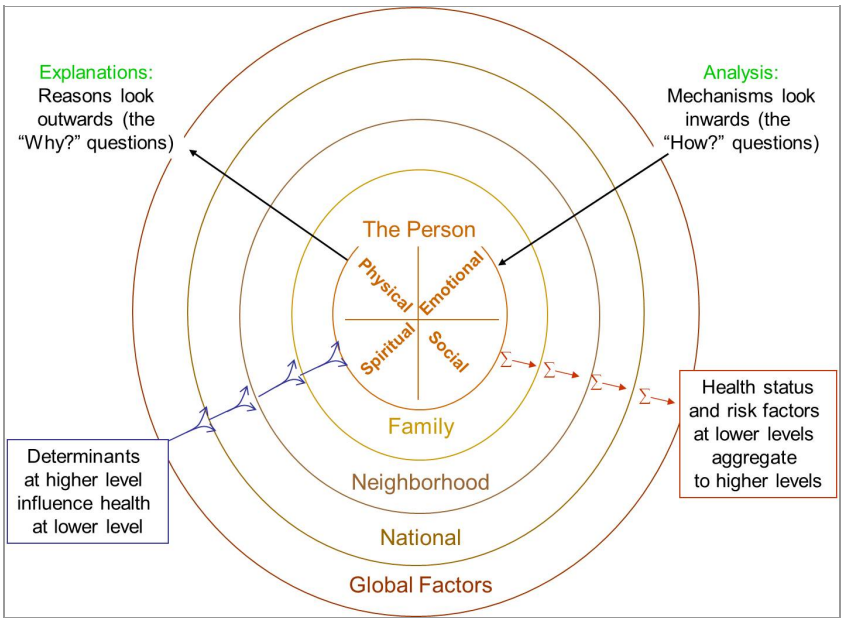


Figure 2.5: An ecological model of the influences on health

Links

Those interested in further reading could begin with the articles by Pearce N. Am J Public Health 1996;86:678-683; Susser M. Am J Public Health 1996;86:674-677; or

Rothman KJ. Am J Public Health 2005;95(1, supplement):S144-S150.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

Reductionism versus emergence

Chapter 1 introduced two ways of thinking about population health. The first viewed population health as the sum total of the health of individuals in that population. This is a ?bottom-up' or reductionist view: the health of the group can be reduced to the average of its members. The alternative, emergent view is that the health of a population is more than the sum of the health of its members, but also reflects the functioning of the whole organism: the product of the interactions among its members, such as the institutions that are created within the society. This represents a dynamic ecological model: healthy individuals contribute to the health of the population but, in turn, individual health is promoted by living in a healthy population, as implied by the arrows in the lower half of Figure 2.5. The emergent view would consider a society of fit people who are co-operating harmoniously with each other as being healthier than a society of equally fit people who were living in a state of tension and mutual conflict. Thus the emergent perspective holds that the health of the whole cannot be described, and even less explained, only in terms of the health of individuals: the functioning of the society has to be considered.

In the same way, patterns of health, such as health disparities, can only be understood by examining the dynamics of the whole society. Health disparities do not exist simply because there are always some people at the tail of the health distribution and they just happen to also be at the tail of the income distribution. They exist because of systematic characteristics of the way that society works?of the interactions among groups of people. In this view, individual characteristics, such as **ethnicity**, are interpreted at the population level in a way that produces systematic reactions, and these may influence the health of individuals within those groups. This ?top-down' perspective reflects Rose's view: the causes of patterns of health at the population level are different from the causes of health at the individual level.[6,7]

The evidence in support of the emergent view includes the correlation of national mortality statistics with national income statistics: up to a certain level of national wealth, a country's mortality is correlated to its income. Beyond this baseline level, the correlation between mortality and national income becomes weaker (Figure 2.2), but a correlation between mortality and the *distribution* of income within the country becomes apparent. Wealthy countries with a wide income gap between rich and poor tend to have less good health status (higher mortality, shorter life expectancy) than those with more equitable distribution of income. This is taken to indicate that, above a certain level of income, the nature of income distribution in a country has a more powerful effect than the aggregate incomes of its citizens. (See Nerd's corner: Relative versus absolute income) Interpretations of this type are being actively debated. See, for example, a discussion in the International Journal of Epidemiology on this topic.[9] The article by Frolich and Potvin will lead you to other articles in the series.

Determinants of Health

Almost every characteristic of a society can affect the health of its citizens, so could be considered a health determinant. To make this topic manageable, agencies such as Health Canada and the WHO have identified key health determinants that deserve special attention; the major ones are reviewed below, while the supplementary materials box shows a fuller listing of determinants.

Links

Public Health Agency of Canada list of health determinants[10]	World Health Organization list of social determinants of health[3]
Income and social status	Improve daily living conditions:
Social support networks	Early child development
Education and literacy	Urban planning
Employment & working conditions	Fair employment & decent work
Social environments	Social protection systems
Physical environments	Universal health care
Personal health practices and coping skills	Tackle the inequitable distribution of power, money and resources:
Healthy child development	Make health equity a priority for all government sectors
Biology and genetic endowment	Allocate resources to health
Health services	International regulations
Gender	Gender equity & empowerment
Culture	Good global governance
	Measure and understand the problem and assess the impact of action:
	Monitoring, research, training

Early childhood development

In treating a disease, a doctor intervenes in a process that was often established many years earlier. In thinking about causes, many take a **life course** perspective that extends the disease origins even further than that shown in Figure 2.4, sometimes even to exposures faced by a person's parents. The life course perspective emphasizes that early experiences have a profound formative impact on an adult. Early nutrition, physical development and fitness are important, as is emotional development which, if positive, builds resiliency, and if negative, enhances vulnerability. Timing of exposures and experiences can be critical.[11] For example, traumatic experiences in early childhood shape personality and have a lasting impact on how a person views his world, how he relates to others, and how he interprets events. The impact of broken homes, chronic childhood

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

stresses and so forth have non-specific effects, acting mainly to increase emotional vulnerability in adult life. The recognition of critical periods in early child development has led to the popularity of infant stimulation programs, such as Head Start.

Links

Head Start

Head Start was established in the U.S. in 1965 as a way to ensure that young children from impoverished families were ready for entry into kindergarten at age five. They offer a comprehensive preschool and family support program for families with low incomes, preparing children to succeed in their early educational experience. Programmes typically offer access to prenatal care, child care and nutritious meals for young children. In Canada, Head Start programmes were developed for indigenous peoples starting in 1995; there are now 125 Aboriginal Head Start sites in urban and northern communities across Canada (seehttp://www.phac-aspc.gc.ca/dca-dea/programs-mes/ahs_main-eng.php).

Income and social status

Figure 2.1 gave just one example of a link between a person's social status and their health. A person's social status is defined by a combination of their wealth, education, occupation and lifestyle, and (secondarily and to a varying extent) by other factors such as ethnicity, personality, and happenstance. Each of these, alone or in combination, can exert positive or negative influences on a person's health. Whatever marker of social status is used (wealth, education, occupation, or power), and whatever the health indicator (longevity, death rates, morbidity, or self-reported distress), there is a universal tendency for those in higher social positions to enjoy better health. Exceptions are rare and often transient (an example occurred in the 1930s, when heart disease was on the rise, and cases often occurred in richer people who could afford a cardiogenic diet and sedentary lifestyle. Subsequently, the decline in heart disease has occurred more steeply among richer people so that the familiar social class gradient now holds).[12] The association between social status and health is now termed **social inequalities in health** (see Health inequality in Glossary) and is summarized in many sources.[1,13,14]

Nerd's Corner

Indicators of social position

A person's position in society influences his or her health in a myriad of ways: through lifestyle (diet and opportunities for healthy exercise will vary according to wealth), educational level (which may affect understanding of health risks, and may influence ability to follow treatment guidelines), occupation (which may pose health hazards for those in dangerous or stressful occupations?Paul's accident was caused by working in a dangerous occupation at a particularly stressful time), living conditions (Mrs. Sulawesi's son Peter's asthma attacks are likely exacerbated by the damp conditions in their apartment), and so on.

Social position cannot be measured directly, but must be described via one or several indicators (rather as health itself cannot be measured directly). The commonest indicators are income (or overall wealth), education, and occupation. There is no simple guideline as to which indicator of social position is best, and the choice will depend on availability of information, and sometimes the nature of the health condition will dictate one indicator or another (a study of occupational lung disease will probably use occupational classification as the main indicator). In non-specific applications each indicator has advantages and limitations. Income offers a useful predictor of lifestyle characteristics, such as diet or living conditions, which may be relevant for some health conditions. But income usually has to be adjusted to consider the number of people the income is supporting and income may give a poor indication of lifestyle for groups such as students, whose poverty is presumably temporary. Educational attainment has the advantage of remaining relatively stable for people over about 25 and to some extent it will predict income, and can also prove useful in predicting conditions such as cognitive impairment. Occupational group can be difficult to classify and, like income, can change rapidly making an occupational history difficult to summarize. But occupation can be very important as a health predictor, for people with identical incomes and education but who work in very different environments are exposed to very different health risks. Occasionally, studies use a composite indicator of social position based on all three indicators, but more commonly one is chosen, based on the nature of the health condition and the availability of the data.

The most obvious of these influences is income – especially insufficient income – described either in relative or absolute terms. Absolute poverty refers to having inadequate resources to meet basic needs for shelter, nutritious food, clothing, and education. People living in poverty lack the resources and opportunities to make choices that promote good health. Being poor may also expose them to inferior physical environments that place them at risk for health problems. Absolute poverty is the leading health determinant in low-income countries; infants and children are particularly susceptible to its effects. In wealthy countries, however, it is not only the very poor who suffer health disparities, as was shown in the health gradient in Figure 2.1.[15] The existence of the gradient in health across income groups indicates that relative income, rather than absolute poverty, should be used in analyzing health inequalities in middle and higher income countries. Relative poverty denotes an income level that is substantially lower than that of other groups within the community, and moves beyond a simple binary contrast between rich and poor, to represent the social gradient (see Nerd's Corner box).

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

Relative versus absolute income

An income gradient in health does not necessarily mean that it is income itself that is producing the effect; income may be a marker for a range of possible causal factors (living environment, lifestyle, etc). Indeed, relative social position is often a stronger predictor of health than absolute income level, suggesting that it is not the actual money that is influential. A person who earns $20,000 per year tends to have better health if that is the average income in the country versus when this income places him low down on the income ranking.

If relative position influences health, it follows that population health will be less good in a society where there is a wide spread in the social hierarchy, where there are both very rich and also very poor people, compared to a society with a narrow range of hierarchy or income. Hence income disparity is an important determinant of health: among richer countries the *range* of income may have a far greater impact on population health than the *mean* income level. There has been much discussion of this provocative finding, and several plausible explanations have been proposed. Some of the channels through which social status and disparities in status within a region affects health are described by Wilkinson and Marmot, and by Kawachi et al.[15,16]

Case Study

We know Paul's disability and precarious work situation may be a source of economic hardship and may affect his standing in the community. He has smoked and gained weight, both of which contribute to his risk of cardiovascular and other diseases. But it may be his perception of his position in life, his status that drives his behaviours. Belief that one's status is low is more likely to be an issue in a society with wide disparities than in a more egalitarian society. Dr. Rao may need to incorporate this awareness in his approach to Paul's behaviour.

Illustrative Materials

Working with patients in poverty

Many health agencies are designed to deliver care to poor people. For example, Community Health Centres in Ontario often treat uninsured and homeless people, creating an atmosphere in which these people can feel at ease. CHCs are staffed by a medical and social services team that includes physicians, nurse practitioners, nutritionists, social workers, and community outreach workers. They tailor programmes to the needs of their local clientele, and may provide interventions such as a harm reduction unit, workshops for recent immigrants, classes on cooking on a low budget, and so forth. Members of ethnic groups can often use the centre after hours for meetings and to arrange mutual support.

Education and literacy

Education is one of many characteristics that both contribute to, and also result from, social position. A person's social position in childhood influences their access to educational opportunities. The resulting education influences their social position in various ways: by its impact on employment opportunities, by determining their income, and by influencing whom they meet and where they live. Each of these factors indirectly influences a person's health. But education also has a more direct influence on health in that it affects a person's ability to navigate the health care system, to interpret health information and to communicate effectively with physicians and other professionals.[17,18] ?Health literacy' refers to the patient's ability to understand health information and to follow guidelines for their treatment. Physicians need to be aware that a large fraction of their patients may not be able to understand information concerning their health in the format that it is given. Educational attainment can also act as a risk factor for certain conditions; dementia offers an example. Here, a lack of education may lead to careers that expose the worker to neurotoxic substances that damage brain function. Conversely, higher education, and the stimulating careers that go with it, may build complex neural networks that protect the aging brain.[19]

Case Study

As a child, Paul was not interested in school. He had difficulty keeping up with his class. His parents didn't see the importance of schooling: there were solid jobs in mining which didn't require schooling. However, as the mining industry came under pressure, company managers began to take short cuts around safety legislation. Paul began to realise that his job was insecure, this worried him, but he couldn't see a way out, he was forced ignore safety procedures and work long hours. It was at the end of a 12 hour shift that Paul, carrying a heavy, awkward load, fell and twisted his neck, resulting in the injury that put him off work.

Social support networks

Mrs Sualwesi lacks a supportive social network. No-one else here speaks her native language and she feels alone in a strange place. Social support benefits health in several ways. It is a source of emotional reassurance and provides a safe place for a person to discuss his problems, which helps him to cope with adversity. Social support provides information and practical support, such as knowing someone who can assist in a time of need. It can also support people in making healthier behaviour choices. The 2003 General Social Survey showed that self-reported health was positively associated with having a network of strong social ties, with belonging to organizations, with reciprocity (those who both give and receive assistance) and volunteering, and with receiving help to carry out daily activities.[20]

The same association between social ties and health holds at the population level: healthy communities establish collaborative networks that help

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

them address social and economic issues. **Social capital** refers to people's willingness to collaborate in groups and engage in collective action, which in turn reinforces trust and confidence within the networks; neighbourhood watch programmes are an example. By contrast, low social capital is characterized by an unwillingness to collaborate with others who are perceived as different and typically occurs where there are wide disparities in income and a perception of social inequalities. Kawachi et al. showed that low social capital was related to higher mortality, while membership in social groups was linked to lower all-cause mortality.[16] A cross-national review of research identified links between social capital and improved health, greater well-being, better care for children, and lower crime.[21]

Employment, working conditions, and occupational health

The WHO recognizes fair employment and decent work as a cornerstone of health, and advocates for fair minimum wages, full employment, and occupational health and safety standards.[14] Although Canada's overall unemployment rate (6.3% in 2008) is enviable in international comparisons, there are inequities across the regions of Canada and across different population groups (for example, 11.5% of recent immigrants are unemployed and many more are underemployed).[22]

Work stress affects many Canadians, especially women.[23] It coincides with other determinants, such as income, in that the lowest income households report high rates of work stress due to job insecurity and dissatisfaction. Workers who perceive work insecurity experience significant adverse effects on their physical and mental health.[24] Work strain arises from a combination of high psychological demands (such as having to work fast) at the same time as having little freedom to make decisions affecting the job (e.g., being in a subordinate position)[25]. All of these challenges describe Mrs. Sulawesi's job. It also results from a mismatch between work effort and reward: jobs that demand high effort for low gain produce feelings of strain that predict poor health.[26] A study that followed 10,000 British civil servants over 10 years showed that those who had little control over their work environment had an increased risk of subsequent heart disease. Providing a greater variety of tasks and more decision-making power at work may decrease risk.[27]

Definition

Work stress (see **Work-related diseases** in Glossary) derives from the amount of a person's perceived control over demands at work, their work satisfaction, perceived levels of physical risk, and job security.

Case Study

Currently, Paul is not employed and is receiving Worker's Compensation benefits. We do not know his level of work stress at the time of his accident, but it is likely that mining created significant physical and perhaps mental stress associated with high-risk work. While Worker's Compensation may address some of his financial worries, it is not going to match the income he earned as a miner, and his prospects for future employment are limited. The symptoms he is experiencing are worsened by this stress. This further reduces his likelihood of finding work.

Illustrative Materials

The strain of family duties: caregiving

Having a sick or aging relative often imposes unpaid work, especially for women. For caregivers who are also employed, it can increase work stress and harm their job security due to absences to care for the relative. The resulting reduced income affects the whole family. In 2006, the Compassionate Care Benefit under the Employment Insurance Program was introduced to reduce stress by providing additional income to individuals caring for sick relatives.

Occupational diseases are disorders that result from conditions in the workplace, typically from exposures to physical, chemical, and perhaps psychological hazards. Asbestos exposure leading to mesothelioma is an example. These are environmental hazards that occur in concentrated form in workplaces; treating and preventing such diseases forms the purview of **occupational medicine**(see Occupational health in Glossary). Occupational diseases may be distinguished from **work-related disorders** (see work-related disease in Glossary), which are conditions that originate in other domains of the patient's life but are exacerbated by their working conditions. Cardiovascular disease or low back pain, for example, may develop over the long term and could have occurred in the absence of work strain, but may be aggravated by a person's work. Occupational diseases may be distinguished from **occupational injuries**, which result from trauma such as strains or sprains, lacerations, burns, or bruises acquired in the workplace. Occupational injuries mainly result from mechanical factors, such as lifting or bending, or from failures in safety measures. Work injuries are an important source of disability and mortality in Canada. In 2003, 630,000 workers had at least one activity-limiting occupational injury, occurring more commonly among men than women (5.2% versus 2.2%). The lower a worker's income, the more likely they are to experience a work-related injury.[28]

The field of occupational health involves the management and prevention of occupational diseases and injuries, as well as the improvement of work settings in general. It is based on the simple idea that work and health influence each other, at times positively and at times negatively. The role of the occupational health physician is to maintain a positive relationship between the two.

The physical environment

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Environmental influences on health can be positive or negative, and cover a wide range of factors, from global (climate change) to national and regional issues (economic recessions, strife, air, and water pollution) to issues in the local built environment (indoor air quality), to the social environment. The positive benefits of spending time in beautiful surroundings are well understood, but most medical research focuses on negative aspects of the environment. Exposures to contaminants in air, water, food and soil, are associated with many chronic diseases and with emerging communicable diseases. Climate change and the associated weather extremes will also affect health: hyperthermia from extreme heat, injury from extreme wind and rain, social disruption from sea level change and agricultural effects, and changing distribution of vectors and infectious agents introduce disease to previously unaffected regions.

Definition

Environmental health considers all the physical, chemical, and biological factors external to a person that may affect their health, and also social factors that influence health behaviours. It encompasses the assessment and control of these factors in order to prevent disease and create health-supporting environments

Links

For further information on environmental health, the Canadian National Collaborating Centre for Environmental Health is one of six public health collaborating centres.It undertakes systematic reviews of environmental health hazards and prepare policy guidelines. The web site includes literature reviews, news reports, and policy statements and describes innovative practices. See http://www.ncceh.ca/

On a global scale, the World Health Organization maintains and environmental health website at: http://www.who.int/topics/environmental_health/en/

Nerd's Corner

Air pollution in Canada

Environment Canada reports trends in emissions in the air. Emissions come mainly from metal smelting and refining, followed by electricity generation. Emissions of sulphur oxides in Canada declined steadily from 3.73 million tonnes in 1985 to 1.9 million in 2007. Trends for nitrogen oxides, of which about 50% come from transportation emissions, have held roughly steady over the same time period, running from 2.66 million tonnes to 2.47 million in 2007. Carbon monoxide (CO) emissions have declined from around 16.8 million tonnes in 1985 to 11 million in 2007. The chief source of CO is transportation, which declined steadily with tighter vehicle emissions standards, but CO also comes from natural sources such as forest fires. The level varies widely, from a low of 760,000 tonnes in 1985 to a high of 9.4 million in 1989. Respiratory health is also influenced by total particulate matter (TPM) in the air. TPM from industrial sources was reduced by almost half in the period 1985 to 2007, but this forms only a tiny fraction of the overall problem: roughly 95% of all TPM comes from "open sources" (dust from roads, agriculture, construction sites), which contribute around 17 million tonnes of TPM per year?and these sources have been steadily rising.

Air pollutants can also cause climate change which will affect communities differently. "Communities may have a limited capacity to adapt to climate-related events, due to poor infrastructure, limited knowledge about the risks, lack of human and social capital, or economic disparities. Outdoor workers will be more vulnerable as they are directly exposed to extreme heat events and increased levels of ultraviolet (UV) radiation. Those who live on the land and whose livelihood is tied to natural resource-based employment will also be at greater risk."[29]

Further information from a Canadian perspective can be found in reports from Natural Resources Canada.[30]

Outdoor air quality is steadily improving in Canada, although people in urban centres still suffer periods of poor air quality. Indoor air may contain constituents such as asbestos, moulds, radon gas, carbon monoxide and methane that affect health. The ?sick building syndrome' describes a range of ailments related to living or working in closed buildings with poor air quality. However, the greatest threat to healthy indoor air quality in Canada remains tobacco smoke, especially for children living in the homes of smokers. These children experience more asthma and bronchitis, are at a higher risk for sudden infant death syndrome and have more lung and ear infections than children living in homes without tobacco smoke.

Design of the built environment also influences health. Overcrowding in housing and community design (which affects transportation and neighbourhood walkability) are increasingly identified as risk factors for chronic diseases, especially respiratory conditions.[31] As people live further from the city centre they drive more, resulting in more vehicle collisions, "as well as higher rates of heart and respiratory diseases and obesity, and elevated stress related to both commuting among congested traffic and increased noise levels."[14] In rural areas, particularly, where the private car is the only means of transport and where walking or cycling is impractical?and often dangerous because of road conditions?obesity is higher than in urban areas and road traffic injuries are frequent.

Individual and Public Health Services

Advances in health care have conferred immense benefits on health and longevity for countries of the developed world and in many developing countries. The invention of antibiotics, antisepsis, life-saving surgical procedures, as well as advances in pharmaceuticals, have all revolutionized the management of a wide range of diseases.

At the population level, public health interventions (sanitation, communicable disease control measures, etc.) and disease prevention (immunization, tobacco control measures, and screening) have made major contributions to saving lives and improving health.[32] Relatively simple measures, such as protecting water supplies, often have the greatest impact on overall health. This is famously illustrated by the decline in tuberculosis in England and Wales. As Figure 2.6 illustrates, tuberculosis mortality had been declining for at least 100 years before the introduction of effective therapy for individual cases; this early decline was achieved by non-specific means such as improvements in housing and nutrition, and through legislative

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

measures such as banning spitting in public which reduced the transmission of the disease. Similar historical declines occurred for polio, smallpox, and whooping cough, all of which responded dramatically to improvements in hygiene long before effective medical treatments were developed.

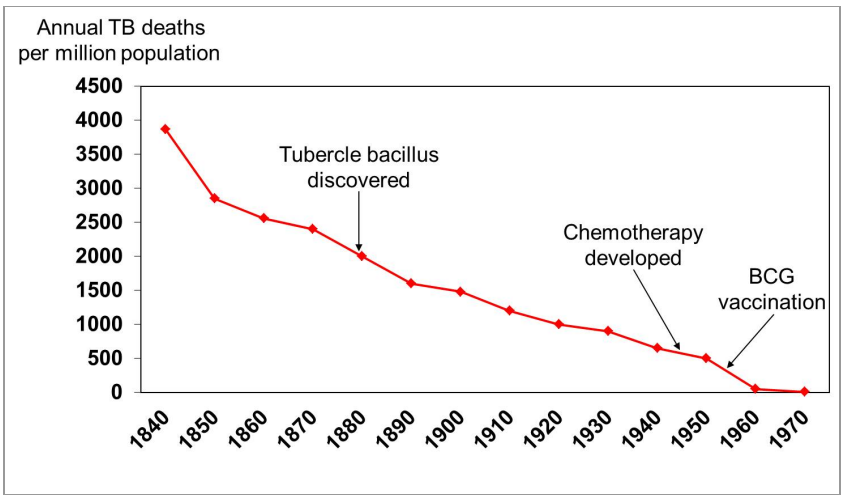


Figure 2.6: The historical decline in tuberculosis in England and Wales, 1840?1970, showing the timing of critical discoveries in understanding the disease

Links

Although most communicable diseases were in decline before immunizations were developed, vaccination has made a profound difference on incidence and death rates. For example, the Public Health Agency of Canada website includes a graph that illustrates how the rising use of hepatitis B vaccination is linked to a significant reduction in the incidence of the disease. (http://www.phac-aspc.gc.ca/im/vpd-mev/hepatitis-b-eng.php).

Public health interventions to improve air or water quality benefit all citizens, whereas the benefits of individual medical care are less evenly spread. Roughly four million Canadians do not have access to a family physician.[33] Despite universal insured health care, access to medications, dentistry, home care, and home support can be significantly affected by income, place of residence, and even ethnicity. Practitioners' offices tend to be built in richer neighbourhoods, where property values will be preserved and quality of life for their employees is more attractive, but where the need for care is lower. The capacity to take advantage of health promotion and prevention programs is influenced by income and education level: facilities relevant to promoting health, like commercial exercise facilities, are less likely to be built in lower-income or immigrant neighbourhoods.[34] Even when financial barriers to care are removed, we run the risk of continued health inequalities if programs are not accessible because of location or because of the person's feeling of discomfort at using facilities. Note that health care facilities are typically located in places that are more readily accessible to the urban rich. A number of specialized services in Canada are now regionalized, making it more difficult for people living in low-income, rural neighbourhoods to access them. It is important to develop clinics that simplify access for disadvantaged groups: clinics for street dwellers, workers in the sex trade and drug addicts; family planning clinics for teenagers; and clinics where staff members speak minority languages and are culturally sensitive.

Gender

Sex refers to our biological identity as male or female. Women in contemporary Canada on average live longer than men, as illustrated in Figure 2.1. Biological differences between the sexes, along with contrasts in the distribution of other health determinants (lifestyle, status and access to services) contribute to this difference. The former are termed ?sex differences', while the latter are ?**gender** differences' and the impact of these two may pull in different directions. Largely as a result of the women's movement, gender disparities have received considerable attention. These occur in all societies in terms of power, freedom, resources and values, and these disparities may affect health. Because disparities (see **Health disparities** in Glossary) are socially generated they should, in principle, be correctable. The most egregious examples include the exploitation of women in the sex trade, underage marriage, or the work conditions in sweat shops in the garment industry. In developed countries there are often gender inequities in income, such that wages for women (outside of the agriculture industry) may be around 20% lower than those for men.[35] Because single parents are almost always women, the lower income of women also affects their children.[14]

Definition

Gender refers to "the array of socially constructed roles and relationships, personality traits, attitudes, behaviours, values, relative power and influence that society ascribes to the two sexes on a differential basis.[36]"

Gender inequities do not only result from income disparities; gender is also linked to differential access to health services, to unequal obligations to provide unpaid family care duties, and to disparities in nutrition.[13] Gender inequities also exist in health research: studies evaluating various medications historically were more likely to use male experimental subjects. The interaction between sex- and gender-related factors and cardiovascular disease outcomes in women remains an important research area.[37] For instance, women present cardiac disease differently than men, and are less likely to be diagnosed accurately and to receive timely treatment, and the benefits of many interventions vary between men and

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

women.[38,39]

Culture

"Culture is a system of ideas, values, and metaphors that are consciously and unconsciously used or enacted by people in their everyday lives. It is not a rigid set of behaviour traits but a fluid and adaptive system of meaning."[40] A person's cultural background has an important influence on his or her beliefs, behaviours, perceptions, emotions, language, diet, body image, and attitudes to illness, pain, and misfortune?all of which can influence health and the use of health care.[41] Culture "explains what one must know and do to function in a given society.[42]" This section outlines the ways in which culture may influence health, while Chapter 3 explores the ways that clinicians can incorporate cultural awareness into their daily practice of medicine.

Science may be universal, but access to it and its interpretation are filtered differently in different cultures. The knowledge we acquire is influenced by our culture, which determines, for example, the subjects we study and books we read in school. Furthermore, when knowledge is lacking, cultural beliefs often take over. No one knows what happens after death, but all cultures prescribe characteristic rituals concerned with death that are based on beliefs. Culture also underpins values, which are deeply held beliefs that define what is desirable and moral. Values influence expectations of behaviour, such as the way a doctor's actions are perceived by the patient. Interventions designed to support healthy behaviours have been most successful when they consider the culture of the target population and when the community is actively engaged in designing and implementing the intervention.[43] However, although cultures may be shared, people are far from homogeneous, and it is dangerous to assume that all people defined as belonging to a certain culture will hold the same norms and values or will react the same way to new ideas and knowledge. Some errors can be avoided by being careful to view culture as influencing behaviour within each specific context, rather than in general.[41]

Illustrative Materials

Sharing and Transmitting Culture

Culture is learned and is conveyed from generation to generation through the process of socialization. Parents transmit cultural values, but so do peer groups and schools. Children of immigrant families often receive conflicting messages from parents and peers. Much adolescent rebellion, with significant health consequences such as drug use, has roots in the young person's struggle to define which cultural code to follow.

Culture is not unitary. Most societies identify a mainstream culture and various subcultures. These may be defined in terms of age (teen culture), lifestyle (gay culture), ethnicity (West Indian), location (street culture), or even health problems (drug culture, Alcoholics Anonymous). Most individuals occupy two or more cultures at the same time.

Definition

Ethnicity is an imprecise term that refers to a collective identity based on a combination of race, religion or a distinctive history. An ethnic group shares cultural customs that distinguish it from neighbouring groups. Ethnicity differs from race in that the shared characteristics are values, norms and ideas rather than physical characteristics. Ethnic groups are generally sub-groups within a culture or within a racial grouping. Ethnicity may refer to how a person describes him- or herself in terms of ancestry, history and culture.

Race is a quasi-biologically defined classification of people based on shared genetically transmitted physical characteristics: "A division of humankind possessing traits that are transmissible by descent and sufficient to characterize it as a distinctive human type." Race is not a scientifically rigorous classification: there is a huge amount of mixing among races; defining racial characteristics do not appear in all individuals, and there may be more genetic differences within a race than between races.

Multiculturalism is the recognition of racial and cultural diversity and respect for the customs and beliefs of others. It includes the right to equal opportunity and recognition regardless of race,colour or religion.

Prejudice is the holding of unfounded ideas (generally negative, but can also be positive) about a group (whether a race, class or ethnic group). These ideas are resistant to change and are rarely open to logical discussion.

Combinations of determinants

The preceding paragraphs discussed each health determinant individually. However, the different determinants often occur in association in individuals and in communities. For instance, a person with little education is likely have an unsatisfactory, poorly paid job and to live in poor housing. At the community level, the neighbourhood in which this person lives is likely to be undesirable, perhaps near an industrial complex with heavy traffic. It unlikely to have many services close by and, having little social cohesion, it is unlikely to develop enough political power to force improvements. This aggregation makes it difficult to tease out the individual determinants that are linked with individual health outcomes. Furthermore, the different determinants can create feedback loops. For instance, overcrowded housing causes increased transmission of infection which causes increased time off work which causes decreased income which forces people to live in overcrowded housing. This idea of multiple associations, links and feedback loops is known as the web of causation.[44]

Health Behaviours

Many of the underlying health determinants influence health via personal behaviours. For example, a lack of exercise, high-fat diet, and alcohol or tobacco consumption form personal risk factors for disease, whereas regular activity and a healthy diet confer health benefits.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Smoking. Over the past 25 years, the proportion of Canadians who smoke cigarettes regularly has declined steadily?a true public health success story. The annual Canadian Tobacco Use Monitoring Survey (CTUMS) has reported a decline in Canadians who report that they "smoke daily or occasionally" from 25% in 1999 to 18% in 2008; this translates into a decrease of about 1.2 million smokers. Smoking rates in Canada are lower than those in many other **OECD** countries, and yet there are important opposing pressures from the tobacco industry. Tobacco manufacturing is an important Canadian industry, forming $1.7 billion of our gross domestic product in 2004.[45,46]

Diet and Exercise. The food we eat and our physical activity directly affect our body weight. About 49% of Canadians were overweight or obese (Body Mass Index or BMI > 25) in 2007, compared to 25% in France and 21% in Japan.[47] A high BMI, sedentary lifestyle and poor diet puts people at risk of diabetes, cardiovascular disease, osteoarthritis, as well as other negative medical and social consequences. As with smoking, there are important social and industry determinants that promote obesity. For example, healthier eating options are more expensive and often unaffordable for low income families. Furthermore, salt, sugar and fat may be addictive in very much the same way as nicotine. Kessler argues that the addictive natures of salt and sugar have allowed the food industry to seduce people into overeating.[47] Exercise patterns are also affected by community design, as when people living in high-crime neighbourhoods avoid outdoor physical activities.

Understanding Health Behaviour

Clinicians generally find that altering patients' health behaviour (helping them to stop smoking, take up exercise, etc.) is slow and difficult. Psychology offers several theoretical models that identify the personal and situational factors likely to influence health behaviour; these help explain why behaviour is often so hard to change.

Links

For more information on health behaviour theories see Theory at a glance: a guide for health promotion practice which can be accessed at: http://www.cancer.gov/PDF/481f5d53-63df-41bc-bfaf-5aa48ee1da4d/TAAG3.pdf

Illness Behaviour. ?**Illness behaviour**' refers to the actions people take in response to their illness, including whether or not they seek health care and whether or not they follow the doctor's recommendations. In 1951, an American sociologist, Talcott Parsons, described the ?sick role', as a socially defined attitude towards illness that allows a person who has been diagnosed as sick to be exempt from his or her normal social roles. However, in return the sick person is obligated to try to regain health, partly by seeking competent medical care and following the doctor's recommendations.[48,49] A patient's decision to consult a physician is commonly termed demand for care. Demand may overlap with, but is conceptually distinct from, **need for care**. Quite often, patients may seek care in the absence of any clear need (particularly if need is defined from the perspective of the biomedical model of disease). Conversely, patients in need of care may not seek it: unmet need, which forms a motive to be concerned with health problems in the broader population. Need is further discussed in the box below, while ways of estimating need for care as part of planning health services will be addressed in Chapter 7.

Nerd's Corner

Defining need

Need for care can be defined in absolute terms: a person is injured and needs emergent care to save his life. One difficulty with this is to know when to stop; dying patients may be considered to need heroic interventions, but these may prolong the condition without curing it. Reflecting the CanMEDS manager role, an alternative approach is to define need in relative terms, considering the benefit that would accrue from any care that could be provided.

Using the relative approach, Acheson defined need as "the ability of an individual to benefit from health care, and [need] exists where he or she has a condition for which there is an effective and acceptable intervention."[50] In this conception, if there is no effective care that would help the patient, he or she has no need for care. While this can appear counterintuitive, from a health services perspective it sensibly points out that we should direct expenditures of scarce resources towards activities that provide benefit, and not waste resources on ineffective treatments.

A more recent synthesis between the absolute and relative perspectives holds that need and demand should not lie in conflict, but that patient and doctor should negotiate expectations for care. The service providers provide only services, and the patient and family must play their part in determining what services are to be received.

The distinction between need and demand also reminds us of the factors affecting the definition of disease. This is influenced both by demand pull (requests for treatment) and by the supply push (e.g., pharmaceutical companies marketing new products whose sale helps to fund research to develop new products to meet future demands).

The Health Belief Model (HBM)

The Health Belief Model (HBM) was originally proposed by G.M. Hochbaum in 1958 to try to explain participation in TB screening programs. It outlines three elements that affect whether or not a person will follow a doctor's recommendation to change their behaviour (begin to exercise, quit smoking, etc):

1. Does the person feel a need to take action? This is influenced by how susceptible they perceive themselves to be to the disease in question, and by how severe they judge the disease to be.

2. How effective do they judge the recommended action to be? Will it prevent the disease or reduce its severity? Are there psychological, financial, and other costs or barriers involved in the proposed action?

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

3. There must also be some cue that triggers an actual change in health behaviour. This can be internal (e.g., appearance of symptoms) or external (e.g., a friend gets sick).[51]

Figure 2.7 sketches the components in this model.

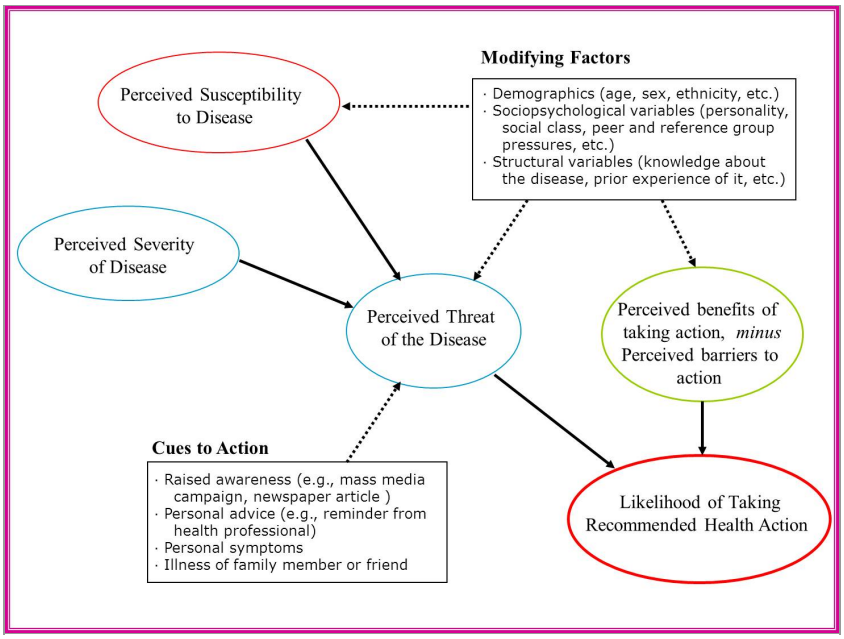


Figure 2.7: Schematic outline of the Health Belief Model

Nerd's Corner

How Accurate is the Health Belief Model?

Janz and Becker reviewed studies of the predictive validity of the HBM, concluding that prospective studies supported its validity, as did several cross-sectional studies. Perceived barriers to action appear to be the single best predictor of subsequent behaviour.[52]

Several studies have evaluated the HBM in predicting uptake of mammography screening, and the HBM has been used in guiding the design of interventions to promote screening. Janz et al. concluded that interventions that address the factors covered by the HBM tend to produce superior results, although it is often not possible from the studies to isolate the effects of the HBM from other characteristics of the intervention.[53]

The Theory of Planned Behaviour

During the 1980s, a more extensive analysis of health behaviour led to Ajzen's Theory of Planned Behaviour (TPB). This takes a cognitive perspective, meaning that it assumes that health behaviour can be analysed in terms of rational planning.[54] The TPB incorporates the ideas in the HBM (summarized in the top line of Figure 2.8), but it also considers the social context via the idea of subjective norms, which refer to perceived social pressures to behave in a certain way. These incorporate the patient's beliefs about how others view his current and proposed behaviours, and how much he values the opinions of his reference group. The third element of the TPB concerns the person's perceived behavioural control?in other words, whether or not he feels capable of making the proposed change. This derives from his perceptions of barriers to the action and his confidence in being able to overcome such barriers, a notion that resembles self-efficacy.

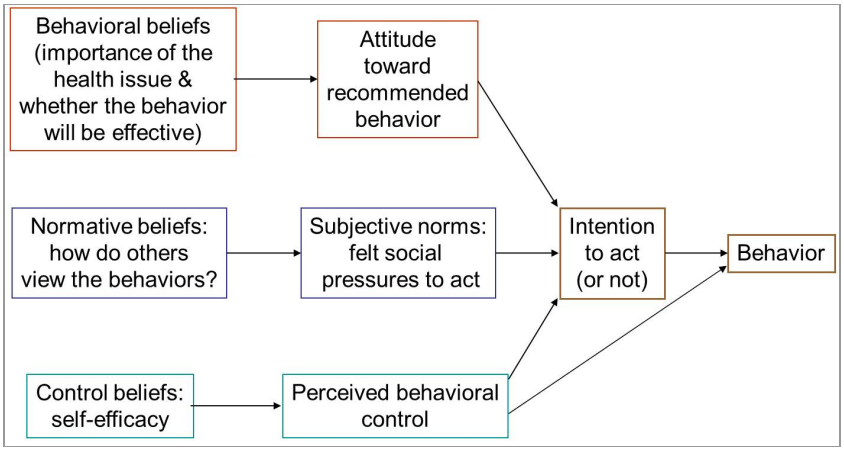


Figure 2.8: Schematic outline of the Theory of Planned Behaviour

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

How accurate is the Theory of Planned Behaviour?

There is an impressive body of literature on the validity of the TPB; there are several meta-analyses and systematic reviews of the literature. For example, a 1991 meta-analysis found that the variables in the model explained 50% of the variance in behavioural intention (based on 19 studies) and 26% of variance in predicting actual behaviour (based on 17 studies).[54] A 1996 systematic review concluded that 41% of variance in intention and 34% in actual health behaviour was explained by the TPB.[55]
In 2001, Armitage and Conner published a meta-analysis of 185 published studies, giving a sample size of well over 300,000. Most of the studies were cross-sectional, but 44 longitudinal studies gave predictive validity evidence for behavioural intentions, and 19 studies predicted actual behaviour. The results are similar: roughly one-third of the variance in behaviour was predicted from the model.[56]

These models demonstrate the broad range of factors that influence health behaviour, helping to indicate why merely offering a patient advice is unlikely to alter their behaviour. But the models also indicate that the factors that do influence behaviour are logical, so health behaviour need not appear mysterious to the clinician. Strategies for changing health behaviours, the obvious central interest for the physician, build on the Health Belief Model and the Theory of Planned Behaviour, and are described in Chapter 8.

Self-test questions

Self-Testing

1. What are the "social determinants of health"?
Aspects of the social structure, functioning, and institutions of a society that represent underlying causes of patterns of health, through numerous different channels. There are many lists of determinants; such as the Public Health Agency of Canada's list of 12: Income and social status; Social support networks; Education and literacy; Employment/working conditions; Social environments; Physical environments; Personal health practices and coping skills; Healthy child development; Biology and genetic endowment; Health services; Gender; Culture.
2. Describe the channels through which social determinants affect the health of a population.
This is the really big question that is under debate: how, precisely, do these external forces "get inside the skin"? At a superficial level, you can cite routes such as differential exposures (where you live and work, etc); differing socially determined patterns of health behaviours & lifestyle; differential access to resources such as adequate nutrition and health care that lead to differences in prevention & care when a problem does arise; different patterns of feelings and emotions that may help us cope with adversity, including health problems (despair versus confidence; sense of control or self-efficacy, etc); different levels of social connection that can offer practical assistance and reliable information, as well as emotional support.
3. What do we mean by health inequities?
Inequalities in health that place an identifiable group at a systematic disadvantage and that are (in theory, at least) preventable and correctable: things such as lack of access to health care for uninsured people; shorter life-span for poor people. Inequities offer a moral imperative for action.
4. Illustrate how gender can produce health inequities.
Note that this question can equally apply to identifiable social groups besides gender. For gender, one may begin by questioning social values: are women treated equally, allowed equal access to resources and opportunities, and genuinely respected as equals? Are there programmes in place to achieve genuine equality of opportunity (e.g., leadership programs for women)? While being respected as equals, are any differences between the sexes also respected, for example in terms of approaches to medical diagnosis and treatment?
5. What could you, as a physician, do about inequities in health?
Advocacy role: by advocating for effective, evidence-based preventive actions that can be applied universally and thereby reduce inequities. Physicians carry immense respect in society and can effectively lobby governments to draw attention to social issues. Drawing attention to the health consequences of inequities by citing specific and real-life examples forces the audience to consider the issue in personal terms: they, too, may be at risk of this disease; they are paying for expensive care for preventable conditions through their taxes, etc. Physicians can also outline effective actions to take to correct inequities: often these are very simple regulations such as banning unsafe toys; installing gates on stairs in homes with toddlers; advocating for child car seats.
6. At what stage of the life course would a public health practitioner seek to intervene to prevent the greatest amount of illness?
Most would argue that ensuring healthy early child development offers the best return on the dollar. See the work produced by Fraser Mustard.

References

1. Evans R, Barer M, Marmor T. Why are some people healthy and others not? The determinants of health of populations. New York: Aldine de Gruyter; 1994.

2. Wilkins R. The Canadian census mortality follow-up study, 1991 through 2001. Health Reports 2008;19(3):25-43.

3. WHO Commission on Social Determinants of Health. Closing the gap in a generation: Health equity through action on the social determinants of health. 2008.

4. Marmot M. Editorial: Inequalities in health. New Engl J Med 2001;345:134-136.

5. Reading C, Wein F. Health inequalities and social determinants of Aboriginal people's health. 2009; Available at: http://www.ahrnets.ca/NCCAH-report-LoppieWein-download(1)%5B1%5D.pdf.

6. Rose G. Sick individuals and sick populations. Int J Epidemiol 1985;14:32-38.

7. Rose G. The strategy of preventive medicine. Oxford: Oxford University Press; 1992.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

8. National Institutes of Health website on gene-environment interactions. Available at: http://www.niehs.nih.gov/health/topics/science/gene-env/index.cfm. Accessed July, 2009.

9. Frolich K, Potvin L. Commentary: structure or agency? The importance of addressing both for addressing social inequalities in health. Int J Epidemiol 2010;39:378-379.

10. Public Health Agency of Canada. What determines health? Available at: http://www.phac-aspc.gc.ca/ph-sp/determinants/determinants-eng.php. Accessed October, 2010.

11. Lynam M. Health as a socially mediated process: theoretical and practice imperatives emerging from research on health inequalities. Adv Nurs Sci 2005;28(1):25-37.

12. Marmot M, Elliot P. Coronary heart disease epidemiology. New York: Oxford University Press; 1994.

13. Berkman L, Kawachi I (editors). Social epidemiology. New York: Oxford University Press; 2000.

14. World Health Organization. Closing the gap in a generation: health equity through action on the social determinants of health. 2009; Available at: http://www.who.int/social_determinants/thecommission/finalreport/en/index.html. Accessed January, 2010.

15. Wilkinson R, Marmot M (editors). Social determinants of health: the solid facts. 2nd ed. ed. Geneva: World Health Organization; 2003.

16. Kawachi I, Kennedy B, Lochner K, Prothrow-Stith D. Social capital, income inequality and mortality. Am J Public Health 1997;87(9):1491-1498.

17. Health Disparities Task Group of the Federal/Provincial/Territorial Advisory Committee on Population Health and Health Security. Reducing health disparities - roles of the health sector: discussion paper. 2005.

18. Anderson A, Mortensen L. Socioeconomic inequality in birth outcomes: what do the indicators tell us, and where do we find the data? CMAJ 2006;174(10):1429-1430.

19. McDowell I, Xi G, Lindsay J, Tierney M. Mapping the connections between education and dementia in the elderly. J Clin Exp Neuropsychol 2007;29:127-141.

20. van Kemenade S, Roy J, Bouchard L. Social networks and vulnerable populations: Findings from the GSS. Health Policy Research 2006;12:16-20.

21. OECD. The Well-being of nations: The role of human and social capital. 2001; Available at: http://www.oecd.org/dataoecd/36/40/33703702.pdf. Accessed July 28, 2011.

22. Butler-Jones D. The Chief Public Health Officer's Report on the state of public health in Canada, 2008. Available at: http://www.phac-aspc.gc.ca/ph-sp/determinants/index-eng.php. Accessed July, 2009.

23. Park J. Work stress and job performance. Perspectives Ottawa: Statistics Canada; 2007.

24. Ferrie J, Shipley M, Stansfield S, Marmot M. Effects of chronic job insecurity and change in job security on self-reported health, minor psychiatric morbidity, psyiological measures, and health related behaviours in British civil servants: the Whitehall II study. J Epidemoiol Community Health 2002;56:450-454.

25. Karasek R, Theorell T. Healthy work: stress, productivity and the reconstruction of working life. New York: Basic Books; 1990.

26. Siegrist J. Adverse health effects of high-effort/low-reward conditions. J Occup Health Psychol 1996;1:27-41.

27. Bosma H, Marmot M, Hemingway H, Nicholson A, Brunner E, Stansfield S. Low job control and risk of coronary heart disease in Whitehall II, (prospective cohort) study. BMJ 1997;314:558.

28. Wilkins M. Work injuries. Statistics Canada: Health Reports 2007;18(3):1-18.

29. Judek S, Jessiman B, Stieb D, Vet R. Estimated number of excess deaths in Canada due to air pollution. Air Health Effects Division, Health Canada, and Meteorological Service of Canada, Environment Canada, April 2005. Available at: http://www.hc-sc.gc.ca/ahc-asc/media/nr-cp/2005/2005_32bk2_e.html. Accessed September, 2008.

30. Natural Resources Canada. Climate change impacts and adaptation: a Canadian perspective. Available at: http://adaptation.nrcan.gc.ca/perspective/index_e.php. Accessed July, 2010.

31. Jackson R. The impact of the built environment on health: an emerging field. Am J Public Health 2003;93(9):1382-1384.

32. Sasieni P, Castanon A, Parkin D. How many cervical cancers are prevented by treatment of screen-detected disease in young women? Int J Cancer 2009;124(2):461-464.

33. Canadian Medical Association. Presentation to the House of Commons Finance Committee. 2005.

34. Powell L, Slater S, Chaloupka F, Harper D. Availability of physical activity-related facilities and neighborhood demographic and socioeconomic characteristics: a national study. Am J Public Health 2006;96(9):1676-1680.

35. UNICEF. The state of the world's children 2007. 2007; Available at: http://www.unicef.org/sowc07/. Accessed July 28, 2011.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

36. Health Canada. Health Canada's gender based analysis policy. 2000; Available at: http://www.hc-sc.gc.ca/hl-vs/pubs/women-femmes/sgba-policy-politique-ags-eng.php. Accessed July 28, 2011.

37. Pilote L, Dasgupta K, Guru V, et al. A comprehensive view of sex-specific issues related to cardiovascular disease. CMAJ 2007;176(6):S1-S44.

38. Berger J, et al. Aspirin for the primary prevention of cardiovascular events in women and men: a sex-specific meta-analysis of randomized controlled trials. JAMA 2006;295(3):306-313.

39. Spitzer D. Gender and sex-based analysis in health research: a guide for CIHR researchers and reviewers. 2007; Available at: http://www.cihr.ca/e/34032.html. Accessed July 28, 2011.

40. Warry W. Ending denial. Toronto: University of Toronto Press; 2007.

41. Helman C. Culture, health and illness. 5th edition ed. New York: Oxford University Press; 2007.

42. Hruschka D, Hadley C. A glossary of culture in epidemiology. J Epidemiol Community Health 2008;62:947-951.

43. Panter-Brick C, et al. Culturally compelling strategies for behaviour change: a social ecology model. Soc Sci Med 2006;62(11):2810-2825.

44. McMahon B, Pugh T, Ipsen J. Epidemiologic Methods. Boston: Little, Brown; 1960.

45. Production and Disposition of Tobacco Products, June 2008. 2008; Available at: http://www.statcan.gc.ca/daily-quotidien/080729/dq080729b-eng.htm. Accessed November, 2009.

46. Statistics Canada. Production and Disposition of Tabacco Products, December 2007. 2007; Available at: http://dsp-psd.pwgsc.gc.ca/collection_2008/statcan/32-022-X/32-022-XIE2007012.pdf. Accessed July 28, 2011.

47. Kessler D. The end of overeating: taking control of the insatiable American appetite. New York, NY: Rodale Press; 2009.

48. Parsons T. The sick role and the role of the physician reconsidered. Milbank Mem Fund Q Health Soc 1975;53(3):257-278.

49. Williams S. Parsons revisited: from the sick role to...? Health (London) 2005;9(2):123-144.

50. Acheson R. The definition and identification of need for health care. J Epidemiol Community Health 1978;32:10-15.

51. Rosenstock I. Historical origins of the Health Belief Model. Health Educ Monogr 1974;2:328-335.

52. Janz N, Becker M. The Health Belief Model a decade later. Health Educ Q 1984;11:1-47.

53. Janz N, et al. In: Glanz K, Rimer B, Lewis F, editors. Health behavior and health education: theory, research, and practice San Francisco: Jossey-Bass; 2002. p. 45-66.

54. Ajzen I. The Theory of Planned Behavior. Organizational Behavior and Human Decision Process 1991;50:179.

55. Godin G, Kok G. The Theory of Planned Behavior: a review of its applications to health-related behaviors. Am J Health Promotion 1996;11:97.

56. Armitage C, Conner M. Efficacy of the Theory of Planned Behavior: a meta-analytic review. Br J Soc Psychol 2001;40:471.

AFMC Primer on Population Health

Chapter 3 Cultural Competence and Communication

Chapter Objectives

After completing this chapter the reader will be able to:

- Discuss how **culture** and spirituality influence health and health practices and how they are related to other **determinant**s of health.
- Describe the importance and impact of good culturally-appropriate communication with the patient, the patient's family and, if necessary, the community as a whole with regard to **risk factor**s and their modification.Describe how the medical, social, and spiritual determinants of health and well-being for First Nations, Inuit, and Métis peoples impact their health.
- Describe the connection between historical and current government practices towards First Nations, Inuit, and Métis peoples (including but not limited to colonization, residential schools, treaties and land claims), and the intergenerational health outcomes that have resulted.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Case Study

In the clinic staff room, Dr Rao is having lunch with the most recent recruit to the clinic. Dr Middleton joined the practice six months ago just after she finished her residency training. She has had a difficult morning. During her morning round of the palliative care unit, the chief nurse said he was having problems with the family of one of her dying patients, Darryl Crow, who is a member of a local First Nations community. Five or six family members are in the room with Darryl and they often pray out loud together with him. Other patients in the unit and their families were finding this behaviour disrespectful of their need for tranquility.

Later on, in clinic Dr Middleton had seen Myriam Fortin, a young single mother of two small children, for follow-up of her mild hypertension. Dr Middleton believes that the cause of Myriam's hypertension is her obesity and she has been trying, in vain, to get her to eat better and to get more exercise. She has advised Myriam to join the gym and also to register for the food information sessions that are offered for a modest fee in the public library. Dr Middleton confessed to feeling very frustrated when Ms Fortin countered her advice saying that all the women of her family enjoy eating and are overweight; her grandmother always said that it was a sign of good health.

Dr Middleton's last patient before lunch was Mary Monye, a young Nigerian woman, whose family moved to Canada when Mary was a child. She was complaining of symptoms that Dr Middleton felt were due to a flu-like illness that a number of patients had been complaining of recently. Mary is a well-educated, working woman in her mid-twenties who had asked for an urgent appointment that morning, but had then arrived late and had brought her mother and one of her sisters with her. Dr Middleton had found the consultation quite difficult as she had to deal with the three women all of whom insisted on contributing to the history and questioned the diagnosis and proposed treatment.

Cultural awareness, sensitivity, and safety

In the last chapter, culture was defined in terms of the shared knowledge, beliefs, and values that characterize a social group. Humans have a strong drive to maintain the sense of identity that comes from membership in an identifiable group. In primeval and nomadic times, a person's survival likely benefited from establishing strong bonds with an in-group of trusted relatives or clan-mates with whom one co-operated and shared, versus an out-group against which there was competition for scarce resources. Within the intermixing of modern society, many of us seek to retain a sense of cultural identity and may often refer to our cultural roots, or use double-barrelled descriptions such as Ukrainian-Canadian. It is important that we are all aware of our own cultural influences and how these may affect our perceptions of others, especially in the doctor-patient encounter. In many subtle ways, the cultural identities of both doctor and patient affect their interaction, and in a country as diverse as Canada this can form an exciting challenge.

Illustrative Materials

Culture and individual identity

Everyone is unique, but there are patterns of similarity. Culture refers to the shared parts, but of course we are all members of many different sub-cultures (your medical class, the soccer team, Irish-Canadian, etc.), so the joint influence of all of these may make a seemingly unique individual. But even though we are unique, most of our characteristics are shared with some identifiable group.

We all perceive others through the filter or perspective of our own cultural upbringing, often without being aware of it: communication can go wrong without our understanding why. The clinician must become culturally aware and sensitive, then culturally competent so that she or he can practice in a manner that is culturally safe.

Definition

Cultural awareness is observing and being conscious of similarities and contrasts between cultural groups, and understanding the way in which culture may affect different people's approach to health, illness, and healing.

Cultural sensitivity is being aware of (and understanding) the characteristic values and perceptions of your own culture and the way in which this may shape your approach to patients from other cultures.

Cultural competence refers to the attitudes, knowledge, and skills of practitioners necessary to become effective health care providers for patients from diverse backgrounds. Competence requires a blend of knowledge and conviction, plus a capacity for action. "A culturally competent physician considers a patient's cultural background when discussing and providing medical advice and treatment, and communicates effectively to enable patients to understand their treatment options."[1]

Cultural safety goes a step beyond cultural sensitivity (being accepting of difference) to understanding that there exist inherent power imbalances and possible institutional discrimination that derive from historical relationships with people of different origins. It implies that the care provider has reflected on her own identity and the perceptions of them that others from different cultures may hold. Culturally safe practice implies the ability to keep these differing perspectives in mind whilst treating the patient as a person worthy of respect in her own right.

Cultural competency in medical practice requires that the clinician respects and appreciates diversity in society. Culturally competent clinicians acknowledge differences but do not feel threatened by them. "Culturally competent communication leaves our patients feeling that their concerns were understood, a trusting relationship was formed and, above all, that they were treated with respect."[2] While a clinician will often be unfamiliar with the culture of a particular patient, the direct approach is often the best: ask the patient what you need to understand about her culture and background in order to be able to help her. A direct approach helps establish mutual respect and tailor the best and most appropriate care for each patient. Awareness of one's own culture is an important step towards awareness of, and sensitivity to, the culture and ethnicity of other people. Clinicians who are not aware of their own cultural biases may unconsciously impose their cultural values on other people. "As physicians, we must

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

make multiple communication adjustments each day when interacting with our patients to provide care that is responsive to the diverse cultural backgrounds of patients in our highly multicultural nation."[2]

Cultural safety refers to a doctor-patient encounter in which the patient feels respected and empowered, and that their culture and knowledge has been acknowledged. Cultural safety refers to the patient's feelings in the health care encounter, while cultural competence refers to the skills required by a practitioner to ensure that the patient feels safe.[3] To practice in a manner that is culturally safe, practitioners should reflect on the power differentials inherent in health service delivery. Taking a culturally safe approach also implies acting as a health advocate: working to improve access to care; exposing the social, political, and historical context of health care; and interrupting unequal power relations. Given that the patient exists simultaneously within several caring systems, influenced by their family, community, and traditions, the culturally safe practitioner allows the patient to define what is culturally safe for them.[4,5]

Case Study

Dr Rao and Mr Crow

Dr Rao explains to Dr Middleton that he has had similar experiences with other dying patients from Mr Crow's community. He had discussed courses of action with the staff in charge of the palliative care unit and it was decided that people from the local First Nations community should be admitted to a room at the end of the corridor, from which other patients would not hear the prayers. He remembers that it had been difficult to get agreement on this solution as some members of the administrative staff were concerned about giving the impression to other patients that they were offering preferential treatment to members of Mr Crow's community. The matter was eventually resolved after a community spokesperson was invited to a staff meeting to discuss the community's practices and beliefs relating to death. Unfortunately, since that time, there has been staff turnover so the needs of Mr. Crow's community members are being forgotten again.

Nerd's Corner

The cultural lens and how culture influences your perceptions of others

Our culture influences the way we perceive virtually everything around us, often unconsciously. Several useful concepts describe issues that can arise:

- *Ethnocentrism*. The sense that one's own beliefs, values, and ways of life are superior to, and more desirable than, those of others. For example, you may be trained in Western medicine, but your patient insists on taking a herbal remedy. You may be tempted to say "So, why are you consulting me, then?" Ethnocentrism is often unconscious and implicit in a person's behaviour. Personal reflection is a valuable tool for physicians to critically examine their own ethnocentric views and behaviours.

- ***Cultural blindness***. This refers to attempts (often well-intentioned) to be unbiased by ignoring the fact of a person's race. It is illustrated in phrases such as ?being colour blind', or ?not seeing race'. However, ignoring cultural differences may make people from another culture feel discounted or ignored; what may be transmitted is the impression that race or culture are unimportant, and that values of the dominant culture are universally applicable. Meanwhile, the person who is culturally blind may feel they are being fair and unprejudiced, unaware of how they are making others feel. Cultural blindness becomes, in effect, the opposite of cultural sensitivity.

- ***Culture shock***. Most physicians come from middle-class families and have not experienced poverty, homelessness or addictions. Exposure to such realities in their patients therefore requires great adaptations and can be distressing. This is a common experience in those who have visited a slum in a developing country, but may also arise at home in confronting abortion, infanticide, or female circumcision.

- *Cultural conflict*. Conflict generated when the rules of one's own culture are contradicted by the rules of another. For example, in Mr Crow's First Nation cultures it is normal when a person is dying to invite extended family members to the bedside, sing songs, and undertake rituals to help the person's soul on its journey; the noise can easily frustrate other dying patients who seek peace and quiet.

- ***Cultural imposition*** *(or cultural assimilation or colonialism).* The imposition of the views and values of your own culture without consideration of the beliefs of others. The history of residential schools in Canada is an illustration of cultural imposition, of which the CBC web archive gives an excellent overview, including government propaganda video clips.

- *Stereotyping and generalization*. What may be true of a group need not apply to each individual. Hence, talking about cultures can lead to dangerously prejudicial generalizations. Prejudice is the tendency to use preconceived notions about a group in pre-judging one of the group's members, so applying cultural awareness to individuals can be hazardous. Yet, on the other hand, ignoring culture (cultural blindness) can be equally detrimental. The key is to acknowledge and be respectful of differences, and to ask patients to explain their perspective when in doubt.

The relevance of culture for health

Culture influences health through many channels:

1. *Positive or negative lifestyle behaviours.* While we often focus on the negative influences of lifestyle behaviour?such as drug cultures, or the poor diet of some teen cultures, for example?we should not neglect the positive cultural influences on behaviours and practices. For example, Mormons and Seventh Day Adventists have been found to live longer than the general population, in part because of their lifestyle including the avoidance of alcohol and smoking, but also because of enhanced social support.[6]

2. *Health beliefs and attitudes*. These include what a person views as illness that requires treatment, and which treatments and preventive measures he or she will accept, as with the Jehovah's Witness prohibition on using whole blood products.

3. *Reactions to being sick*. A person's adoption of the **sick role** (and, hence, how he or she or he reacts to being sick) is often guided by his or her cultural roots. For instance, "machismo" may discourage a man from seeking prompt medical attention, and culture may also influence from whom a

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

person will accept advice.

4. *Communication patterns,* including language and modes of thinking. Beyond these, however, culture may constrain some patients from expressing an opinion to the doctor, or may discourage a wife from speaking freely in front of her husband, for example. Such influences can complicate efforts to establish a therapeutic relationship and, thereby, to help the patient.

5. *Status*. The way in which one culture views another may affect the status of entire groups of people, placing them at a disadvantage. The resulting social inequality or even exclusion forms a health determinant. For example, women in some societies have little power to insist on condom use.

Illustrative Materials

The culture of poverty

Socio-economic status influences most aspects of health, but the nature of that influence is complex and varied. Nonetheless, there are some consistencies that may be relevant for the clinician to note. The term **culture of poverty** describes certain characteristics of behaviour and outlook that may be seen among people who are living in socially disadvantaged conditions. Here are some thoughts to bear in mind when communicating with people living in poverty:

1. People lower in the social hierarchy have the same diseases as anyone else, but tend to have more of them.

2. When you are poor and sick the future looks bleak, and so you try to ignore it, to live in the present, and plan only in the short term.

3. When a person with few choices is trying to survive from day to day, long-term health is less of a priority than getting through today.

4. Behaviours that may seem irresponsible to people who are not members of the culture often have a different meaning to people in marginal situations. For instance, adolescent pregnancy may offer a route to self-esteem. Substance abuse may be an antidote to reality. Both, ultimately, serve to maintain poverty, but recognize item 3, above.

5. A person living in poverty may not be able to follow the doctor's advice; he may not have the time, the money, or the opportunities that the doctor probably takes for granted. A poor person may not have a drug plan, so may not be able to afford the antibiotic you just prescribed.

6. Do not confuse schooling with intelligence. Many people with little formal schooling have no difficulty understanding new concepts when they are properly explained.[7] "When you don't got a lot of education, you sure gotta use your brains."(Attributed to Yogi Berra)

Categories of culture

Culture concerns patterns of behaviours and values. At a higher level, these patterns can themselves be grouped and classified to help make sense of what might otherwise seem somewhat random. A basic classification will be useful for the practising clinician. A distinction is often drawn between collectivistic and individualistic societies or cultures. Collectivist cultures (including some traditional Chinese families and some African groups) hold values of sharing, of group or family solidarity, and emotional interdependence that emphasize duties and mutual obligations, as well as group decision making. By contrast, individualistic societies (such as mainstream North America) value autonomy, individual initiative and emotional independence, the right to privacy, pleasure seeking, and financial security. People from collectivistic cultures who come to North America often feel very isolated and find the way we expect people to take responsibility for their own health difficult to understand. Collectivistic cultures will generally expect families to provide strong mutual support, and decisions concerning treatment will often be made by the group. Accordingly, families from collectivist cultures commonly care for elderly relatives in the family home, forming three-generational families. By contrast, families from individualistic cultures are often those who will place their aging relatives in residential care homes.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

Further details on cultural types

The following table was based on the work of a Dutch sociologist, Geerd Hofstede, and portrays an extension of the individualist-collectivist contrast into other areas of daily life. Some general implications for health and medical care are suggested.

Identity: individualism vs. collectivism.	In individualistic countries (Canada, USA), people are responsible for themselves and initiative is valued; people have relatively weak ties to their group (work, university, organization). In collectivist societies, such as China or Japan, a person's identity is based on their group membership, so they value tight social frameworks and a feeling of belonging to a group. Mutual obligations between people will be strong. *Relevance to medicine*: People from individualistic cultures may expect to make their own decisions regarding their health. It is common for people from collectivist cultures to make therapeutic decisions as a member of a group; group pressures may strongly influence the individual, who may fear a sense of humiliation if he or she does not comply. Collectivist societies value harmony, whereas individualistic societies value self-respect, autonomy and personal achievement.
Power distance: How does the society deal with inequality? How tolerant is it of unequal power?	In some cultures, such as those in Arabic countries, in Mexico or India, subordinates tend to follow authority, to respect and accept their boss merely because of his position. In Hofstede's term, ?power distance' is high ("everyone has his place"). *Relevance to medicine*: Patients from a society that accepts power distance tend to follow the doctor's orders without question. By contrast, patients from societies with lower power distance may not automatically respect a doctor's opinion and may tend to question the reasons for making a recommendation.
Avoidance of uncertainty: The norms and beliefs regarding how people react to ambiguous situations.	Countries such as Japan, France, or Greece tend to avoid uncertainty; they prefer predictability and, therefore, develop strict hierarchies, laws, and procedures. Deviant ideas are discouraged and consensus is important, so there is typically a strong sense of nationalism. In Nordic or Anglophone countries, people seem more tolerant of uncertainty; they dislike structure and there is less nationalism. But they are viewed as disorganized and confused by people from uncertainty-avoidant countries. *Relevance to medicine*: People from cultures that avoid uncertainty may wish their doctor to provide clear guidance on what they must do. People from cultures that tolerate uncertainty may accept that there are always alternative approaches; a therapy may be tried and if it works, great, but if it does not, another one may.
Masculinity: How polarized are men's and women's roles in society? Are men assertive and women submissive, or are they more equal?	*Relevance to medicine*: Masculine values tend to coincide with a clear distinction between gender roles: often the husband will make decisions for the wife.
Orientation to time: Do people focus on long-term or short-term goals?	Western societies typically have a short-term focus and also view time as a valuable resource. There is an emphasis on focusing on one thing at a time. For other societies (African, Caribbean, and some Indian groups) time urgency is much less central. They may be polychronic (lots of things can happen at once and things can be put off to later: *mañana*). *Relevance to medicine*: Patients from such cultures may not be punctual ("eight o'clock Jamaica time" means any time after around nine or nine-thirty - roughly). In some Arabic cultures, setting deadlines may be viewed as an insult: important things take a long time and cannot be rushed.

(Source: Hofstede G. Cultures and organizations: software of the mind. London: McGraw-Hill; 1991.)

Indigenous people's health in Canada

A prominent Canadian cultural theme concerns the history of our Indigenous peoples, and many historical events have exerted a lasting effect on the health of indigenous peoples. These determinants include the creation of the reserve system, forced relocations, forced placement of children in residential schools, inadequate services for those living on reserves, systemic racism, and a lack of comprehension of the effects of these experiences in the mainstream society.

By imposing Western cultural values and laws, Canada profoundly influenced many determinants of health for First Nations, Inuit and Métis peoples. This colonization has been described as a "process of encroachment and subsequent subjugation of Indigenous peoples since the arrival of the Europeans. From the Indigenous perspective, it refers to the loss of lands, resources, and self-direction and to the severe disturbance of cultural ways and values."[8] Through attempts at cultural assimilation, indigenous groups lost their land, self-government systems, cultures, languages, health care and education systems and traditional economies. This has left Indigenous peoples worse off than other Canadians in social and economic terms, and this has had profound implications for their health. The National Aboriginal Health Organization (NAHO) suggests that there is some preliminary evidence that the transfer of authority over service delivery will lead to better health outcomes for Aboriginal Peoples. A NAHO presentation[9] refers to the adverse impact of colonization on environmental change; of the destruction of traditional economies and of self-sufficiency; of the damage of migration and cultural influences on traditional nutrition patterns; of the removal of traditional forms of care; and of the impact of loss of self-determination on identity and consequently on suicide rates.

A more recent issue has been the lasting impact of residential schooling. Residential schooling was a systematic attempt to assimilate indigenous groups. Many children were forced away from their families and communities into the schools, although some families believing that school would enable their children to live in the white society were not against it. At school, children were forbidden to speak their native languages and they were expected to respect values and norms, which, particularly in relation to child-rearing practices, were very different from those in the children's homes. The experience for the children was, at the very least, distressing. Some children experienced sexual and physical violence. The residential

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

school programme started around 1874, taking over from the mission schools whose aim was to convert Aboriginal people to Christianity. Compulsory attendance began in the 1920s. The last residential school closed in 1996. The residential schools programme did not succeed in its end, which was the destruction of the languages and traditional cultures of First Nations, Inuit and Métis peoples, but they did lead to the erosion of their cultures and thereby their dignity. The legacy of the residential school among Indigenous peoples the loss of identity, of alienation and cynicism towards the rest of society. The spiral of personal health problems that have arisen from this trauma include addictions, abusive relationships (victims who have been abused and have not healed in turn abuse others) and suicide, and are at risk of being transmitted to younger generations. Indigenous communities have been working hard to heal from this trauma and ensure the resilience of their young people and their cultures. A health care workforce that provides culturally safe care is part of that healing.

The cultural trauma of colonization has lasting health effects: ". . . cumulative, emotional and psychological wounding across generations, including the life span, which emanates from massive group trauma."[10] "This trauma has affected Indigenous attitudes, beliefs and behaviours, and forms a crucial social determinant of Indigenous health. The present-day effects of colonization include a lack of cultural understanding between Indigenous and non-Indigenous Canada and strong feelings of distrust between the two."[11] "At the same time, First Nations, Inuit, and Métis peoples have shown great resiliency in dealing with these challenges and have a rich body of knowledge and traditions to share. Traditional knowledge and ways of healing continue to be facilitated through healers, midwives, and traditional medicine persons who constitute a significant Indigenous health provider system."[5] Indigenous medicine contains innumerable herbal remedies, some of which have formed the basis for pharmaceutical treatments that we routinely use in Western medicine. An example the Pacific Yew tree whose bark has been used traditionally for centuries and which was the source of taxol for breast cancer. Indigenous healing approaches are holistic in that they consider mind, body and spirit, as implied in the WHO definition of health. Indigenous *medicine* is distinguished from *healing*, which goes beyond mere treatment of sickness and often involves a spiritual dimension.

The contrasts between traditional Aboriginal approach to healing and Western medicine can make it challenging for practitioners of Western medicine to meet Aboriginal patients' needs; tension and misunderstandings can arise. Accordingly, in many cities medical facilities are being developed that try to integrate Western medicine with Aboriginal teachings; the Anishnawbe Health centre in Toronto is a good example, and is described at http://www.aht.ca/. NAHO offers eight guidelines on practising culturally safe health care for Aboriginal patients. These cover the provision of Aboriginal rooms in the hospital; the need to allow Aboriginal patients access to ceremony, song and prayer; respect for a patient's need for ceremonial items; the need for information and for family support; guidelines for the appropriate disposal of body parts; and guidelines for handling death.[3]

Illustrative Materials

The following are some healing practices associated with First Nations cultures, which are commonly performed by community elders.

Smudging
A smudge is smoke used for ritual cleansing. Smudging is a ceremony traditionally practiced by some Aboriginal cultures to physically or spiritually purify or cleanse negative energy, feelings, or thoughts from a place or a person. Sacred medicines such as cedar, sage, sweet grass, or tobacco are burned in an abalone shell. The person puts their hands in the smoke and carries it to their body, especially to areas that need spiritual healing (mind, heart, body).

Healing circles
Meetings held to heal physical, emotional, and spiritual wounds. A symbolic object, often an eagle feather, may be given to a person who wishes to speak, and then it is passed around the circle in sequence to others who wish to speak. Shamans may conduct the ceremony.

Sweat lodge (a.k.a. purification lodge)
A ceremonial sauna used for healing and cleansing. It is made of a wooden framework covered by blankets or skins, usually dome-shaped, about 1.5 metres high, and large enough for eight people to sit in a circle on the ground. Hot stones are placed in a shallow hole in the centre of the lodge. A medicine man pours water on the stones to produce steam and participants may spend an hour sweating in the lodge. The lodge combines the four elements of fire, water, air, and earth. Ceremonies include offerings, prayers, and reverence. At times, excessive exposure to the heat of the lodge may have negative health effects; environmental toxins can also be released if grasses that have been exposed to pesticides are placed on the rocks.

Sun Dance (a.k.a. Rain Dance, Thirst Dance, Medicine Dance)
A ritual that celebrates the harmony between man and nature, and spiritual dedication. Originally practiced at the summer solstice, the Sun Dance represents continuity between life, death, and regeneration. The symbolism often involved the buffalo, on which Plains Indian groups depended, so deserving reverence, but which they also had to kill for sustenance. Four days before the ceremony, the dancers prepare by purifying themselves, at times in a Sweat Lodge, by meditating and collecting ceremonial items of dress to use in the Sun Dance. The Sun Dance itself takes another four days and generally involves drumming, singing and dancing, but also fasting and, in some cases, self-inflicted pain. This symbolized rebirth and often involved piercing the skin and attaching cords that the person had to tear out. This element led governments to suppress the Sun Dance around 1880, but it has since been re-introduced.

Pipe ceremony
The pipe is used individually and in groups for prayer and ceremonial purposes. Participants gather in a circle. A braid of sweet grass is burned to purify the area and those present to make a sacred place for the spirits to visit. Tobacco or kinnickkinnick, a traditional mixture of bearberry and wild herbs or red willow shavings, is smoked so that prayers can be made to the Great Spirit or requests made of the spirits. The pipe may also be smoked to open other meetings or ceremonies. When not in use, the bowl and stem are separated and carried by one individual, the pipe holder.

Potlatch

A ceremonial feast among northwest Pacific coast Native peoples held to celebrate major family events, such as a marriage or birth. The host distributes gifts according to the status of each guest, reinforcing the perceived hierarchical relations between groups. At times the gift-giving became competitive, the host giving away personal possessions in anticipation that others would reciprocate in their turn. Such largesse enhanced the host's prestige. Missionaries encouraged government to outlaw the Potlatch around 1885, but this ban proved impractical to enforce and was eventually repealed. The ceremony is now commonly practiced.

Self-test questions

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Self-Testing

1. What elements of a patient's culture would you consider when deciding how best to manage the case?
Cultural influences may affect the patient's reaction to the disease, to your suggested therapy, and to your efforts to help them prevent recurrences by changing risk factors. Therefore, it may be important to find out about such possibilities; you can explain that you need them to tell you about their family's and community's feelings about your recommendations. Explain that you are not familiar with their community and want them to tell you if they may have beliefs or obligations that you should be aware of, such as any restrictions on diet, medications, etc., if these could be relevant.
2. Outline at least one difference between cultural competence and cultural safety.
Cultural competence is included within cultural safety, but safety goes beyond competence to advocate actively for the patient's perspective, to protect their right to hold the views they do. When a patient knows that you will honour and uphold their perspective and not try to change it, they will be more likely to accept your recommendations. A physician who practices culturally safe care has reflected on her own cultural biases recognizes them and ensures that her biases do not impact the care that the patient receives. This pattern of self-reflection, education and advocacy is also practiced at the organizational level.
3. How do the effects of colonization continue to impact on the health of Indigenous Peoples in Canada?
By imposing Western cultural values and laws, Canada profoundly influenced all the determinants of health for First Nations, Inuit and Métis peoples. Through attempts at cultural assimilation, they lost their land, self-government systems, cultures, languages, health care and education systems, economies, etc. One cannot have health and wellness when so many of the determinants of health are not met. A more recent issue has been the lasting impact of residential schooling. This formed a systematic attempt to destroy the languages, traditional cultures and thereby the dignity of First Nations, Inuit and Métis peoples. Most Canadians are only now learning about the abuses of this system which began with the forcible removal of children from their parents, families, communities and cultures and sometimes included sexual and physical violence against children. This has left a legacy among indigenous peoples of feelings of a loss of identity, alienation and cynicism towards the rest of society. The spiral of personal health problems that have arisen as a result of this trauma include addictions, abusive relationships (victims who have been abused and have not healed in turn abuse others) and suicide, and are at risk of being transmitted to younger generations. Indigenous communities have been working hard to heal from this trauma and ensure the resilience of their young people and their cultures. Having a health care workforce that can practice culturally safe care is part of that healing.

Further Reading

1. The Association of Faculties of Medicine of Canada (AFMC) has a module on cultural awareness designed for those who teach physicians from other countries, but it contains some useful general material on culture and cultural awareness. http://www.afmc.ca/img/modules_en.htm

2. Various resources outline the importance of cultural safety in Aboriginal health and healing. For example, a paper by Brascoupé and Waters was entitled "Exploring the applicability of the concept of cultural safety to Aboriginal health and community wellness" (Journal of Aboriginal Health, November 2009, pages 6 ? 41). This, and other materials, are available through the National Aboriginal Health Organization (NAHO) web site at http://www.naho.ca/

References

1. Hruschka D, Hadley C. A glossary of culture in epidemiology. J Epidemiol Community Health 2008;62:947-951.

2. Caron N. Caring for Aboriginal patients: the culturally competent physician. Royal College Outlook 2006;3(2):19-23.

3. National Aboriginal Health Organization (NAHO). Cultural competency and safety: a First Nations, Inuit, and Métis Context & guidelines for health professionals. Presentation to the Royal College of Physicians and Surgeons of Canada: Advisory Committee on First Nations, Inuit and Métis Health Education in PGME & CME 2001 April 21.

4. Spence D. Hermeneutic notions illuminate cross cultural nursing experiences. J Adv Nurs 2001;35(4):624-630.

5. Indigenous Physicians Association of Canada, Association of Faculties of Medicine of Canada. First Nations, Inuit, Métis health: core competencies: a curriculum frameword for undergraduate medical education. 2008; Available at: http://www.afmc.ca/pdf/CoreCompetenciesEng.pdf. Accessed April, 2010.

6. Merrill R. Life expectancy among LDS and Non-LDS in Utah. Demographic Research 2004;10(3):61-82.

7. Benson D. Providing health care to human beings trapped in the poverty culture: reconciling the inner self with the business of health care. Physician Executive 2000;26(2):28-32.

8. National Aboriginal Health Organization. Broader Determinants of Health in an Aboriginal Context. Available at: http://www.naho.ca/documents/naho/publications/determinants.pdf. Accessed July 28, 2011.

9. LaRocque E. Violence in Aboriginal communities. 1993; Available at: http://dsp-psd.pwgsc.gc.ca/Collection/H72-21-100-1994E.pdf. Accessed July 28, 2011.

10. Dodgson J. Indigenous women's voices. J Transcult Nurs 2005;16(4):339.

11. Navaneelan T. Determinants of health status of Aboriginal people in Canada. Available at: http://www.medicine.uottawa.ca/sim/data/Aboriginal_Health_Determinants_e.htm. Accessed April, 2010.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Chapter Objectives

After completing this chapter, the reader will be able to:

- Discuss the concept of life course and the **natural history** of disease, particularly with respect to possible public health and clinical interventions.
- Understand the four levels of prevention (primordial, primary, secondary, and tertiary).
- Appreciate the role that physicians can play in promoting health and preventing diseases at the individual and community level (e.g., prevention of low birth weight, immunization, obesity prevention, smoking cessation, cancer screening, etc.).
- Describe strategies for community needs assessments, health education, community engagement, and **health promotion**.
- Describe the five strategies of health promotion as defined in the Ottawa Charter and apply them to relevant situations.
- Describe the advantages and disadvantages of identifying and treating individuals versus implementing population-level approaches to prevention.

Natural history of disease

Chapter 1 introduced the **natural history**of a disease, from exposure to causal agents through to its progression and final outcomes. Understanding the characteristic natural history of a disease enables physicians to anticipate prognosis and to identify opportunities for prevention and control.[1] For instance, based on what is known of Type 2 diabetes, Dr. Rao can be concerned that Catherine Richards may face visual impairment, kidney failure, and possible amputation if she does not take measures to control her disease. Understanding the natural history can also tell us roughly the time frame within which we have to intervene to alter the clinical course of her diabetes and prevent the development of more serious consequences.

Ideally, prevention occurs before people contract a disease, so preventive programs are often delivered to currently healthy people in the general population. To design such a programme we must, therefore, understand the distribution of the condition in the population and know how to identify future cases. The metaphor of the "**iceberg of disease**" suggests that for every case that comes to a clinician, there are likely to be many more people with pre-clinical disease in the community, and even more with risk factors for the condition. Figure 4.1 links the phases in the natural history and **clinical course** of a disease to successive stages of prevention. Different groups, in different locations, will deliver the different preventive strategies.

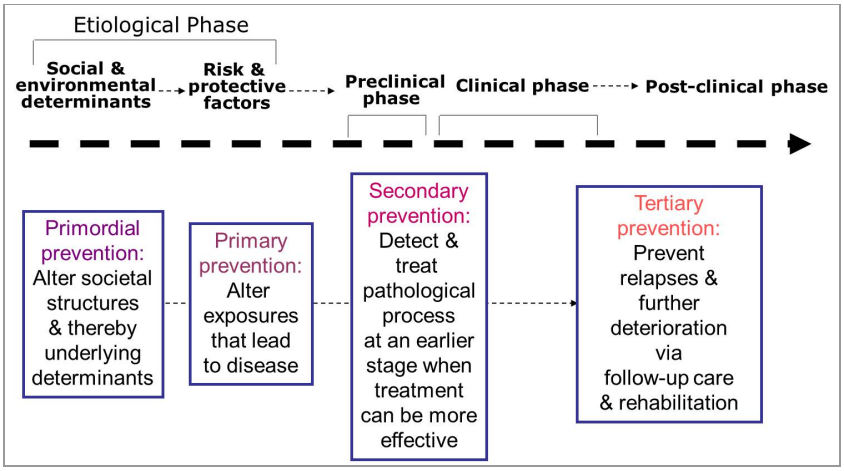


Figure 4.1: Clinical course of a disease: four prevention stages

The stages of prevention

Chapter 1 illustrated how the development of any disease in a patient progresses through a natural history that can, for convenience, be broken into a series of stages. Preventive measures can be applied at any stage along the natural history of a disease, with the goal of preventing further progression of the condition (see **Prevention** in Glossary). For the purposes of introduction it is convenient to think of preventive actions in terms four main stages, but in reality the stages blur one into the next.

Primordial prevention consists of actions to minimize future hazards to health and hence inhibit the establishment factors (environmental, economic, social, behavioural, cultural) known to increase the risk of disease.[2] It addresses broad health determinants rather than preventing personal exposure to risk factors, which is the goal of primary prevention. Thus, outlawing alcohol in certain countries would represent primordial prevention, whereas a campaign against drinking and would be an example of primary prevention.

Examples of primordial include improving sanitation (such that exposure to infectious agents does not occur), establishing healthy communities, promoting a healthy lifestyle in childhood (for example, through prenatal nutrition programs and supporting early childhood development programmes), or developing green energy approaches. Starfield et al. give more examples.[2] So, in preventing Catherine Richards's diabetes, subsidized fitness programmes at the sports centre could have made make such activities more affordable for women like her, and could help to make exercise a norm for women in her community. Similarly, increasing sports programmes in schools may help reduce obesity in the subsequent generations.[3] As these are all population-level programmes, primordial prevention is conceptually linked to population health and health promotion, but clinicians can play a role bringing problems to notice and advocating for action on determinants.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Illustrative Materials

Healthy communities.

During the early 1980s the European regional office of the World Health Organization (WHO) proposed actions to improve the quality of life in cities by making the urban environment conducive to healthy living: providing recreational resources, improved transportation, cleaner environments, more pleasant housing and so on. Toronto was an early participant in the healthy cities movement.[4]

Primary prevention seeks to prevent the onset of specific diseases via risk reduction: by altering behaviours or exposures that can lead to disease, or by enhancing resistance to the effects of exposure to a disease agent. Examples include smoking cessation and vaccination. Primary prevention reduces the incidence of disease by addressing disease risk factors or by enhancing resistance. Some approaches involve active participation, as with regular tooth brushing and flossing to prevent dental caries. Other approaches are passive: adding fluoride to the municipal drinking water to harden tooth enamel and prevent caries. Primary prevention generally targets specific causes and risk factors for specific diseases, but may also aim to promote healthy behaviours, improve host resistance, and foster safe environments that reduce the risk of disease, for instance, thorough cleaning of operating rooms to prevent post-operative infection. Preventive efforts can be fitted into the agent-host-environment model of causation introduced in Chapter 2.

Secondary prevention includes procedures that detect and treat pre-clinical pathological changes and thereby control disease progression. Screening procedures (such as mammography to detect early stage breast cancer) are often the first step, leading to early interventions that are more cost effective than intervening once symptoms appear. Routine blood sugar testing for people over 40 would be an example relevant to detecting Catherine's diabetes early. Screening is usually undertaken by health professionals, either at the level of individual doctor-patient encounters (e.g., routine blood pressure checks) or via public health screening programs (e.g., mammography screening). The criteria for implementing a screening programme are described in Part 3.

Once a disease has developed and has been treated in its acute clinical phase, tertiary prevention seeks to soften the impact caused by the disease on the patient's function, longevity, and quality of life. Examples include cardiac rehabilitation following a myocardial infarction, seeking to alter behaviours to reduce the likelihood of a reinfaction. Tertiary prevention can include modifying risk factors, such as assisting a cardiac patient to lose weight, or making environmental modifications to reduce an asthmatic patient's exposure to allergens. In the example of Catherine Richards, it might include ensuring regular check-ups to monitor her condition, including eye exams to check for possible adverse outcomes of her diabetes. Where the condition is not reversible, tertiary prevention focuses on rehabilitation, assisting the patient to accommodate to his disability. For reversible conditions, such as many types of heart disease, tertiary prevention will reduce the population prevalence, whereas for incurable conditions it may increase prevalence if it prolongs survival. The key goal for tertiary prevention is to enhance quality of life.

Table 4.1illustrates the primary, secondary and tertiary levels of prevention.

Table 4.1: Examples of primary, secondary, and tertiary prevention interventions targeting individuals and populations

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Disease	Intervention level	Primary	Secondary	Tertiary
Colorectal cancer	Individual	Counselling on healthy lifestyles: dietary counselling for people at risk of colorectal cancer, etc.	Hemoccult stool testing to detect colorectal cancer early	Follow-up exams to identify recurrence or metastatic disease: physical examination, liver enzyme tests, chest x-rays, etc.
	Population	Publicity campaigns alerting the public to the benefits of lifestyle changes in preventing colorectal cancers; promotion of high fibre diets; subsidies to help people access exercise programmes; anti-smoking campaigns	Organized colonoscopy screening programs	Implementation of health services organizational models that improve access to high-quality care
Infectious diseases: hepatitis C	Individual	Counselling on safe drug use to prevent hepatitis C virus (HCV) transmission; counselling on safer sex	Screening for HCV infection of patients with a history of injection drug use	HCV therapy to cure infection and prevent transmission
	Population	HCV prevention includes safer sex practices, programmes to discourage needle sharing among intravenous drug users, etc.	Establish a universal testing system for HCV in high risk groups	(Similar to primary prevention): ensuring close control of high risk sites such as tattoo parlours that have been associated with outbreaks
Metabolic syndrome	Individual	Nutrition and exercise counselling	Screening for diabetes	Referral to cardiac rehabilitation clinics
	Population	Built environment favourable for active transport (walking, bicycling rather than using a car)	Community level weight loss and exercise programs to control metabolic syndrome	Implementation of multidisciplinary clinics

Here Be Dragons

Beware: you may see the term secondary prevention mistakenly used in reference (for example) to preventing a second stroke in a patient who has had a first one. This is incorrect, for it would be an example of tertiary prevention, in terms of preventing further damage to the patient's vascular system.

Illustrative Materials

Bodies in the river

A passer-by sees a body floating down a river and calls 911. Firemen arrive and haul the person out. Paramedics start resuscitation and rush the victim to the ER, where the hospital public relations office announces that while everything possible was done to save the victim, sadly he died. Suicide is presumed.

Some time later a similar event happens again. Sensing an ideal political opportunity, the local member of parliament lobbies the Ministry of Health to have an ambulance station built nearby, complete with advanced rescue equipment and special resuscitation training "so that this tragedy need never happen again in our community." A journalist at the press conference pointedly asked why nets or fences could not be installed on the bridge located upstream as this would be more effective and cheaper than maintaining services downstream; he was promptly escorted away from the meeting by the politician's aides.

Applying the metaphor of ?upstream' and ?downstream' interventions that was introduced in Chapter 2 to the case of the bodies in the river, tertiary prevention would imply downstream efforts at resuscitation at the scene and in hospital. But of course these may be judged as being "too little, too late." Secondary prevention activities might focus (metaphorically) further upstream and include attempts to identify suicidal people before they make an attempt, perhaps via screening for depression in primary care practices using a brief screen for depressive symptoms. Those who give indications of depression could be more fully evaluated, perhaps by referral to a psychologist. The success of secondary prevention would depend on many factors but the current evidence suggests that it is not effective in preventing suicide. Therefore, primary or primordial prevention may be more effective in this instance. Primary prevention might include social programs for high-risk youth in areas of high suicide rates, or putting up safety nets to prevent the act of suicide. Primordial prevention falls in the domain of population health approaches; these involve a wide range of government agencies and focus on developing healthy public policies and altering underlying determinants of health. Applied to the bodies in the river, primordial prevention would likely focus on the social ecology of suicide, seeking to identify underlying determinants that explain why people in

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

that area are throwing themselves into the river. Is there, for example, a connection between the world economic downturn, local unemployment, debt and feelings of despair? Do such problems occur more commonly in small, one-industry towns in Canada?

Nerd's Corner

More on the stages of prevention

Thinking of prevention as a series of stages offers a useful introduction to the topic, but classifying a given action into a stage may not be simple and depends greatly on the context. On closer examination, you can only take the idea of stages of prevention so far.

For example, checking (and then controlling) blood pressure may represent primary prevention if the condition you aim to prevent is a heart attack. But it may be seen as secondary prevention if the patient has a family history of hypertension and you are screening for this?it may even be tertiary prevention if the patient has already had a heart attack and the goal is to prevent a recurrence. Bother!

Secondary prevention has been defined as slowing the progression of a disease or its sequelae at any point after its inception. From this perspective, treating a disease can be seen as prevention if it slows the progression of the disease or prevents sequelae. In an interview aired on CBC on 19 July 2010, Michel Kazatchkine, Executive Director of the Global Fund to fight AIDS, TB and Malaria, noted that treatment of HIV/AIDS can qualify as primary prevention because it reduces the risk of transmission of the virus.

Conversely, excess use of antibiotics may be viewed as the opposite of prevention if it contributes to the development of resistant organisms. Clinicians should always weigh the costs and benefits of treating against those of not treating in much the same way as they do in prevention. Reflecting this theme, you may also hear the term ?quaternary prevention': where a patient is at risk of over-medication, quaternary prevention protects them from new medical invasion, suggesting interventions which are ethically acceptable, and assuring quality of the care process.[5]

Implementing prevention, health protection and health promotion

Preventive programs are delivered by a range of public health agencies, or individual clinicians (frequently by general practitioners, nurses and nurse practitioners), or through community agencies (including volunteer groups such as students against drunk driving, or non-profit groups). The ideal programme involves the coordinated work of many agencies; we have many examples of good programs in Canada, but they tend to be coordinated in relative isolation by the single agencies responsible for them, resulting in gaps and overlaps between programs. A challenge is that programmes need to be tailored to the local situation: one size does not fit all. Demographics and the local situation have to be considered; this theme is discussed in programme planning in Chapter 12.

Health protection

Health protection refers to a wide range of activities undertaken by public health departments and also by some government agencies, such as the Public Health Agency of Canada (PHAC). Health protection spans primordial and primary prevention, such as "ensuring safe food and water supplies, providing advice to national food and drug safety regulators, protecting people from environmental threats, and having a regulatory framework for controlling infectious diseases in place. Ensuring proper food handling in restaurants and establishing smoke-free bylaws are examples of health protection measures.[6]" Public health protection also deals with reducing environmental threats to the health of the population, such as biological, chemical, or physical agents that may cause an epidemic if not controlled. Public health agencies are often required by law to deal with identified threats, which can be detected via surveillance systems (see **Surveillance** in Glossary), as described in Chapter 7. The health protection approach will vary according to the type of biological, chemical, or physical risk involved. For example, federal, provincial and local inspectors will handle the biological risks of specific communicable diseases such as those associated with contamination of water or food. Legislation and regulations address chemical hazards, such as risks due to environmental tobacco smoke, and fines are applied. A traditional public health protection or risk reduction approach to the problem of the bodies in the river might involve adding protective netting to the bridge. Even though nets might not deter a determined suicide attempt, it could reduce impulse suicides and it would help to make the problem visible by reminding passers-by that suicides have occurred in that spot. This recognition may help indirectly, by raising vigilance in the public and encouraging people to take early warning signs seriously.

Health promotion

While health protection focuses on removing negative influences on health, **health promotion** aims to enhance health in terms of developing healthy public policy, healthy environments and personal resiliency; this reflects a philosophy of supporting communities and individuals to take charge of their own health. Beginning in the 1970s and 1980s, health promotion built on health education, which was considered insufficient in itself. The aims of health promotion include, but also go beyond, preventing disease. "Health promotion includes strengthening the skills of individuals to encourage healthy behaviours, and it also includes building the healthy social and physical environments to support these behaviours.[7]" It involves "any combination of health education and related organizational, economic, and political interventions designed to facilitate behavioural and environmental changes conducive to health.[8]"

A health promotion programme might include a specific intervention such as smoking cessation, but this would be presented as part of a broader set of interrelated interventions including environmental and lifestyle changes that support non-smoking. Enhancing supportive environments and encouraging healthy behaviours using a health promotion approach contributes to primary prevention of disease, but is also intended to have a broader beneficial effect in terms of encouraging people to take responsibility for their health. This idea addresses the concerns over the erosion in individual responsibility for health that may result from growing reliance on the availability of therapy once disease develops, which was discussed in Chapter 1 (see Nerd's Corner box). Applied to the bodies in the river, a health promotion approach might begin by asking what was causing people to throw themselves into the water, and then attempt to correct this cause. It would typically focus on programs aimed at helping people to cope with

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

stress in their lives: arranging mutual-help groups and working to improve living conditions, and so forth.

Recognizing that health behaviours are unlikely to change in a lasting manner unless environmental factors that give rise to them are also changed, health promotion takes an ecological approach and focuses on community engagement and environmental change in addition to individual behaviours. An ecological approach often begins with a community needs assessment. Reflecting the models of behaviour changed described in Chapter 2, people rarely change a particular behaviour unless they consider it a priority. Likewise a community will have a more or less well-voiced set of priorities that may support or inhibit efforts to promote health. An assessment of perceived community needs is therefore often the first stage in a local health promotion campaign: what does this community perceive as its priorities and how do these match the goals of the health promotion team? A community needs assessment will collect data, typically from a combination of surveys and interviews with community opinion leaders, on the health and social problems of individuals, families and the community as a whole. Community Health Centres commonly apply a needs-based approach to planning health promotion programmes; these Centres typically have community representation on their board of directors and plan programmes based on community input. Engaging community members in the planning process not only ensures that the programme is likely to be relevant to local needs, but also helps to promote community support for the programme and participation in it. The actual design of the programme may be developed following a case conference of experts in relevant specialties who suggest approaches that are designed to match the particular local situation.

The principles on which health promotion strategies are designed were described in the Ottawa Charter for Health Promotion.

Nerd's Corner

The philosophy of health promotion

Health promotion reflects a characteristic set of liberal values, such as self-responsibility for health, that characterize much of the WHO thinking: "Health promotion is the process of enabling people to increase control over, and to improve, their health." This definition derived from the WHO conception of health (see **Health** in Glossary) described in Chapter 1: "The extent to which an individual or group is able to realize aspirations and to satisfy needs, and to change or cope with the environment.[8]" Along with self-responsibility, other core values of health promotion include many "upstream" factors, such as:

- promoting equity and social justice;
- applying a holistic definition of health;
- consideration of the full range of health determinants;
- recognizing environmental influences on health;
- empowering people and building individual and collective capacity;
- seeking to enhance people's social participation; and
- fostering collaboration between agencies.

The Ottawa Charter for Health Promotion

Chapter 1 described the evolution of thinking about prevention when diseases of lifestyle became common in the 1950s and 1960s. Health education was initially seen as a key approach to modifying health behaviours, but subsequent recognition that education alone was insufficient broadened the field to health promotion. In 1986, Ottawa was the venue for an international conference sponsored by the WHO to establish the basic design principles for health promotion programmes.[8] The resulting charter proposed a plan of action to achieve health for all by the year 2000. The charter sets out a range of upstream and downstream approaches outlined in the box below.

Illustrative Materials

The **Ottawa Charter for Health Promotion**

The Ottawa Charter and the Health for All 2000 manifesto included the following strategies:

- *Building healthy public policy*. The aim is to put health on the agenda for all policymakers, and to ensure that they consider the health implications of their decisions. Healthy public policy is policy that does not have the side effect of damaging health while seeking some other goal.
- *Creating supportive environments*. The emphasis on environment reflects an awareness of the impact of natural, built, and social environments on health, and proposes a socio-ecological approach to health.
- *Strengthening community action*. Health promotion requires community empowerment and involvement in setting priorities, planning and implementing strategies to achieve better health.
- *Developing personal skills*. Health promotion supports personal and social development through providing information and enhancing life skills.
- *Re-orienting health services*. Health promotion argues for shifting health resources towards a more equal distribution between treating disease and preventing it. Essentially, health services should be expanded to include the 4 strategies above in addition to conventional medical care. Responsibility for health promotion services should be shared among individuals, community groups, health professionals, health services, and governments.

The Charter also identified seven prerequisites for health: peace, shelter, education, food, income, a stable eco-system, and sustainable resources. These prerequisites are closely related to the macro-social determinants of health and are essential in understanding why we have failed to reach the goal of health for all by the year 2000 even though it is 10 years past the deadline.[9,10] Discussion points: to what extent has the world failed in this goal, and why? How can a physician influence these factors in his or her practice?

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Illustrative Materials

An application of the Ottawa Charter

A public health programme in Glasgow, Scotland, illustrated an application of the Ottawa Charter in an experiment to improve the dental health of 5-year-old children living in deprived neighbourhoods.[11] The oral health of Glasgow children had been documented as being among the poorest in Western Europe.

The ?from birth' caries prevention programme addressed early lifestyle determinants of dental caries. Oral health action teams in each of the city's 15 health care administrative areas led the interventions.

Here are some examples of the activities they undertook:

Building healthy public policy: staff education in all nursery schools; implemented healthy snacks policies in nursery schools; subsidised utensils and food blenders; provided free fluoride toothpaste

Supportive environments: community oral health promotion events organized; changes implemented in nursery schools; held community consultations; arranged ? get cooking' classes;

Developing personal skills: literature was translated into jargon-free language; training in toothbrushing skills; dental health song books

Strengthening community action: there was a strong emphasis on community engagement. This included creating networks of voluntary community activists to give outreach into the communities, leading community groups in identifying ways to promote caries-protective events and behaviours; they trained the trainers

Reorienting health services: the oral health action teams promoted perinatal oral health sessions in doctors' offices; created dental registration programmes.

The programme included an evaluative component that will be summarised in Chapter 7.

Self-test questions

Self-Testing

1. Distinguish between the natural history and the clinical course of a disease.
Natural history refers to the sequence of changes and progression in a disease that is untreated. Treatment may modify this, producing the clinical course. If treatment is ineffective (as in using antibiotics for a viral illness) the two may be the same.
2. The Pap smear test detects cervical cancer in women at an early stage of the disease when there are no symptoms and the disease is not evident on visual examination. Which of the following statements applies to a woman with no history of cervical cancer who undergoes a Pap smear?
A. She is practicing primordial prevention.
B. She is practicing primary prevention.
C. She is practicing secondary prevention.
D. She is practicing tertiary prevention.
E. This is a screening test, not a preventive procedure.
C) Secondary Prevention
Rationale:
A. Primordial prevention acts early in the causal chain, to alter general social or economic circumstances that give rise to risk factors. The Pap test does not address general circumstances, but a specific disease process.
B. Primary prevention is defined as the avoidance of disease ? having a Pap smear would not prevent the disease from occurring.
C. Secondary prevention is defined as the interruption of any disease process before the emergence of recognized symptoms or diagnostic findings of the disorder. The Pap smear test forms an essential step in this process: it identifies the disease process before the emergence of symptoms.
D. Tertiary prevention is defined as the avoidance of negative sequelae of a disease process, once the disease has been diagnosed and treated. The Pap smear concerns early detection; it has nothing to do with subsequent effects of the disease.
E. Yes, it is a screening test, but that forms an integral component of an approach to prevention. So this response is splitting definitional hairs and presumably the physician's intent in administering the test was to prevent further progression of the cancer.
3. Contrast the underlying philosophies of health promotion and health protection.
Health promotion seeks to foster the capacity for self-responsibility in a community, enabling it to improve its health via collective action. Health protection is set of actions, often supported by legislation, applied by an external authority to avoid adverse health consequences. Involvement of the community may be limited.
4. Summarize the elements in the Ottawa Charter for Health Promotion
Building healthy public policy; Creating supportive environments; Strengthening community action; Developing personal skills, and Re-orienting health services. Now give examples of how these might be put into practice in a typical family medicine centre in the city where you live.

References

1. Bhopal R. Concepts of epidemiology. Oxford: Oxford University Press; 2002.

2. Porta M. Dictionary of epidemiology. New York: Oxford University Press; 2008.

3. Starfield B, et al. The concept of prevention: a good idea gone astray? J Epidemiol Community Health 2008;62:580-583.

4. Beaglehole R, Bonita R, Kjellström T. Basic epidemiology. Geneva, Switzerland: WHO; 1993.

5. Flynn B. Healthy cities: toward worldwide health promotion. Annu Rev Public Health 1996;17:299-309.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

6. Public Health Agency of Canada. Health goals for Canada. 2009; Available at: http://www.phac-aspc.gc.ca/hgc-osc/new-1-eng.html. Accessed Nov 21, 2009.

7. Green L. National policy on the promotion of health. Int J Health Educ 1979;22:161-168.

8. World Health Organization. Ottawa charter for health promotion. 1986; Available at: http://www.who.int/hpr/NPH/docs/ottawa_charter_hp.pdf. Accessed July 28, 2011.

9. Bryant J, Zuberi R, Thaver I. Alma Ata and health for all by the year 2000. The roles of academic institutions. Infec Dis Clin North Am 1991;5(2):403-416.

10. Gunning-Schepers L. "Health for all by the year 2000": a mere slogan or a workable formula? Health Policy 1986;6(3):227-237.

11. Blair Y, Macpherson L, McCall D, McMahon A. Dental health of 5-year olds following community-based oral health promotion in Glasgow, UK. Int J Paediatric Dentistry 2006;16:388-398.

Chapter Objectives

After completing this chapter, the reader will be able to:

- Know how to access and collect health information to describe the health of a population:
  1. Describe the types of data and common components (both qualitative and quantitative) used in creating a community health needs assessment.
  2. Critically evaluate possible sources of data to describe the health of a population, including the importance of accurate coding and recording of health information.
  3. Describe the uncertainty associated with capturing data on the number of events and populations at risk.
- Analyze population health data using appropriate measures:
  a. Apply the principles of epidemiology in analyzing common office and community health situations.
  b. Discuss different measures of association, including relative risk, odds ratios, attributable risk, and correlations (see **Correlation** in Glossary).
  c. Understand the appropriate use of different graphical presentations of data.
  d. Describe criteria for assessing causation.
  e. Demonstrate an ability to critically appraise and incorporate research findings, with particular reference to the following elements:
     i. Characteristics of study designs (RCT, cohort, case-control, cross sectional).
     ii. Measurement issues (bias, confounding).
     iii. Measures of health and disease (measures of central tendency) and sampling.

Case Study

Julie Richards is worried about her menopause. She gets hot flashes and feels generally tired. She feels it is a time in life that brings many changes, including the risk of osteoporosis and cancers. She mentioned this to her daughter Audrey, who searched the Internet and found lots of information about hormone therapy, calcium supplements and products such as evening primrose oil. There was also information about physical exercise as a way of improving well-being. Julie Richards consults Dr. Rao on the matter and shows him the information Audrey found.

Dr. Rao looks at the texts and wonders how to judge their claims. In particular, Julie wants to know if a magnet will help her symptoms of menopause. She read about it on the web and shows the printout to Dr. Rao. The website gives quite a bit of information about menopause and cites some peer-reviewed articles that suggest that static magnets are effective in the treatment of dysmenorrhoea.

Dr. Rao uses Medline and other Internet sources to check this out. He finds that the author founded and now runs a private clinic specializing in menopause problems. Through Medline, Dr Rao finds a number of articles on magnets in pain management. There is a systematic review of the evidence that finds that magnets might be minimally effective in osteoarthritic pain, but of no demonstrated value in other types of pain. Promoters of the magnets say that their mechanism of action is either direct interference with nerve conduction, or action on small vessels to increase blood flow.

Assessing information and evidence

People have claimed to be able to cure ills for as long as history has been recorded. Some cures are based on science, in that their mode of action and the principles underlying them are known. Some are known empirically to improve health, but we do not fully understand how they work. Other treatments have not been shown to be of benefit. Many have never been rigorously tested. Finally, some have been shown to have only a placebo effect, a benefit achieved by suggestion rather than by direct chemical action.

Every year, approximately 2,000,000 articles summarize the results of scientific research. Many of the conclusions disagree with one another. There are many different reasons for these disagreements, ranging from biases in the study design, to the perspective of the investigator, to unique characteristics of the study subjects, or to methods used in the analysis of study results. No study is perfect, and yet the ideal of practising evidence-

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

based medicine demands that clinicians base their decisions on scientific evidence. Because some ?evidence' is flawed, the clinician must be able to judge the validity of information published in the medical literature, and this forms the theme of critical appraisal of the literature. To guide medical practice, various agencies routinely review publications and develop clinical guidelines based on the current evidence, but (like Dr. Rao discussing the magnets and menopause with Julie) each clinician has to understand the basics of how to critically appraise medical research.

Critical appraisal

The core of evidence-based medicine is to judge the quality of evidence via critical appraisal: the process of applying rules of science and common sense in judging the quality of written articles. Critical appraisal refers to judging the validity of the methods and procedures used to collect data, identifying possible biases that may have arisen, assessing the adequacy of the analysis and completeness of reporting, the conclusions drawn, and reviewing the study's compliance with ethical standards of research. There are a number of checklists that help guide the critical appraisal process, but ultimately, clinicians must use their judgement to assess the study quality and the usefulness of its results to their particular clinical situation.

A first approach to critical appraisal begins with common sense. Here is an approach from 1990 to critically evaluating data quality under the acronym FiLCHeRS, which stands for Falsifiability, Logic, Comprehensiveness, Honesty, Replicability, and Sufficiency.[1]

Table 5.1: Standards for evaluating information quality under the acronym FiLCHeRS

Falsifiability	It must be possible to conceive of evidence that would prove the claim false (for example, there is no logical way of proving that God does not exist).
Logic	Arguments offered as evidence in support of a claim must be logically coherent.
Comprehensiveness	The evidence offered in support of any claim must be exhaustive?all of the available evidence must be considered and one cannot just ignore evidence to the contrary.
Honesty	The evidence offered in support of any claim must be evaluated with an open mind and without self-deception.
Replicability	It must be possible for subsequent experiments or trials to obtain similar results.
Sufficiency	The evidence offered in support of any claim must be adequate to establish the truth of that claim, with these stipulations: the burden of proof for any claim rests on the claimant, extraordinary claims demand extraordinary evidence, and evidence based on authority and/or testimony is always inadequate for any paranormal claim.

Research evidence that is critically appraised as being of good quality is cumulated to form the basis for evidence-based medicine.

Evidence-based medicine

Evidence-based medicineis defined as "the consistent use of current best evidence derived from published clinical and epidemiologic research in management of patients, with attention to the balance of risks and benefits of diagnostic tests and alternative treatment regimens, taking account of each patient's unique circumstances, including baseline risk, co-morbid conditions and personal preferences".[2]

In clinical practice, evidence-based medicine means integrating clinical experience with the best available external clinical evidence from systematic research. The approach was primarily developed in Canada by Dr. David Sackett and others at McMaster University during the 1970s, and is now recognized as a key foundation of medical practice.[3] Sackett described evidence-based medicine as the process of finding relevant information in medical literature to address a specific clinical problem, the application of simple rules of science and common sense to determine the validity of the information, and the application of the information to the clinical question. The aim was to ensure that patient care is based on evidence derived from the best available studies. Sackett argued that the "art of medicine" lies in taking the results of several sources of evidence and interpreting them for the benefit of individual patients: the opposite of what he called "cookbook medicine." The approach has subsequently been applied beyond clinical medicine to propose, for example, evidence-based public health and evidence-based policy-making.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Mnemonic: The 5 As of evidence-based medicine

Here is a sequence that a clinician may follow in applying evidence-based medicine in deciding how to handle a challenging clinical case:

- *Assess:* Recognize and prioritize problems.
- *Ask:* Construct clinical questions that facilitate efficient searching for evidence.
- *Acquire:* Gather evidence from quality sources.
- *Appraise:* Evaluate evidence for its validity, importance, and usefulness.
- *Apply:* Apply to individuals taking account of patient's preferences and values.

For more information on the 5 As, please visit: http://www.jamaevidence.com/ and http://www.ebm.med.ualberta.ca/

Appraising scientific evidence: qualitative versus quantitative research

Medical knowledge is derived from a combination of qualitative and quantitative research. Qualitative research refers to the use of non-numerical observations to answer "Why?" questions, while **quantitative methods** use data that can be counted or converted into numerical form to address "How?" questions. As summarized in Table 5.2, each approach serves a different purpose, so most researchers view the two as complementary and accept a "mixed methods" approach.

Here Be Dragons

Qualitative variables versus qualitative studies

Quantitative studies often examine qualitative variables. For example, a researcher who wants to know about satisfaction with services might ask a question such as: "How satisfied were you with the care you received?" The researcher might use an answer scale that allows the response to be crudely quantified, perhaps through a series of statements: very satisfied, satisfied, unsatisfied, or very unsatisfied. These could be scored 1, 2, 3 or 4 and the researcher could report the mode, or the median score. The study, although measuring the quality of something, expresses the results as numbers and is, therefore, a quantitative study.

Meanwhile, a qualitative study on this topic might involve a focus group of patients where a group facilitator asks focus group participants the same question but allows the participants to talk about what they consider important to their satisfaction and then asks follow-up questions to explore participants' ideas in depth. The information produced is then reviewed and analyzed to identify common themes and sub-themes in the focus group discussion.

Table 5.2: Comparison of qualitative and quantitative research methods

Qualitative research	Quantitative research
Generates hypotheses	Tests hypotheses
Is generally inductive (works from the particular instance to the general conclusion)	Is generally deductive (works from the general theory to the particular explanation)
Focuses on studying the range of ideas; sampling approach provides representative coverage of ideas or concepts	Focuses on studying the range of people; sampling provides representative coverage of people in the population
Answers "why?" and "what does it mean?" questions	Answers "what?", "how much?" and "how many?" questions
Captures rich, contextual, and detailed information from a small number of participants	Provides numeric estimates of frequency, severity, and associations from a large number of participants
Example of a study question: What is the experience of being treated for breast cancer?	Example of a study question: Does treatment for breast cancer reduce mortality and improve quality of life?

Links

Qualitative versus quantitative

For a fuller comparison of qualitative and quantitative designs, see http://www.pngted.uconn.edu/siegle/research/Qualitative/qualquan.htm

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Here Be Dragons

Numbers may not count

In quantitative studies, numbers can be used to categorize responses to qualitative questions, such as "How satisfied were you? Answer 1= very unsatisfied to 4 = very satisfied". Beware: these numbers are arbitrary, and we cannot claim that they represent an even gradient of satisfaction. In technical jargon, these are ?ordinal' numbers (equivalent to house numbers along a street) (see Scales of measurement in Glossary) but the change in satisfaction between each number is not necessarily equal. Accordingly, such data have to be analysed using nonparametric statistical methods ? for example, using a median rather than a mean.

By contrast, measuring body temperature in degrees Celsius forms an ?interval' measure, in which the amount of change in temperature is equal across successive numbers on the scale. Data from such measurements can be analysed using parametric statistics: mean values can legitimately be calculated.

Qualitative research

Qualitative research is defined as any type of research "that employs non-numeric information to explore individual or group characteristics, producing findings not arrived at by statistical procedures or other quantitative means. Examples of the types of qualitative research include clinical case studies, narrative studies of behaviour, ethnography, and organizational or social studies."[4] Qualitative research is useful for understanding an area where little is known, or where previous work seems to be inadequate or not applicable to a particular situation. Qualitative research is explanatory: it seeks to make sense of complex situations where reducing the data to quantifiable elements would cause the loss of too much information. Hence its sampling methods focus on sampling content, rather than persons. It shows how particular situations are experienced and interpreted. Qualitative methods are useful in constructing a theory or a theoretical framework from observable reality and in understanding phenomena deeply and in detail. The precise research question dictates the qualitative data collection and analytic methods to be used. There are many methods, but the three main approaches and their key features are described in the Nerd's Corner box.

Nerd's Corner

Table 5.3: Types of qualitative study

Qualitative method	Type of question	Data source	Analytic technique
Phenomenology		Primary: audiotaped, in-depth conversation. Secondary: poetry, art, films.	Theming and phenomenological reflection; memoing and reflective writing.
Ethnography	Observational questions (e.g., How does the surgical team work together in the OR?) and descriptive questions about values, beliefs and practices of a cultural group (e.g., What is going on here?)	Primary: participant observation; field notes; structured or unstructured interviews. Secondary: documents, focus groups.	Thick description, re-reading notes and coding by topic; storying (narrating the story); case analysis; diagramming to show patterns and processes.
Grounded theory	Process questions about how the experience has changed over time or about its stages and phases (e.g., How do medical residents learn to cope with fatigue?) or understanding questions (e.g., What are the dimensions of these experiences?)	Primary: audiotaped interviews; observations. Secondary: personal experience.	Theoretical sensitivity; developing concepts for theory generation. Focussed memoing; diagramming; emphasis on search for core concepts and processes.

Source: Adapted from Richards et al.[5]

Judging the quality of qualitative research

In judging qualitative research you should look for answers to the following questions:

1. Was the design phase of the project rigorous?

Elements to consider are the skill and knowledge of the researcher and the completeness of the literature review. The research question should also be clear and should be appropriate to qualitative analysis. The researcher should state the perspective from which the data will be gathered and analyzed.

2. Was the execution of the study rigorous?

The final sampling should represent all relevant groups. For instance, a study of patient satisfaction should cover all types of patients that attend the clinic and sample from both sexes, the full age range, and the full range of complaints. In qualitative research, sample size is not necessarily fixed.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Sampling may continue until no new ideas or concepts emerge, a situation known as *saturation*. Nor is the interview script fixed. Questions need not be uniform, but should capture participants' verbal and non-verbal cues so that the topic can be fully explored. As the project continues, the interview script may evolve in response to the findings of previous interviews.

Nonetheless, data collection methods, while flexible, should be systematic and organized, and must be documented. Having more than one researcher analyze the same data is a way to identify possible biases in interpretation; how differences in interpreting the results were reconciled should be noted. In some studies, study participants are asked to validate the interpretation. The reader should look for evidence that the research was conducted in an ethical manner. In particular, confidentiality and anonymity pose challenges in qualitative research, as the data are built on highly personal reports from small samples.

Bias is inherent in qualitative research. Observational data collection, whether or not those being observed can see the observer, can influence their behaviour. The interaction between interviewers and interviewees can also influence responses. Similarly, the results of data analysis can depend on the knowledge and perspective of the person doing the analysis. These are the same kind of problems that are seen in quantitative research, (See the section on bias) but the methods to counteract them cannot be the same. Quantitative research aims for uniformity and standardization to reduce bias. Qualitative research, by its nature, responds to context, it should also explain the context or perspective of the researcher so the reader can assess the researcher's influence on the findings. Two or more researchers often carry out analysis of qualitative research so that personal perspectives can be recognised in order to reduce their effects on the interpretation of results. The knowledge and perspective of each analyst should also be stated for the benefit of the reader.

3. Can I transfer the results of this study to my own setting?

Readers must decide if the context and subjects of the study are sufficiently like their own context and patients for the results to be applicable. The results can also be compared to the published literature: How closely does it corroborate other work? If it corroborates closely, it is likely to be generalizable and, therefore, transferable to a similar context.

Illustrative Materials

How qualitative and quantitative approaches complement each other

Cockburn examined patient satisfaction with breast screening services in Australia. She used qualitative methods to develop measures of satisfaction. She reviewed the literature and interviewed patients and staff about their experience of the services. From this she developed a standardized questionnaire to measure aspects of screening services that reflect previous studies and that are also relevant to patients' experiences of the service. She analyzed data from this questionnaire in a quantitative manner.[6]

Quantitative research

The Western scientific paradigm has been refined over the centuries to understand the general principles underlying observable phenomena; qualitative research focuses on describing specific instances of phenomena. As the name suggests, quantitative research is based on a foundation of counting and mathematical analysis and uses rigorous study designs, both of which seek to minimize the role of human judgment in collecting and interpreting evidence. In general, researchers use quantitative methods to look for statistical associations between variables, or for differences in patterns of health between one population and another. A fundamental motive for many (although not all) such studies is to identify causal factors that influence health status. Examples include studying patterns of risk factors between people with diabetes and those without, patterns of recovery in patients who get a treatment and those who don't, or lifestyle patterns in different sectors of the population. Clinical trials that test the efficacy of a new therapy in curing a disease are also studies of causation: did the therapy cause the patient to recover? Because causal interpretation underlies so much of medical research, a brief description of the logic of inferring causes will be given before study designs to identify causes are described.

Criteria for inferring causation

Epidemiological studies and reviews can demonstrate associations between variables, but an association is not necessarily causal. Sadly, there is no sufficient way to prove for certain that an association between a factor and a disease is a causal relationship. In 1965, Austin Bradford Hill proposed a set of criteria for assessing the causal nature of epidemiological relationships; he based these in part on Koch's postulates. The criteria have been revised many times, so you may find different versions with different numbers of criteria.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

Koch's postulates

Robert Koch (1843 ? 1910) was a Prussian physician and, in 1905, the Nobel Prizewinner in Physiology for his work on tuberculosis. Considered one of the fathers of microbiology, he isolated *Bacillus anthracis*, *Mycobacterium tuberculosis* (once known as Koch's bacillus) and *Vibrio cholerae*. His criteria (or postulates) to establish a microbe as the cause of a disease were that the microbe must be

- Found in all cases of the disease examined
- Prepared and maintained in pure culture
- Capable of producing the original infection, even after several generation in culture
- Retrievable from an inoculated animal and cultured again.

These postulates built upon criteria for causality formulated by the philosopher John Stuart Mill in 1843. Microbiology nerds will be able to cite diseases caused by organisms that do not fully meet all the criteria, but nonetheless, Koch's postulates provided a rational basis for the study of medical microbiology.

Table 5.4: Criteria for inferring a causal relationship

Criteria	Comments
1. Chronological relationship: Exposure to the presumed cause must predate the onset of the disease.	This is widely accepted. But beware of the difficulty in knowing when some diseases actually began if they have long latent periods. For instance, it can be difficult to be sure when exactly a cancer began.
2. Strength of association: If all those with the disease were exposed to the presumed causal agent, but very few in the healthy comparison group were exposed, the association is a strong one. In quantitative terms, the larger the **relative risk**, the more likely the association is causal.	This criterion can be disputed: the strength depends very much on how many other factors are considered, and how these are controlled in a study. A strong relationship may result from an unacknowledged confounding factor. An example is the strong link between birth order and risk of Down's syndrome. This is actually due to maternal age at the child's birth. A weak association may also be causal, particularly if it is modified by other factors.
3. Intensity or duration of exposure (also called **biological gradient**or **dose-response**relationship): If those with the most intense, or longest, exposure to the agent have the greatest frequency or severity of illness, while those with less exposure are not at as sick, then it is more likely that the association is causal.	A reasonable criterion if present, but may not apply if a threshold level must be reached for the agent to have an effect. Hence the absence of a dose response does not disprove causality.
4. Specificity of association: If an agent or risk factor is found that consistently relates only to this disease, then it appears more likely that it plays a causal role.	This is a weak criterion, and was derived from thinking about infectious diseases. Factors such as smoking or obesity are causally associated with several diseases; the absence of specificity does not undermine a causal interpretation.
5. Consistency of findings: An association is consistent if it is confirmed by different studies; it is even more persuasive if these are in different populations.	A good criterion, although it may lead us to miss causal relationships that apply to only a minority of people. For instance, the drug induced haemolysis associated with glucose-6-phosphate dehydrogenase (GPD) deficiency could be difficult to demonstrate in populations with low prevalence of GPD deficiency.
6. Coherent or plausible findings: Do we have a biological (or behavioural, etc.) explanation for the observed association? Evidence from experimental animals, analogous effects created by analogous agents, and information from other experimental systems and forms of observation should be considered.	A good criterion if we do have a theory. But this can be subjective: one can often supply a *post hoc* explanation for an unexpected result. On the other hand, the lack of a biological explanation should not lead us to dismiss a potential cause. Knowledge changes over time, and new theories sometimes arise as a result of unexpected findings. Truth in nature exists independently of our current ability to explain it.
7. **Cessation of exposure**: If the causal factor is removed from a population, then the incidence of disease should decline.	This may work for a population, but for an individual, the pathology is not always reversible.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Illustrative Materials

Does asbestos cause lung cancer?

The more criteria that are met in a given instance, the stronger is the presumption that the association is causal. For example, did exposure to asbestos fibres among construction workers cause lung cancer in some of them?

1. *Chronological relationship*: Can we be sure that the exposure to asbestos predated the cancer (which may have taken years to develop)?
2. *Strength of the association*: Did groups of workers who had the greatest exposure to asbestos show the highest rates of cancer?
3. *Intensity and duration of the exposure*: Were those with the longest work history the most likely to get sick?
4. *Specificity*: Did they just get lung cancer?
5. *Consistency*: Have similar findings been reported from different countries?
6. *Coherence and plausibility*: Does it make biological sense that asbestos could cause lung cancer?
7. *Cessation of exposure*: After laws were passed banning asbestos, did lung cancer rates decline among construction workers?

In the end, whether or not a factor is accepted as a cause of a disease remains a matter of judgement. There are still defenders of tobacco who can use scientific, evidence-based arguments to point out the flaws in the conclusion that smoking causes cancer and heart disease.

The following sections describe the main quantitative research designs relevant to medicine; Chapter 6 will review measurement methods used in health research. Every study has to minimize two types of error in its design and data collection: bias and random errors. In addition, studies also have to address confounding, a challenge in the interpretation of the results. These are explained in detail later in the chapter, but brief definitions may help the reader at this point.

Definition

Error: "A false or mistaken result obtained in a study or experiment."[1] (We may distinguish between random and systematic errors.

Random error: deviations from the truth that can either inflate, or reduce estimates derived from a measurement or a study. They are generally assumed to be due to chance and, if the sample is large, to have little distorting effect on the overall results. Statistics, such as the confidence interval, are available to estimate their magnitude (see "Sampling and chance error" below).

Systematic error, or bias: the systematic deviation of results or inferences from the truth: a consistent exaggeration of effect, or an underestimate. These may arise from defects in the study design, including the sampling, or may arise from faulty measurement procedures.

Confoundinga challenge in interpreting the results of a study in which the effects of two processes are not distinguished from each other (see "Confounding", below).

Research designs

Unlike qualitative methods, quantitative research is based on systematically sampling people, and uses standardized measurement, analysis, and interpretation of numeric data. Quantitative research uses a variety of study designs, which fall into two main classes: experimental studies (or trials) and observational studies (see Observational study in Glossary).

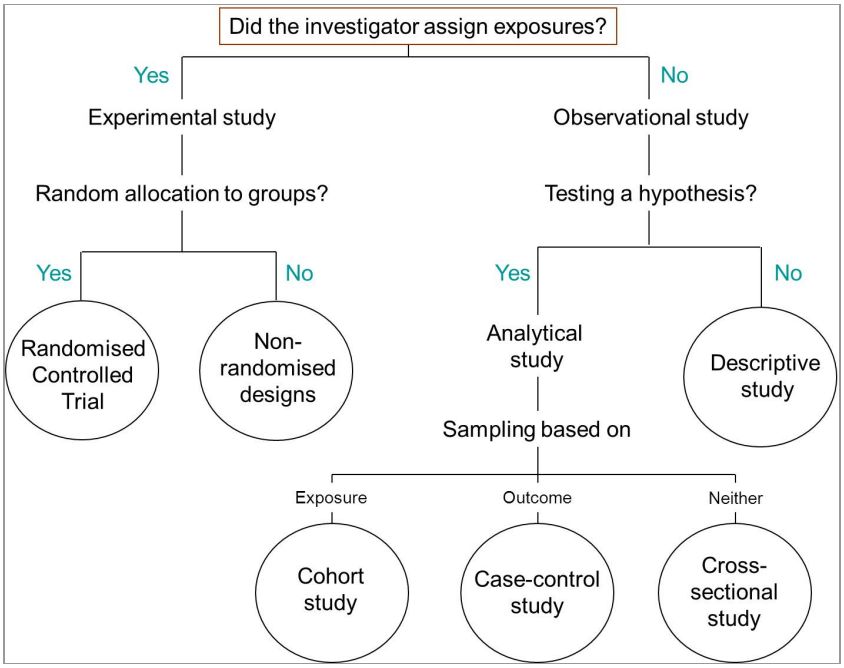


Figure 5.1: What kind of study is it?

Experimental (or interventional) studies

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

As the name implies, these are studies in which the participants undergo some kind of intervention in order to evaluate its impact. An intervention could include a medical or surgical intervention, a new drug, or an intervention to change lifestyle. Because they are the most methodologically rigorous design, experiments are the default choice for providing evidence for best practice in patient management, so this discussion will begin with them. The experimental researcher has control over the intervention, its timing, and dose or intensity. In its simplest form, an experimental study to test the effect of a treatment will follow these steps:

1. The researcher formally states the hypothesis to be tested

2. The researcher selects people eligible for the treatment

3. The sample is divided into two groups

4. One group (the experimental, or intervention group) is given the intervention while the other (the control group) is not

5. Outcomes of interest are recorded over time, and the results compared between the two groups.

Step 3 leads to a critical distinction, shown at the left of Figure 5.1: the distinction between a randomized controlled trial and non-randomized designs. In the former, people are allocated to intervention and control groups by chance alone, while in the latter the choice of who receives the intervention is decided in some other way, such as according to where or when they were recruited, or by the sequence in which they enter the study. There are many types of non-randomized studies and, because the researcher often does not have complete control over the allocation to experimental or control group, they are regarded as inferior to true randomized designs (see Nerd's corner box). They are often called *quasi-experimental designs*.

Nerd's Corner

Quasi-experimental designs

An example of a quasi-experimental study would be to treat hypertensive patients attending one hospital with one protocol and compare their results to patients receiving a different treatment protocol in another hospital. This has the advantage of being simple: there is no need to randomise patients in each hospital, and staff training is greatly simplified. However, many biases might arise in such a study: patients might choose which hospital or clinician they attend (self-selection); one hospital may treat more severe patients; other aspects of care in the two hospitals may be different, and so forth.

Another quasi-experimental study is the time series design. This refers to a single-group research design in which the comparison involves measurements made before and after some intervention, thereby allowing trends to be compared to detect the impact of the intervention. For example, to examine whether introducing a new textbook on public health has any impact on student learning, LMCC exam marks could be compared for successive cohorts of medical students, before and after the introduction of the new textbook. This design can be considered an experiment if the investigator has control of the timing of the introduction of the book; otherwise it is an observational study. Again, this design has the virtue of feasibility: it would be difficult to randomly allocate some student to have the book and others not, because the book might be shared between the two groups.

Nonetheless, quasi-experiments have sufficient sources of potential bias that they are regarded as substantially inferior to true randomized experiments, so their findings are rarely considered conclusive.

The key advantage of a random assignment is that the random allocation of people into different groups means that other factors that could affect the outcome (i.e., **confounding** factors) are likely to be equally represented in each study group?including unknown ones, such as genetic characteristics that affect prognosis. On average, assuming no systematic error or bias, the only thing that differs between two randomly allocated groups is the intervention, so that any differences in results are probably related to the intervention. The larger the study sample the more confident we can be that other factors are equivalent in the two groups, so any difference found really is due to the intervention. But this is still a matter of probabilities, and this is why we need tests of statistical significance. These show the likelihood that observed differences between experimental and control groups have arisen merely by chance.

Nerd's Corner

Not always truly random

For practical reasons, some trials use non-random patient allocation. For instance, alternate patients may be assigned to treatment and control groups. This is superior to participants themselves choosing which group to join, and may approach the quality of a random allocation. The method of allocation should be scrutinized for possible biases. For instance, if all Monday morning's patients are assigned to the treatment group and Monday afternoon's patients go to the control group, are there likely to be systematic differences between the groups that could affect differences in outcomes?

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Here Be Dragons

Random selection of subjects versus random allocation to study groups

Distinguish between random *selection* of subjects from a sampling frame and random *allocation* of subjects to experimental or control groups. Non-random selection of subjects is mainly an issue in descriptive research and can lead to results that cannot be generalized to the general population (see the section on sampling bias). Non-random allocation to experimental and control groups can give results that are confounded by inherent differences between groups. For instance, Monday morning's patient may be from the local area, while Monday afternoon's patients may be from another town, arriving in the afternoon because that is when the bus from that town arrives. (See Nerd's corner "Not always truly random")

Randomized controlled trials

The most common experimental design in medical research is the **randomized controlled trial** (RCT). An RCT is a true experiment in that the investigator controls the exposure and, in its simplest form, assigns subjects randomly to the experimental or control group (which may receive no treatment, the conventional treatment, or a placebo). Both groups are followed and assessed in a rigorous comparison of their rates of morbidity, mortality, adverse events, functional health status, and quality of life. RCTs need not be limited to two groups; a number of different treatment regimens may be compared at once. RCTs are most commonly used in therapeutic trials but can also be used in trials of prevention. They are often conducted across many centres, as illustrated by clinical trials of cancer treatments (see box).

Links

Cancer therapy trials in Canada

The National Cancer Institute of Canada Clinical Trials Group undertakes multi-centre trials to increase sample size and also to enhance generalizability of findings. For more information on the NCI Clinical Trials Group, visit the Canadian Cancer Society website at:

http://www.cancer.ca/research/partnerships%20and%20programs/national%20programs/ncic%20ctg.aspx?sc_lang=en

The steps in an RCT are:

1. State the hypothesis in quantitative and operational terms

2. Select the participants. This step includes calculating the required sample size, setting inclusion and exclusion criteria, and obtaining free and informed consent

3. Allocate participants randomly to either the treatment or control group. Randomization removes allocation bias, increases the chances that any confounding factors will be distributed evenly between both groups, and it allows the valid use of statistical tests. Note that there may be more than one intervention group, for example, receiving different doses of the experimental medication

4. Administer the intervention. This is preferably done in a blinded fashion; so that the patient does not know which group she is in. Ideally, the researcher (and certainly the person intervening and monitoring the patient's response) should also not know which group a given patient is in (this is called a *double-blind experiment*). This helps to remove the influence of the patient's and the clinician's expectations of the treatments, which could bias their assessment of outcomes

5. At a pre-determined time, the outcomes are monitored (e.g., physiological or biochemical parameters, morbidity, mortality, adverse events, functional health status, or quality of life) and compared between the intervention and control groups using statistical analyses. This indicates whether any differences in event rates observed in the two groups are greater than could have occurred by chance alone. Sometimes, those who are analysing the data and interpreting the results do not know which group received which treatment until the analysis is complete. This would be called a triple blind experiment.

While RCTs are regarded as the best research design we have, they do have limitations. By design, they study the efficacy of a treatment under carefully controlled experimental conditions, so they may not provide evidence on how well the treatment will work in the field. **Efficacy** refers to the potential impact of a treatment under the optimal conditions typical of a controlled research setting. **Effectiveness** refers to its impact under the normal conditions of routine practice: for example, in trial conditions the medication may be efficacious because patients know that they are participating in a research project and are being supervised. In the real world the medication may not be effective because, without supervision patients may not take all of their medication in the correct dose. An efficacious intervention may also not be efficient enough to put into practice. Breast-self examination has been shown to detect early breast cancer, but only in trials conditions in which women received constant follow-up by trained nurses. This level of intervention was too costly to be generally applicable.

Furthermore, trials are often conducted on selected populations (e.g., men aged 50 to 74 who are current smokers with unstable angina, no co-morbidity and willing to participate in a research study); this may reduce the extent to which results can be generalized to typical angina patients. Trials may also face problems due to attrition if participants drop out of the study. Finally, intervention trials, although designed to detect differences in the known and desired outcomes, may not be large enough to reliably detect previously unknown or rare effects.

An adaptation of the RCT is the ?N of 1' trial which can have particular value in testing the application of a treatment to a particular patient in a way that avoids most sources of bias (see Nerd's Corner box titled "N of 1 trial").

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

N of 1 trial

An **N of 1 trial** is a special form of clinical trial that studies a single patient and can be useful in evaluating a treatment for a particular patient. It requires that the effect of the treatment can be reversed. The patient receives either the active treatment or a control (e.g., placebo), determined randomly and administered blindly. Outcomes are recorded after a suitable time delay, followed by a washout period when the patient is not receiving the medication to eliminate remaining traces of it. The patient then receives the alternate treatment (placebo or active) and outcomes are evaluated. The cycle may be repeated to establish stable estimates of the outcomes. The main advantage is that the study result applies specifically to this patient and allows for careful calibration to optimize the therapeutic dose. The results cannot be generalized beyond this particular patient.

Some studies use an N of 1 approach but in a larger group of people. These can produce highly valid results because almost all sources of bias are eliminated, because each patient acts as his own control.

Ethics of RCTs

Some special ethical issues (see **Ethics** in Glossary) arise in the conduct of all medical experiments. A tension may arise between two basic principles: it is unethical to deny a patient access to an effective treatment, but it is also unethical to adopt a new treatment without conducting rigorous testing to prove efficacy. Therefore, if there is evidence that a treatment is superior, it may be unethical to prove this in a randomized trial because this would entail denying it to patients in the control group. Hence, the RCT can only ethically be applied when there is genuine uncertainty as to whether the experimental treatment is superior; this is termed equipoise. It is also unethical to continue a trial if the treatment is found to be obviously effective or obviously dangerous. Trials are therefore planned with pre-set stopping rules that specify conditions under which they should be prematurely concluded (see box ?Early termination of trials'). It is also unethical to conduct trials that offer only marginal benefits in terms of broader social value (e.g., studies that benefit the publication record of the researcher more than the health of patients or studies which double as marketing projects). These ethical principles mean that many established treatments will probably never be evaluated by a controlled trial:

- Appendectomy for appendicitis
- Insulin for diabetes
- Anaesthesia for surgical operations
- Vaccination for smallpox
- Immobilization for fractured bones
- Parachutes for jumping out of airplanes, as the British Medical Journal humorously noted.[7]

Links

Ethics of research involving humans

For information on the ethical conduct of RCTs and other types of research, please see theTri-Council Policy Statement on Ethical Conduct for Research Involving Humans, available at: http://www.pre.ethics.gc.ca/eng/policy-politique/tcps-eptc/

Nerd's Corner

Early termination of trials

In principle, intervention trials are designed to include the minimum number of patients necessary to demonstrate the effectiveness of the intervention. Therefore, in principle, statistical significance will not be reached until all results have been obtained from the planned number of participants. Early observations of large differences between groups could be due to chance. Researchers may have a choice between stopping a trial before the number of participants is large enough to demonstrate adequately the superiority of one course of action or to continue the trial even though, at the moment, one course of action appears greatly superior to the other. Researchers should use methods that allow continuous monitoring of outcomes so that, as soon as clinically significant differences occur, the trial can be stopped.

Phases of intervention studies

Once a new pharmaceutical treatment has been developed, it undergoes testing in a sequence of phases before it can be approved by regulatory agencies for public use. Randomized trials form one stage within this broader sequence, which begins with laboratory studies using animal models, thence to human testing:

Phase I: The new drug or treatment is tested in a small group of people for the first time to determine safe dosage and to identify possible side effects

Phase II: The drug or treatment is given to a larger group at the recommended dosage to determine its efficacy under controlled circumstances and to evaluate safety. This is generally not a randomized study

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Phase III: The drug or treatment is tested on large groups to confirm effectiveness, monitor side effects, compare it to commonly used treatments, and to collect information for the safe use of the drug. Phase III testing normally involves a series of randomized trials. At the end of this phase, the drug may be approved for public use. The approval may limit how the drug can be used, for instance in specific diseases or in certain age groups

Phase IV: After the treatment enters the marketplace, information continues to be collected to describe effectiveness on different populations and to detect possible side effects. This does not involve an RCT and is called post-marketing surveillance; it is based on reports of side effects from physicians (and patients) so it requires the active participation of treating physicians and is necessary to detect rare or slowly developing side effects.

Links

The clinical trials registry

Sometimes, studies that produce negative results go unreported, and yet these results are important in providing clinical evidence that a treatment may not work. To prevent the bias that would result from failing to report negative results, all planned trials have to be registered and a comprehensive list of clinical trials is maintained by the U.S. National Institutes of Health Clinical Trials. See their website at: http://www.clinicaltrials.gov/

A checklist for judging the completeness of reporting an RCT is provided on the Consolidated Standards for Reporting Trials (CONSORT) website at: http://www.consort-statement.org/index.aspx?o=1031

General information on the CONSORT group is available at: http://www.consort-statement.org/

Observational studies

In observational studies, the researcher observes what happens to people under exposure conditions that have been self-selected or have been determined by influences outside the control of the researcher. The researcher can choose what exposures to study, but does not influence them. As this is a non-randomized design, the major problem in inferring causation is that the exposed and unexposed groups may differ on other key factors that may themselves be true causes of the outcome, rather than the characteristics under study. Such factors are known as *confounders*.

Descriptive studies

Descriptive studies describe how things are; they do not set out to test hypotheses. For instance the Canadian Community Health Survey describes health and health habits in the Canadian population, or a family physician might describe the demography of patients attending her practice. They are usually cross-sectional in design. Surveys are often used in a descriptive manner, for example to establish disease prevalence, or to record who uses health services and what patients think of them. This kind of information can be useful for clinicians deciding what kinds of information to offer their patients, or what services they ought to provide. They are particularly useful for public health and health care planning. Descriptive information is often collected by surveys or by surveillance programmes, covering person, place, and time of disease occurrences.

Analytical studies

The critical distinction between a descriptive and an analytical study is that the latter is designed to test a hypothesis. When an outcome variable, such as heart disease, is studied in relation to an exposure variable such as body weight, the study does more than count: it tests a hypothesis predicting an association between the two. Analytical observational studies can be of three types, depending on the time sequence and sampling procedures used to collect data.

Cross-sectional studies

Here, subjects are selected irrespective of the presence or absence of the characteristics of interest for hypothesis testing. One of the most common cross-sectional analytical studies is the survey, in which a random sample is drawn to give an accurate representation of the population. It is similar to a descriptive survey except that the purpose of the analysis is to record associations between variables, rather than merely to report frequencies of their occurrence.

As an example of a cross-sectional study, a researcher might draw a random sample of people to test hypotheses concerning the association between feelings of stress and the use of medical services. The researcher might ask whether people had visited a doctor in the last 2 weeks, and if they were under stress in the last year. Suppose the sample included over 18,000 people about stress and doctor visits, producing the following result:

Table 5.5: Stress and physician visits: calculating the association between two variables

		Doctor visit in the last 2 weeks?		
		Yes	No	Total
Stress in the last year?	Yes	1,442	3,209	4,651
	No	2,633	11,223	13,856
	Total	4,075	14,432	18,507

Note that this result can be reported in either of two ways:

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

1. Of those who suffered stress in the last year, 31% (1442/4651) visited their doctor in the last 2 weeks compared with only 19% (2633/13856) of those who did not suffer stress.

2. Of those who visited their doctor in the last 2 weeks, 35% (1442/4075) suffered stress in the previous year, compared with 22% (3209/14432) of those who did not visit their doctor.

Either approach is correct. The researcher is free to decide which way to report the results; the study design allows both types of analysis. All that can be concluded is that there is an association between the two variables. It might be supposed that stress predisposes people to visit their doctor, or could it be that the prospect of a visit to the doctor causes stress, or perhaps something else (fear of an underlying illness?) causes both. This study cannot provide support for an inference of causation because in this cross-sectional study it is impossible to know if stress pre-dated the doctor visit.

Nerd's Corner

Ecological studies

Ecological studies measure variables at the level of populations (countries, provinces) rather than individuals. They are the appropriate design for studying the effect of a variable that acts at the population level, such as climate change, an economic downturn, or shortages of physicians. Like a survey, they can be descriptive or analytic. They have the advantage that they can often use data that are already available, such as government statistics. Ecological studies can be useful for generating hypotheses that can be tested at the individual level. For example, the hypothesis that dietary fat may be a risk factor for breast cancer arose from a study that showed a correlation between per-capita fat consumption and breast cancer incidence rates across several countries. Countries with high per-capita fat consumption had higher incidences of breast cancer.

However, there is a logical limitation in drawing conclusions from ecological studies for individual cases. Because the ecological finding was based on group averages, it does not necessarily show that the *individuals* who consumed a high fat diet were the ones most likely to get cancer. Perhaps the breast cancer cases living in countries with high fat diets actually consumed little fat: we cannot tell from an ecological study. This difficulty in drawing conclusions about individuals from ecological data is called "the ecological fallacy." To draw firm conclusions about the link between dietary fat and breast cancer risk, the two factors must be studied in the same individuals. Nevertheless, ecological studies are often used as a first step, to suggest whether or not a more expensive study of individuals may be worthwhile.

Cohort studies

A cohort is a group of people who can be sampled and enumerated, who share a defining characteristic and who can be followed over time: members of a birth cohort share the same year of birth, for example. Cohort studies of health commonly study causal factors; the characteristic of interest is usually some sort of exposure that is thought to increase the likelihood of a health outcome. A **cohort study** typically begins with a sample of people who do not have the disease of interest; it collects information on exposure to the factor being studied, and follows exposed and unexposed people over time (Figure 5.2). The numbers of newly occurring (incident) cases of disease are recorded and compared between the exposure groups. Cohort studies are also known as longitudinal or follow-up studies.

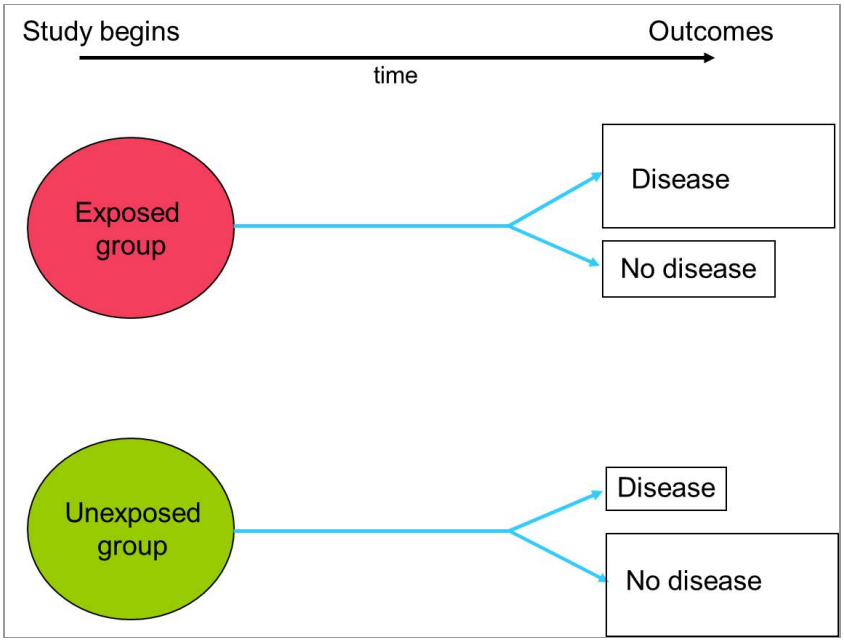


Figure 5.2: Schema of a cohort study

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

Cohort

Cohort: from Latin *cohors,* meaning ?an enclosure.' The meaning was extended to an infantry company in the Roman army through the notion of an enclosed group or retinue. Think of a Roman infantry cohort marching forward; some are wearing new metal body armour, while others have the old canvas-and-leather protection. Bandits shoot at them and General Evidentius can record the outcomes and his trusty scribe, Epidemiologicus, can compare the results using simple arithmetic.

In simple cohort studies the results can be fitted into a "2 by 2" table (2 rows by 2 columns ? don't count the Total column).

Table 5.6: Generic format for a 2 x 2 table linking an exposure to an outcome.

	Outcome (e.g., disease) present	Outcome (e.g., disease) absent	Total
Exposure (or risk factor) present	a	b	a + b
Exposure (or risk factor) absent	c	d	c + d

The incidence, or risk of disease in the exposed group, is calculated as a / (a + b). Correspondingly, the risk in the non-exposed people is c / (c + d). These risks can be compared to get a risk ratio (often called a **relative risk**, or RR) that gives an estimate of the strength of the association between the exposure and the outcome: [a/(a + b)] / [c/(c + d)]. A relative risk of 1.0 indicates that exposed people are neither more nor less likely to get the disease than unexposed people: there is no association between exposure and disease. A relative risk greater than 1.0 implies that, compared to a person not exposed to the factor, a person who has been exposed has a greater chance of becoming diseased, while a relative risk of less than 1.0 implies a protective effect, that is exposed people have a lower chance of becoming diseased compared to unexposed people. The fact that exposure was recorded before the outcomes is the main advantage of cohort studies; they can clearly establish the causal criterion of a temporal sequence between exposure and outcome as long as study participants truly did not have the disease at the outset. Furthermore, because recording of exposures and outcomes is planned from the beginning of the study period, data recording can be standardized.

Definition of exposure groups

Imagine a cohort study designed to test the hypothesis that exposure to welding fumes causes diseases of the respiratory tract. The sample could be drawn on the basis of a crude indicator of exposure, such as using occupation as a proxy (welders are assumed to be exposed; a non-welding occupations would be presumed to be unexposed). This approach is frequently used in occupational and military epidemiology. A more detailed alternative would be to quantify levels of exposure (e.g., from the person's welding history); this requires considerably more information but would permit dose-response to be estimated?one of the criteria for inferring causation (see Table 5.4). In an extension of this quantified approach, a cohort study might not select an unexposed group to follow, but rather select a sample of individuals with sufficient variability in their exposure to permit comparisons across all levels of exposure, or to permit mathematical modelling of exposure. Cohort studies of diet, exercise, or smoking often use this approach, deriving information from a baseline questionnaire. This approach has been used in community cohort studies such as the Framingham Heart Study.

Illustrative Materials

The Framingham Study

Since 1948, the town of Framingham, Massachusetts, has participated in a cohort study to investigate the risk factors for coronary heart disease. The study has now collected data from two subsequent generations of the families initially involved. This has produced quantitative estimates of the impact of risk factors for cardiac disease, covering levels of exercise, cigarette smoking, blood pressure, and blood cholesterol. Details of the Framingham studies may be found at the Framingham Heart Study website: http://www.framinghamheartstudy.org/

Here Be Dragons

A cohort study proves??

In drawing conclusions from a cohort study, it is tempting to assume that the exposure causes the outcome. The study can demonstrate that there is a temporal association between the two; the exposure is associated with the outcome and occurs before the outcome. The study can also demonstrate a dose-response gradient. This meets two of the criteria for a causal relation, but **confounding** can remain an issue (see section on Confounding).

This may explain why so many studies in epidemiology produce conclusions that conflict with one another; for this reason there is a strong emphasis on undertaking a randomized controlled trial, where it is feasible and ethical to do so. An example is that of hormone replacement therapy and cardiovascular disease in women. (See illustration box "An example of confounding")

Case-control studies

Case-control studies (see **Case-control study** in Glossary) compare a group of patients with a particular outcome (e.g., cases of pathologist-confirmed pancreatic cancer) to an otherwise similar group of people without the disease (the controls). As shown in Figure 5.3, reports or records of

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

exposure (e.g., alcohol consumption) before the onset of the disease are then compared between the groups. The name of the design reminds you that groups to be compared are defined in terms of the outcome of interest: present (cases) or absent (controls).

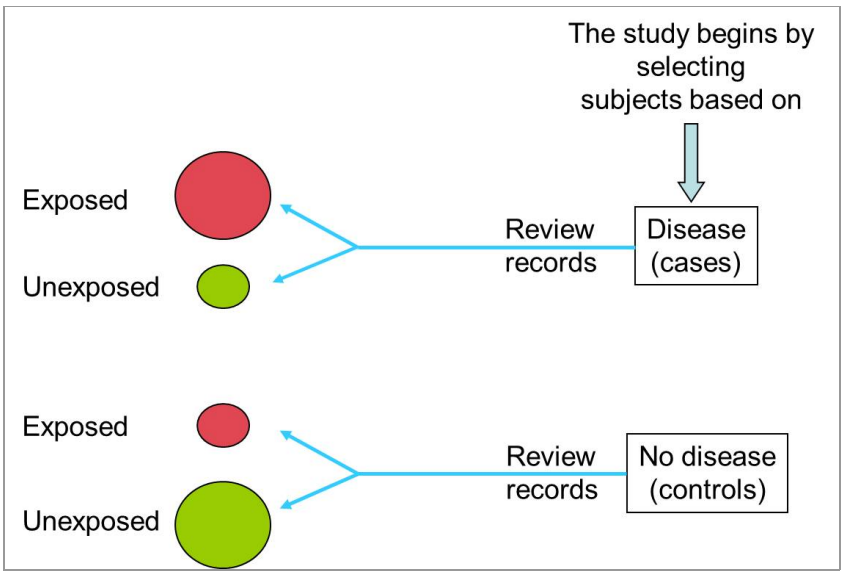


Figure 5.3: Schema of a case-control design

Notice that a case-control study does not allow the calculation of the incidence or risk of the disease, because it begins with people who already have it and a predetermined number who do not. Therefore, a risk ratio cannot be calculated. Instead, the study identifies the exposure status of a sample of cases and another of controls. This information allows the calculation of the odds of a case having been exposed?the ratio of a:c in the 2 x 2 table (Table 5.7). This can be compared to the odds of a control having been exposed, the ratio of b:d. The result of the case-control study is then expressed as the ratio of these 2 odds, giving the **Odds Ratio** (OR): (a/c) / (b/d). To make the calculation easier, this is usually simplified algebraically to ad/bc.

Table 5.7: Generic 2 x 2 table for calculating an odds ratio

	Outcome (or disease) present	Outcome (or disease) absent
Exposure (or risk factor) present	a	b
Exposure (or risk factor) absent	c	d

The OR calculated from a case-control study can approximate a relative risk, but only when the disease is rare (say, up to around 5% in the sample, as is the case for many chronic conditions). The interpretation of the value of an OR is the same as a RR. Like a relative risk, an OR of 1.0 implies no association between exposure and disease. A value over 1.0 implies a greater chance of diseased people having been exposed compared to controls. A value below 1.0 implies that the factor is protective. This might occur, for example, if a case-control study showed that eating a low fat diet protected against heart disease.

Key contrast between cohort and case-control studies

In cohort studies, the participants groups are classified according to their exposure status (whether or not they have the risk factor).

In case-control studies, the different groups are identified according to their health outcomes (whether or not they have the disease).

Here Be Dragons

Prospective or retrospective?

These terms are frequently misunderstood, and for good reason.

Cohort studies define the study groups on the basis of initial exposure, but that could have occurred before the actual study was undertaken. A study beginning today might obtain employment records showing levels of exposure among people who currently work as welders and then follow them over time to check outcomes several years hence. This would be a prospective cohort study. But it would be quicker to work from occupational records and select people who worked as welders 30 years ago, and compare their health status now, according to their previous level of exposure. This could be called a retrospective cohort study, but is better called an historical cohort study.

The word *retrospective* causes confusion because it was formerly used to refer to case-control studies. Most authorities have now abandoned the term entirely.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Measures of risk: attributable risk and number needed to treat

The RR and OR indicate how much an individual's risk of disease is increased by having been exposed to a causal factor, in *relative terms*. Both statistics answer the question "Compared to someone without the risk factor, how many times as likely am I to get the disease?", giving the answer as a ratio: "You are twice as likely", or "10% more likely". A patient's question, however, often concerns *absolute risk,* which relates to disease incidence and answers the question "What is my chance of getting the disease (in the next year, ten years, my lifetime)?" The answer is given as a proportion, such as 1 in 10, or 1 in 100. An important point to bear in mind when communicating with a patient is that if the disease is rare, the RR of having a risk factor can appear quite frightening?100% greater risk of death in the next year?even though the absolute risk is small. A relative increase of 100% on an absolute risk of 1 in a million is still only 2 in a million.

Judging the magnitude of a risk introduces the concept of attributable risk, which indicates the number of cases of a disease among exposed individuals that can be attributed to that exposure:

Attributable risk = Incidence in the exposed group – incidence in the unexposed.

This tells us how much extra disease has been caused by this exposure, in absolute terms: 1 case per million persons in the example above. In the case of a factor that protects against disease, such as a vaccination, it tells us how many cases can be avoided. Sometimes this value is expressed as a proportion of the incidence in exposed persons, yielding the *exposed attributable fraction, EAF*:

$EAF = [Incidence_{(exposed)} - Incidence_{(unexposed)}] / Incidence_{(exposed)}$

This statistic can be useful in counselling an exposed patient: "not only are you at high risk of developing lung cancer, but 89% of your risk is attributable to your smoking. Quitting could have a major benefit".

We can also apply the idea of attributable risk to describing the impact of risk factors on the population. This introduces measures of population attributable risk (PAR) and of population attributable fraction (PAF): statistics that evaluate the impact of a causal factor by substituting incidence in the whole population for incidence in the exposed (see Nerd's corner).

Nerd's Corner

Population attributable risk

In discussing the impact of preventive programmes, the *population attributable risk* indicates the number of cases that would not occur if a risk factor were to be eliminated:

$Incidence_{(population)} - Incidence_{(unexposed)}$

Sadly, this statistic is almost never used, despite its obvious utility in setting priorities for health policy. Expressed as a proportion of the incidence in the whole population, it yields the *population attributable fraction* (which also goes by half a dozen other names):

$[Incidence_{(population)} - Incidence_{(unexposed)}] / Incidence_{(population)}$

This most valuable statistic for public health shows the proportion of all cases of a disease that is attributable to a given risk factor, and was used (for example) to estimate that 40,000 Canadians die from tobacco smoking each year. A little algebra shows that it depends on the prevalence of the risk factor and the strength of its association (relative risk) with the disease. The formula is:

$PAF = P_e (RR_e-1) [1 + P_e (RR_e-1)]$,
where P_e is the prevalence of the exposure (e.g., the proportion who are overweight) and RR_e is the relative risk of disease due to that exposure.

The *population prevented fraction* is the proportion of the hypothetical total load of disease that has been prevented by exposure to the protective factor, such as an immunization programme. The formula is:
$P_e (1-RR)$.

A useful application of the attributable risk is the concept of ?Number Needed to Treat' (NNT). This is a metric that summarizes the effectiveness of a therapy or a preventive measure in achieving a desired outcome. The basic idea is that no treatment works infallibly, so the number needed to treat is the number of patients with a condition who must follow a treatment regimen over a specified time in order to achieve the desired outcome for one person. The NNT is calculated as the reciprocal of the absolute improvement. So, if a medication cures 35% of the people who take it, while 20% recover spontaneously, the absolute improvement is 15%. The reciprocal = 1 / 0.15 = 7. So, on average, you would need to treat 7 people to achieve 1 cure (within the specified time). The NNT can also be applied in describing the value of a preventive measure in avoiding an undesirable outcome, and it can likewise be used in calculating the hazard of treatment such as adverse drug reactions, when it is termed ?number needed to harm'.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Illustrative Materials

Calculating OR, RR, AR, ARR & NNT

A cohort study of the effectiveness of an immunization examined whether or not immunized and non-immunized people became sick. The results are as follows

	Sick	Not sick
Immunized	20 (a)	100 (b)
Not immunized	50 (c)	30 (d)

Total = 200

Odds ratio (OR) (Note that the result is < 1, so immunization protects)	0.12	ad / bc
Relative risk (RR)	0.167 / 0.625 = 0.267	a / (a+b) / c / (c+d)
Attributable risk (AR) (a negative attributable risk indicates protection)	0.167 ? 0.625 = -0.458	a / (a+b) ? c / (c+d)
Absolute risk reduction (ARR) (attributable risk with the sign changed)	0.625 ? 0.167 = 0.458	c / (c+d) ? a / (a+b)
Number-needed-to-treat (NNT)	1 / 0.458 = 2.18	1 / ARR

Sampling and chance error in studies

While there are many different ways that a study could produce a misleading conclusion, it is useful to consider three categories of alternative explanations: chance, bias, and confounding, or CBC.

Chance

The basic purpose of studying a sample is to estimate the value of a parameter (such as the average birth weight of a baby girl) in the broader population from which the sample was drawn. All studies are based on samples of people and people vary, so different samples from the same population are likely to give different results purely by chance. Hence, each estimate of the underlying population parameter derived from a sample carries with it some uncertainty. Minor differences between studies may simply reflect the use of different samples, even when the populations from which the samples were drawn are identical. This is known as ?sampling error' or ?random variation'. Random selection from a population guarantees that, *on average*, the results from successive samples will reflect the true population parameter, but the results of individual samples differ from those of the parent population, and some differ substantially. How do we indicate the likely extent of this uncertainty?

Definition

A **parameter** is the true value in the population, which is the value that a study sample is being used to estimate. By knowing the population parameter, you can interpret values for your patient: Is this child's birth weight reassuringly normal for this population? Parameters in the population are usually designated by Greek letters, while sample estimates of these parameters are indicated by Latin letters:

mean parameter = μ

sample estimate = $\bar{x}$ (called ?x-bar')

standard deviation parameter = σ

sample estimate = s.

Statistics is the branch of mathematics based on probability theory that deals with analysis of numerical data. It is known as biostatistics when it is applied to biological research. In this Primer we provide only a very brief overview of statistical methods; you will have to turn to a statistics text for more information. Statistics provides ways to measure the likelihood that variations in a result (e.g., in mean blood pressure readings in experimental and control groups) could have occurred merely by chance, due to random sampling variation. Remember, you are interested not so much in the actual samples reported in a study, but in whether the results of the study would also apply elsewhere, such as to patients in your practice. This is the essence of evidence-based medicine: if the study results could have occurred by chance alone, you should not base your practice on those results! Statistics can estimate how likely, and by how much, the average characteristic in a sample of a certain size is likely to differ from that of the population as whole. For instance, if you take a sample of people from a population, measure their systolic blood pressure and calculate the mean for the group, statistics can be used to estimate the likely range within which the population mean actually falls.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

Population or descriptive statistics versus sample or inferential statistics

Statistical tests help us to decide whether or not chance variation (i.e., that due to sampling) was a significant factor in the results of a study. The underlying mathematics varies according to what the study aims to do. If it aims to measure the level of a variable in a population, the mathematics of descriptive statistics is used. If a study aims to compare groups, the mathematics of inferential statistics is used. Conveniently, the results of both types of statistical test may be summarized in a similar fashion using *p*-values or confidence intervals, but the formulae and tables used to arrive at the results are different.

The **confidence interval** (or CI) is a statistic used to indicate the possible extent of error in an estimate of a parameter derived from a sample. An expression such as "mean systolic blood pressure was120 mmHg (95% CI 114, 128 mmHg)" indicates that results from the sample suggest that the mean systolic blood pressure in the overall population is 120 mm Hg. Furthermore, based on the size of the sample used and the variability of BP readings within it, there is a 95% probability that the true value of the mean in the whole population will lie between 114 and 128 mmHg. The confidence interval can be represented graphically as a line or error-bar, as seen in Figure 5.7.

Similar logic applies in inferential statistics, when studying differences between groups or associations between variables in a sample. For example, in an RCT of an antihypertensive medication, the researcher needs to know the likelihood that the observed *difference* in mean blood pressure in the experimental and control groups could have occurred by chance. If the results of a statistical test suggest that the difference could readily have occurred by chance, the researcher must conclude that he didn't find any evidence that the therapy reduced blood pressure.

Like mean values, odds ratios and relative risks are also reported with confidence intervals. The confidence interval around an odds ratio indicates whether or not the association is statistically significant, as well as its likely range. For odds ratios and relative risks, if the confidence interval includes 1.0, it is assumed that there is no statistically significant difference between the two groups since an odds ratio or relative risk of 1.0 means that there is no difference between the two groups. For example, a relative risk of 1. 4 (95% CI 0.8, 2.1) means that we can be 95% confident that the true relative risk is somewhere in the range 0.8 to 2.1. Furthermore, because this range includes 1.0, it is quite possible that there is no association in the population. The choice of 95% is arbitrary, but is very commonly used.

Here Be Dragons

Statistical and clinical significance are different

Statistical significance only tells you whether or not the difference observed (e.g., between patients treated with the antihypertensive and others on placebo) could have occurred by chance. It does not directly tell you about the magnitude of the difference, which is important in reaching your clinical decision. For example a drop of 2 mmHg in a trial of BP therapy might be statistically significant, but has negligible clinical importance.

For a new treatment to be useful, the result must be both statistically significant and clinically important. This thinking resembles that behind the Number Needed to Treat statistic, which also offers a way to summarize the amount of improvement produced by a treatment, rather than just whether it was statistically significant.

Nerd's Corner

The limits to statistical significance

When a statistical test shows no significant difference between two groups, this means either that there really is no difference in the population or there may be a difference, but the sample did not reveal it. This is usually because the sample size was not large enough to demonstrate it with confidence (the sample lacked the "power" to detect the true difference). It is intuitively clear that a larger sample will give more precision in any estimate; indeed, if you study the whole population there is no need for confidence intervals or statistical significance because you have measured the actual parameter.

The smaller the true difference (such as between patients treated with a new BP medication and those treated using the conventional therapy), the larger the sample size needed to detect it with confidence. Turning this idea around, if a very large sample size is needed to demonstrate a difference as being statistically significant, the difference must be very small, so you should ponder whether a difference that small is clinically important.

There is no absolute threshold for defining when a finding may be due to chance. The protocol of quantitative studies should indicate the probability level that is going to differentiate between results due to chance and those that are considered statistically significant. The level chosen is usually 5%, which is equivalent to calculating a 95% confidence interval for a relative risk. Then, if the statistical test indicates less than a 5% probability that the result was due to chance, it will be accepted as not merely a chance finding. In reporting the results of a study, this is usually expressed as a *p*-value; $p < .05$ means that the probability of this result happening by chance is less than 5%. The choice of statistical formula used to calculate the *p*-value depends on various elements in the design of the study. The different tests and the conditions of their use will be found in a textbook of biostatistics.

Bias

Bias, or the systematic deviation of results or inferences from the truth, is a danger in the design of any study.[2] Special care is taken by researchers to avoid (or, failing that, to control for) numerous types of bias that have been identified.[8] The many possible biases may be grouped into two general categories: sampling biases (derived from the way that persons in the study were selected) and measurement biases (due to errors in the way that exposures or outcomes were measured).

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Sampling (or selection) bias

Random sampling seeks to select a truly representative sample of people from a broader population; a more formal definition of the idea is that everyone in the population has an equal (and non-zero) chance of being selected. This is especially important in descriptive studies that estimate prevalence. It may be less important in studies that seek to identify abstract scientific truths such as the association between two variables.[9] For instance, a researcher who wants to study the association between arthritis and obesity might be justified in drawing her sample from a population at unusually high risk of obesity in order to get adequate numbers of obese and very obese people in her study.

For practical reasons, very few studies are able to sample randomly from the entire target population, and the researcher usually defines a ?sample frame', which is assumed to be similar to the entire population. The researcher then draws a sample from that frame. So the researcher might sample patients attending Weenigo General Hospital in order to make inferences about patients attending hospitals in Weenigo. Sampling bias may then occur at two stages: first, in the choice of sampling frame, because patients attending Weenigo General may differ from patients attending other local hospitals and, second, in the method used to draw the sample of patients attending the hospital. These ideas are illustrated in Figure 5.4.

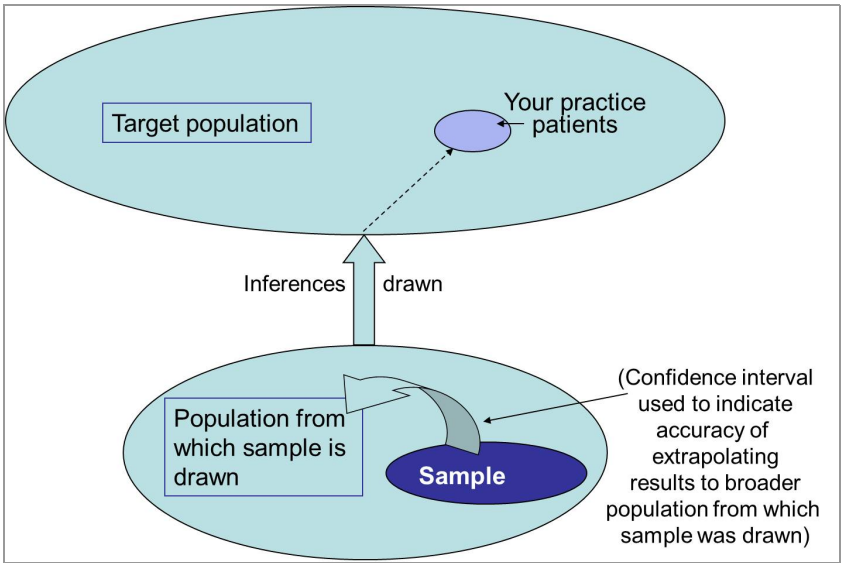


Figure 5.4: Extrapolation from sample to target population

Bias mainly arises when samples are not drawn randomly, so that not everyone in the population has an equal chance of being selected. For example, a newspaper advertisement that reads "Wanted: Participants for a study of blood pressure" might attract retired or unemployed people who have the time to volunteer, especially those who have a personal interest in the topic (perhaps they have a family history of hypertension). If these characteristics are, in turn, associated with blood pressure, an estimate of the population mean BP based on the mean in this sample will be biased. Much research is undertaken in teaching hospitals, but patients seen in these centres differ systematically from patients with the same disease seen in primary care settings?they tend to be sicker, to have more co-morbidities, and often conventional therapy has failed, leading to their referral to tertiary care centres. If studies of tertiary care samples yield different findings than would be seen in the population of all people with a disease, the findings are subject to a specific form of selection bias known as *referral bias*.

Case Study

Dr. Rao notes that the study on the static magnet for menopausal symptoms assembled its sample by placing an advertisement offering women a free trial of the magnet. He worries that women who answered such an advertisement may have been predisposed to believing that the magnet works, a belief that was possibly established by the advertisement itself. They may not be representative of all women with menopausal symptoms.

Illustrative Materials

A biased sample

During the 1948 U.S. presidential elections, a Gallup poll predicted that Dewey, a Republican, was going to win the election against Truman, a Democrat, by a margin of over 10 percentage points. As it turned out, Truman won by 4.4 percentage points. Among the reasons for this poor prediction was the fact that the poll was carried out by telephone. As telephone ownership at the time was limited, and as richer people were more likely to own a phone and also to vote Republican, the sample was probably biased in favour of Republican supporters. This is an example of sampling bias that selected for a confounding variable and that led to a false conclusion. Confounding is discussed below.

Non-response bias

Even if the sampling method is unbiased, not all of those selected will actually participate. If particular types of people did not participate, this *non-response* could bias the study results. One way to detect possible non-response bias is to compare characteristics of participants, such as their age, sex and where they live, with those of people who refused to participate, although, even if these characteristics match, it does not rule out bias on

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

other characteristics that were not recorded. Another way is to compare the eventual sample with information on the population drawn from the Census.

Information bias: systematic measurement errors

Measurement error refers to deviations of recorded values on a measurement from the true values for individuals in the study. As with sampling error, measurement errors may sometimes be random, or they can be systematic. *Social desirability* bias refers to systematic response errors that tend to place the respondent in a socially approved light. For example, most people report that they are more physically active than the average Canadian, which is illogical. Men tend to exaggerate their height and under-estimate their weight.[10] Other biases arise from flaws in the questionnaire design: for example, asking people about their physical activity in February may give a biased estimate of their yearly activity level because many people are more physically active during the warm months than when it is very hot or extremely cold. *Recall bias* commonly occurs in surveys and especially in case-control studies. People's memories often err. For example, questionnaire surveys on time since last mammography report that significantly more women have had mammography in the last two years than studies based on mammography billing records.

Here Be Dragons

Increasing sample size does not reduce sample bias

Increasing sample size can minimize random errors of measurement and sampling but will have no effect on systematic errors; results will remain biased no matter how many subjects participate. A large, biased study may be more misleading than a small one!

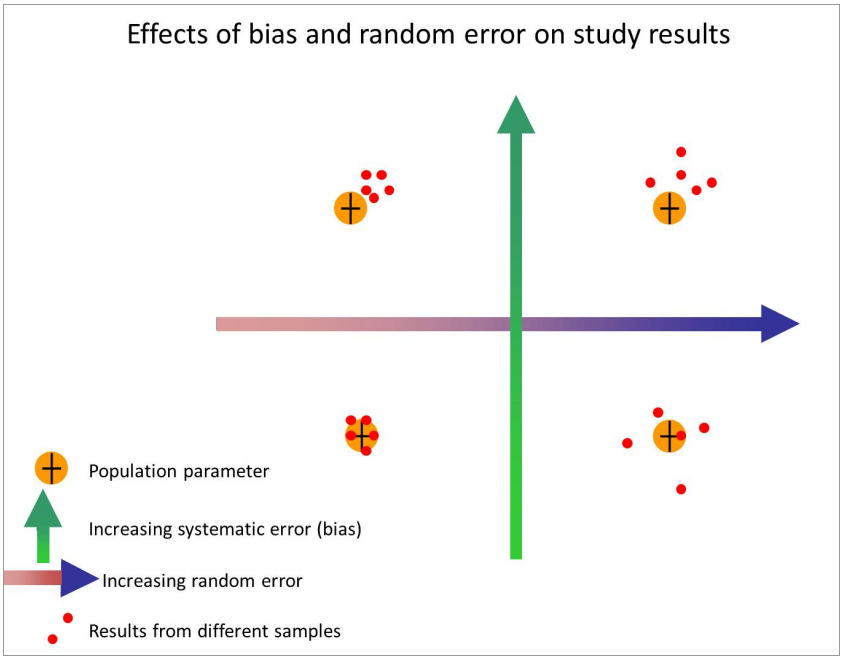


Figure 5.5: Random and systematic study errors

In Figure 5.5, each dot represents an estimate of a parameter obtained from a sample. The top two sections of the figure illustrate the presence of systematic error, and the sample estimates are off-target or biased. In the presence of systematic error, increasing sample size or using more samples will not bring the study results closer to the truth; it may simply make you think that they are more accurate. Increasing the sample size or the numbers of sample in the bottom section, where there is little systematic error, will reduce the uncertainty of the estimation.

The patterns in Figure 5.5 can also be useful in understanding test and measurement validity and reliability, as discussed in Chapter 6. For this, substitute the word "**validity**" for "systematic error" and "**reliability**" for "random error."

Information bias: Objectivity of the investigator

Case Study

When Dr. Rao reads a study report that suggests a relationship between an exposure and an outcome, he needs to be reasonably sure that the results are "true." Journals with a rigorous peer review process generally publish studies of the high quality. By referring to such journals, Dr. Rao may be reassured of the results to a certain extent, but he should still consider other possible explanations for the findings before accepting them as true.

Whether looking at print or Internet information, you should try to find out as much as possible about the source, to check its credibility and to identify possible conflicts of interest. This is why Dr. Rao searched for information on the author of the article on magnets and menopause. Was the author in the medical products business, perhaps selling magnets? Trials published by people with a financial interest in the product under investigation are more likely to conclude in favour of the product than are trials published by people with no financial interest, even when the study results look similar. The U.S. Food and Drug Agency and the Federal Trade Commission have proposed questions to ask when judging the sources of information:

1. Who is behind it?

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

2. Why is the information being published?

3. Where does the information on the website come from?

4. Is the information well documented?

5. Who is responsible for the information?

6. How old is the information?

Links

Assessing claims

Information about how to spot false claims about medical interventions can be found at the Federal Trade Commission website: http://www.ftc.gov/bcp/edu/pubs/consumer/health/hea07.shtm

10 things to know about assessing medical information on the web can be found at:

http://nccam.nih.gov/health/webresources/

This provides only a brief overview of the many potential types of study bias. Several authors have undertaken more extensive reviews, as shown in the Nerd's corner "Types of bias".

Nerd's Corner

Types of bias

Many types of bias can affect the results of research. Epidemiologists are fascinated by bias (perhaps in an attempt to prove they are dispassionate) and in 1979, David Sackett published a ?catalogue' of over one hundred named biases that can occur at different stages in a research project. The following lists only the main headings of Sackett's catalogue.[8]

Literature Review	Study Execution	Analysis
Study Design	**Data Collection**	- Confounding bias
- Selection bias	- Instrument bias	- Analysis strategy bias
- Sampling frame bias	- Data source bias	- Post hoc analysis bias
- Nonrandom sampling bias	- Subject bias	**Interpretation of Results**
- Noncoverage bias	- Recall bias	**Publication**
- Noncomparability bias	- Data handling bias	

Real nerds can look up the original article and review subsequent literature to complete the catalogue. Half-hearted nerds should remember that systematic error can creep in at any stage of a research project, so that research reports should be read critically; the reader should question what happened at every stage and judge how this might affect the results.

Confounding

A study in the 1960s reported a significant tendency for Down's syndrome to be more common in fourth-born or higher order children.[11] There was no obvious sampling or measurement bias and the result was statistically significant, so would you believe the result? The answer may be "yes" in terms of the existence of an association, but "no" if the implication is a causal one. In other words, birth order may be a risk marker, but not a risk (or causal) factor.

Confounding arises when a third variable (or fourth or fifth variable, etc.) in a causal net is associated with both an exposure and the outcome being studied (see Figure 5.6). If this third variable is not taken into account in a study, conclusions about the relationship between two variables may be misinterpreted. In the Down's syndrome example, the mother's age is a confounding factor in that fourth-born and higher order infants tend to be born to older mothers, and maternal age is an independent risk factor for Down's syndrome. In most scientific articles, the first table compares the study groups (e.g., mothers with a Down's infant and others without) on a number of variables that could affect outcome, such as mean maternal age at the child's birth. This allows the reader to determine whether any of these variables is associated with the outcome, and so could potentially act as a confounding factor that should be adjusted in the analysis.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

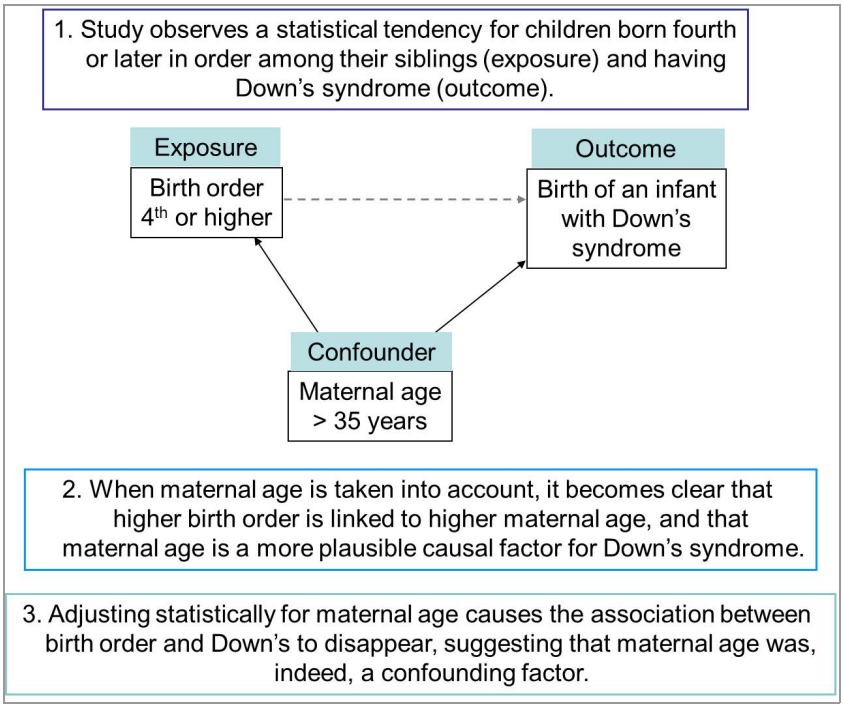


Figure 5.6: An illustration of confounding

Illustrative Materials

An example of confounding: hormone therapy and cardiovascular disease

Before 1990, a number of observational studies concluded that post-menopausal women who took hormone replacement therapy were less likely to develop cardiovascular problems than those who did not. It was, therefore, recommended that all post-menopausal women should take hormone replacement therapy. However, a randomized trial, the Women's Health Initiative Trial, showed quite the opposite: in fact, hormone replacement therapy was associated with an *increase* in cardiovascular disease. HRT recommendations were quickly changed. It seems likely that the observational studies were confounded by social status: women of higher social status were more likely to take hormone replacement therapy and also less likely to have heart disease.[12]

Nerd's Corner

Dealing with confounding

Confounding can be minimized at the design stage of a study, or at the analysis stage, or both.

In experimental designs, random allocation to intervention and control groups is the most attractive way to deal with confounding. This is because random allocation should ensure that all characteristics are equal between the study groups?the more so if the groups are large. Nonetheless, all factors that may confound results should be measured and compared in each group at the start of a study. This should be reported to enable the reader to judge whether, despite randomization, potential confounders were more prevalent in one group than the other.

To complement randomization, the study can be restricted to, for instance, one sex or a narrow age range. This reduces the confounding effect of the factors that are used to restrict the study, but it limits the study' generalizability as its results apply only to that restricted population. Another design strategy is matching: that is, the deliberate selection of subjects so that level of confounders is equal in all groups to be compared. For example, if sex, age, and smoking status are suspected confounders in a cohort study, the researcher records the level of these factors in the exposed group and samples people in the unexposed group who are similar in terms of these factors.

At the analysis stage of a study, stratification can be used to examine confounding. In stratification, the association between exposure and outcome is examined within strata formed by the suspected confounder, such as age. Published reports often mention a Mantel-Haenszel analysis, which is a weighted average of the relative risks in the various strata. If differences arise between the stratum-specific estimates and the crude (unadjusted) estimate, this suggests confounding. Another analytic strategy uses multivariable modelling techniques, such as logistic regression, to adjust a point estimate for the effects of confounding variables. The underlying concept of multivariable modelling is similar to that of standardization (see Chapter 6), the technique used in measurement of health status to allow comparison across populations and time periods. However, selection and measurement biases cannot be corrected at the analysis stage. Here, only careful planning at the design stage for sample selection and the use of standardized measurement procedures can minimize these biases.

The hierarchy of evidence

Study designs vary in the extent to which they can control various forms of error, and hence some designs provide more reliable evidence than others (see section on study designs for information about each). In 1994 the Canadian Task Force on the Periodic Health Examination proposed the idea of a hierarchy of evidence; this arose as a by-product of their work to make recommendations about screening and preventive interventions. The hierarchy implies that clinicians should judge the credibility of evidence according to the type of study that produced it

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

I	Evidence from at least one properly designed RCT
II-1	Evidence from well-designed controlled trials without randomization
II-2	Evidence from well-designed cohort or case-control studies, preferably from more than one centre or research group
II-3	Evidence obtained from multiple time-series studies, with or without the intervention. Dramatic results in uncontrolled experiments (e.g., first use of penicillin in the 1940s) are also included in this category
III	Opinions of respected authorities, based on clinical experience, descriptive studies, reports of expert committees, consensus conferences, etc.

Systematic reviews

Since the Task Force proposed its classes of evidence, advances have been made in undertaking systematic reviewsof the literature and then combining the results from many randomized trials in meta-analyses. The aims of a systematic review are to identify all relevant studies related to a given treatment or intervention, to evaluate them, and to summarize their findings. A key element is the comprehensiveness of the literature review; conclusions should be based on the whole literature and not just on a small, possibly biased selection of studies.[13] A systematic review must follow a rigorous and explicit procedure that allows for its replication. Where the articles included in a review are similar enough, their data can be pooled and meta-analysed.

Meta-analyses

A **meta-analysis** provides a statistical synthesis of data from separate but comparable studies. It can either pool data from all of the people studied and re-analyze the combined data, or else aggregate the results from each study as published, producing an overall, combined estimate. Meta-analyses can be applied to randomized controlled trials and to other study designs including case-control or cohort studies. It is generally accepted that a meta-analysis of several randomized controlled trials offers better evidence than a single trial. In those cases where data cannot be pooled, results of different studies may be summarized in a narrative review or presented in a forest plot, as shown in Figure 5.7.

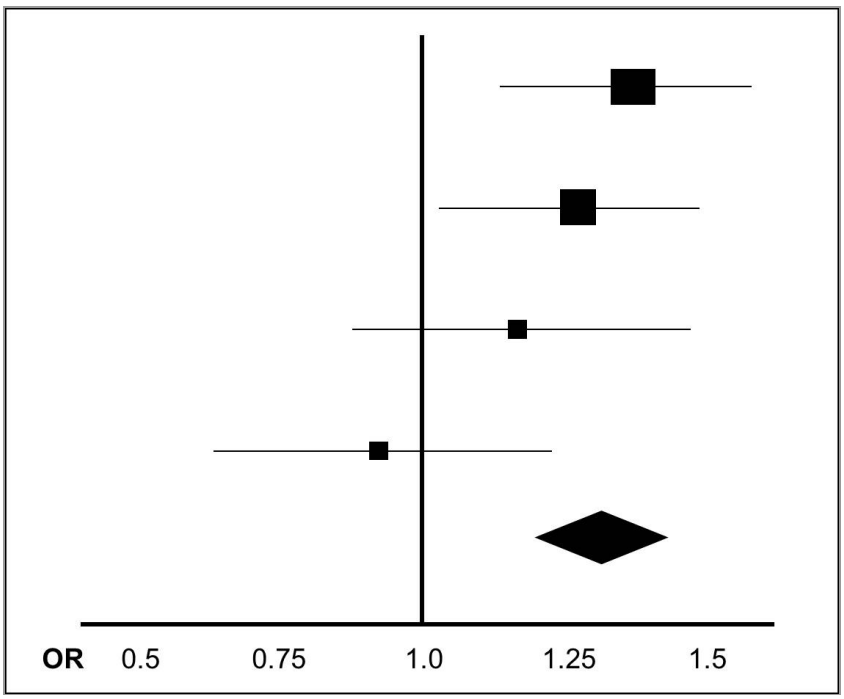


Figure 5.7: An example forest plot, comparing odds ratio results from four case-control studies (square symbols with confidence intervals shown by the horizontal lines), and the result of the meta-analysis (diamond symbol) that pooled the results of the individual studies. The size of the squares indicates the relative sample sizes in each study. The vertical line marks the rate ratio of 1.0, indicating no difference in risk between the study groups. Results on the left of the vertical line indicate a reduction in risk (OR < 1.0; results to its right indicate an increase (OR > 1).

The Cochrane Collaboration

Systematic reviews, including meta-analyses, are normally undertaken by specialized content and research experts, who often work in teams such as those assembled through the Cochrane Collaboration. This is an international organization whose goal is to help scientists, physicians, and decision makers make well-informed decisions about health care by coordinating systematic reviews of the effects of health care interventions. The reviews are published electronically in the Cochrane Database of Systematic Reviews. International collaborative review groups who have expertise in the topic under review and in research methods prepare reviews. An early example of a systematic review was that on the use of corticosteroids in mothers in premature labour to accelerate foetal lung maturation and prevent neo-natal respiratory distress syndrome. Babies born very prematurely

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

are at high risk of respiratory distress owing to their immature lungs, a significant cause of morbidity and mortality. The results of 21 randomised trials give no indication that the corticosteroid treatment increased risk to the mother, and it produced a 30% reduction in neonatal death and similar benefits on a range of other outcomes. Therefore antenatal corticosteroid therapy was widely adopted to accelerate foetal lung maturation in women at risk of preterm birth.

Meta-analyses are now considered to provide the highest level of evidence, so this has changed the original Canadian Task Force hierarchy of evidence. With the idea of literature review in mind, level I of the hierarchy is now sub-divided into

- Cochrane reviews
- Systematic reviews
- Evidence based-guidelines
- Evidence summaries.

The final step: applying the study results to your patients

This brings us full circle, back to critical appraisal of the literature. The first stage is to formulate an overall judgment of the quality of the study or studies, and for this there are a number of checklists. The original ones were developed at McMaster and published in a series of papers that appeared in the *Journal of the American Medical Association* from 1993 to 1994. They addressed critical appraisal in judging evidence for causation, prognosis, the accuracy of diagnosis, and effectiveness of therapy. There are now many critical appraisal checklists that help reviewers evaluate papers. To illustrate the general format, the paper-clipped note lists some questions used to appraise an article about the effectiveness of a therapy.

Links

Critical appraisal

The *Journal of the American Medical Association* papers on critical appraisal are included in an electronic manual by Gordon Guyatt that may be available in your university library. It is available on the web at: http://jamaevidence.com/resource/520

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Checklist for the quality of a study of the effectiveness of a therapy

Are the results valid?

- Were the patients randomized?
- Was randomization concealed?
- Were patients aware of their group allocation? (Were the patients ?blinded'?)
- Were clinicians who treated patients aware of the group allocation?
- Were outcome assessors aware of group allocation? (Were the assessors ?blinded'?)
- Were patients in the treatment and control groups similar with respect to known prognostic variables? (For instance, were there similar numbers of smokers in each group in a study of asthma therapy?)
- Was follow-up complete?
- Were patients analyzed in the groups to which they were allocated?

What are the results?

- How large was the treatment effect?
- How precise was the estimate of the treatment effect?

How can I apply the results to patient care?

- Were the study patients similar to patients in my practice?
- Are the likely treatment benefits worth the potential harm and costs?

What do I do with this patient?

- What are the likely outcomes in this case?
- Is this treatment what this patient wants?
- Is the treatment available here?
- Is the patient willing and able to have the treatment?

For questions to ask about randomized controlled trials (RCT), see the Links box, "the clinical trials registry."

Once you are satisfied that a study provides a valid answer to a clinical question that is relevant to your patient, check that the results are applicable to your patient population.

Target population

Is the study population similar to your practice population, so that you can apply the findings to your own practice (see Figure 5.4)? Consider whether the gender, age group, ethnic group, life circumstances, and resources used in the study are similar to those in your practice population. For example, a study of the management of heart disease might draw a sample from a specialist cardiovascular clinic. Your patient in a family medicine centre is likely to have less severe disease than those attending the specialist clinic and, therefore, may respond differently to treatment. Narrow inclusion and broad exclusion criteria in the study may mean that very few of your patients are comparable to the patients studied. Furthermore, the ancillary care and other resources available in the specialist centre are also likely to be very different from a primary care setting. If these are an important part of management, their absence may erase the benefits of the treatment under study. Other aspects of the environment may also be different: the results of a study set in a large city that found exercise counselling effective may not apply to your patients in a rural area where there are fewer opportunities for conveniently integrating exercise into everyday life.

Intervention

Is the intervention feasible in your practice situation? Do you have the expertise, training, and resources to carry out the intervention yourself? Can you to refer your patient somewhere else, where the expertise and resources are available? In many cases, practical problems of this type mean that an intervention that has good efficacy in a trial does not prove as effective when implemented in usual practice. The enthusiasm and expertise of the pioneers account for some of this difference; the extra funds and resources used in research projects may also have an effect.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

How much does it cost?

The cost includes the money needed to pay for the intervention itself, the time of the personnel needed to carry it out, the suffering it causes the patient, the money the patient will have to pay to get the intervention, ancillary medical services, transport, and time off work. An intervention that costs a great deal of time, money or suffering may not be acceptable in practice.

Intervention in the control group

What, if anything, was done for people in the control group? If they received nothing, including no attention by the researchers, could a placebo effect of people in the active intervention group receiving attention have accounted for part of the effect? In general, new interventions should be compared to standard treatment in the control group so the added benefits and costs of the new can be compared with the benefits and costs of the old.

Your patient's preferences

Finally, all management plans should respect patients' preferences and abilities. The clinician should explain the risks and benefits of all treatments? especially novel ones?in terms the patient understands, and support the patient in choosing courses of action to reduce harm and maximize benefit.

Self-test questions

Self-Testing

1. In your capacity as advisor to the Minister of Health, how would you design and implement a series of studies to determine the relationship between personal stereo use and noise-induced hearing loss?
First, discuss study designs. Assuming that a randomized trial on humans would be unethical, what is the feasibility of a cohort study? What would the timeline be? If you resort to a case-control study, how would you collect information on listening volume? Do you actually need to study this at the individual level? Can you get a crude but useful approximation at the population level by correlating the incidence of deafness with sales of devices (hence you have to interview no one)? Second, consider data collection: How accurate may self-reporting be? What might be the accuracy of retrospective self-reporting of usage among people who had become deaf?: Might they report higher use given that they are now suffering a medical problem? Could you instead modify some personal stereo devices to record how often they are used, and at what volume?

References

1. Lett, J. A field guide to critical thinking. Skeptical Inquirer [serial online] 1990;4(14)http://www.csicop.org/si/show/field_guide_to_critical_thinking/(Accessed Nov 2009).

2. Porta, M.e. A Dictionary of Epidemiology. 5th ed. New York (NY): Oxford University Press, 2008.

3. Sackett, D., et al. Evidence-based medicine - how to practice and teach EBM. 2nd edition. London: Churchill-Livingstone, 2000.

4. Schwandt, T. Qualitative inquiry: a dictionary of terms. Oaks (CA): Sage Publications, 1997.

5. Richards, I., Morese, J. Read me first for a users' guide to qualitative methods. 2nd ed. Thousand Oaks (CA): Sage Publications, 2007.

6. Cockburn, J., Hill, D., Irwig, L., De Luise, T., Turnbull, D., Schofield, P. Development and validation of an instrument to measure satisfaction of participants at breast cancer screening programmes. Eur J Cancer 1991;27(7): 827-831.

7. Smith, G., Pell, J. Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. BMJ 2003;327(7429): 1459-1491.

8. Sackett, D. Bias in analytic research. J Chron Dis 1979;32: 51.

9. Miettinen, O. Theoretical epidemiology: principles of occurrence research in medicine. New York (NY): John Wiley, 1985.

10. Connor-Gorber, S., Shields, M., Tremblay, M., McDowell, I. The feasibility of establishing correction factors to adjust self-reported estimates of obesity. Health Reports 2008;19(3): 71-82.

11. Renkonen, K., Donner, M. Mongoloids: their mothers and sibships. Ann Med Exp Biol Fenn 1964;42: 139-144.

12. Anderson, G., Judd, H., Kaunitz, A., et al. Effects of estrogen plus progestin on gynecologic cancers and associated diagnostic procedures: The Women's Health Initiative randomized trial. JAMA 2003;290: 1739-1748.

13. Liberati, A., Altman, D., Tetzlaff, J., et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Ann Intern Med 2009;151: W65-W94.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Chapter Objectives

After reading this chapter, you will be able to:

1. Analyze population health data using appropriate measures:

Describe the concepts of, and be able to calculate, incidence, prevalence, attack rates, case fatality rates (see Case fatality rates in Glossary), and to understand the principles of standardization.

1. Interpret and present the analysis of health status indicators:

1. Demonstrate an ability to use practice-based health information systems to monitor the health of patients and to identify unmet health needs;

2. Understand the appropriate use of different graphical presentations of data (bar charts; line graphs; survival curves) (see figures 6.1, 6.3, and 6.5);

3. Demonstrate an ability to critically appraise and incorporate research findings with particular reference to measurement issues (validity, sensitivity, specificity, positive predictive value, negative predictive value; reliability);

4. Apply the principles of epidemiology by accurately discussing the implications of the measures.

Case Study

Dr. Rao, the Richards family's physician, saw some statistics about Goosefoot in the regional public health department's "Physician Update" pamphlet. There is demographic information on the age and sex breakdown of the population and figures on average income, unemployment, and educational attainment. There is also information on health habits and hospital admission rates, death rates, and consultation rates. He is looking at the page showing the following information about the local area and the Weenigo region:

Table 6.1 Health indicators for Goosefoot community, compared to the region

Major health indicators	Goosefoot	Weenigo region
Number of deaths, 3 year average	132	9,829
Mortality rate (per 100,000), 3 year average	884	808
Standardized mortality rate (per 100,000), 3 year average	670	751
Life expectancy at birth, men (years)	76.2	77.5
Remaining life expectancy at 65, men (years)	20.0	17.9
Life expectancy at birth, women (years)	80.7	82.4
Remaining life expectancy at 65, women (years)	20.1	21.7
Number of live births (2009)	140	13,981
Infant mortality rate (per 1,000 live births, 2009)	4.2	4.9
Perinatal mortality rate (per 1,000 total births, 2009)	5.5	7.0

Dr Rao scratches his head and wonders what these numbers mean and why there are so many different measures of mortality.

The scope of health measures

The medical model of health that was described in Chapter 1 defined and measured health in terms of low rates of adverse health events. Most early measures of population health reported incidence rates of "the five Ds": death, disease, disability, discomfort, or distress. Note that the five Ds form a hierarchy, from objective, numerical measures to more subjective, qualitative indicators, and also from those that are routinely collected (e.g., death certificates) to those that are available only from a research study (e.g., questions on feelings of distress).

The quality of health data varies. Indicators based on death rates are robust and almost complete because death certificates must be completed by law (although the accuracy of the diagnosis may be questioned). Disease records can also be reasonably complete: diagnoses can be taken from hospital discharge summaries, for example, although such data can only be generalized to people treated in hospital. The quality of statistics based on disease records depends on the accuracy with which physicians completed the original forms, but they are relatively accurate. Because of their availability and comparability, mortality and morbidity statistics are used by national and international agencies to compare health status between countries. Some of the commonest statistics for this purpose include death rates per thousand, infant mortality rates, average life expectancy, or a range of morbidity indicators, such as rates of reportable disease. The results can be further analyzed within a region to compare the health of different groups of people, or to track particular health problems such as influenza or HIV/AIDS.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

The dynamic evolution of health indicators

The indicators we select to identify health issues are not static but evolve over time. Agencies commonly collect information on issues that are of interest and concern to them, and the act of measurement focuses attention on that topic. This might be a health problem, such as infant mortality. The resulting interventions will, if successful, tend to reduce the acuity of that particular problem. This, in turn, reduces the value of that indicator as a marker of current health issues. Meanwhile, the emergence of new concerns motivates the collection of information on *those* problems to monitor progress toward new goals, and so the cycle begins again.[1] For example, as infant mortality declines, growing numbers of those who survive exhibit health problems associated with low birth weight or prematurity?problems rarely encountered in populations with high infant mortality. This refocuses attention towards indicators of disability and quality of life, rather than mortality. This whole evolution also links to the evolving definitions of health that were described in Chapter 1.

Individual measures and population-level indicators

Health measures can record information on individuals or on whole populations. Beginning with the latter, the most common measures of population health, mortality or morbidity rates, are derived from counts of individuals, which are then aggregated up to a population level (prevalence and incidence rates are examples). Morgenstern termed these *aggregate measures*, which he distinguished from *environmental* and *global* measures of population health.[2] Morgenstern's environmental indicators record factors external to the individual, such as air or water quality. These indicators can be recorded at the individual or at the ecological level: contrast an individual dosimeter with area air-quality monitoring. In health research, environmental indicators are often used to record factors that affect individual health, but they can also be used to indicate the health consciousness of a society. A society with high levels of air pollution, for example, is likely to experience higher rates of respiratory disease but may tolerate this if it has conflicting priorities, such as building an industrial economy. The distinguishing characteristic of Morgenstern's third category, of global health indicators, is that they have no obvious analogue at the individual level; examples include the existence of healthy public policy, including laws restricting smoking in public places, or social equity in access to care. These characteristics may influence the health of individuals, but can also be viewed as ways to summarize the healthiness of a population: Is this a caring society? Does it try to protect the health of its citizens? The contrast between aggregate and global measures corresponds to the distinction between health *in* the population and health *of* the population that was introduced in Chapter 1 (Link to Nerd's corner Box on Page 18 in chapter 1 "Alternative conceptions?"then delete this reference).

Prevalence and incidence

These are both aggregate indicators of morbidity. Incidence (see **Incidence rate** in Glossary) is a measure of the speed at which new events (such as deaths or cases of disease) arise in a population during a fixed time. It can be expressed either as a proportion or as a rate. The incidence rate is calculated as the number of new cases arising in a susceptible population divided by the person-time of observation. The mortality rate in Goosefoot is 884 per 100,000 per year, or in other words 884 per 100,000 person-years of observation. The incidence proportion is calculated by dividing the number of new cases or events in a fixed time period by the number of people at risk at the beginning of that period. Usually, the period reported is one year, producing the annual incidence. This gives a proportion (ranging from 0 to 1) that is useful in communicating the idea of risk and provides an answer to the question: What is the probability that a person will get this disease within this time frame? Note that you may also see the term *cumulative incidence* instead of incidence proportion, and the result for Goosefoot is 0.00884 per calendar year.

There are two difficulties with the incidence proportion. First, it tends to be biased downward. Those who are lost to follow-up will tend to reduce the numerator as we may not know if they actually got the disease; they may also inflate the denominator if we do not know that they have moved. Likewise, those who die of other causes tend to inflate the denominator, because we don't know if they would have got the disease had they lived. Second, many diseases can recur, and if the recurrences are included, the incidence proportion could exceed 1.0, which seems inappropriate. On the other hand, accepting only a single occurrence for each person can underestimate the true burden of disease. Therefore, an alternative approach is to express incidence as a rate: the number of cases per time of observation, typically the number of cases per 100 or 1,000 person-years of observation. Incidence rates are equivalent to recording speed in km per hour and, like speed, the incidence rate gives an instantaneous reading of the frequency with which the disease is occurring, or the expected time-delay until the next case. A graphical illustration of the difference between incidence rates and proportions is shown later in the chapter, in Figure 6.5. A problem common to both methods of presenting incidence is the potential for an underestimate of the numerator, due to subclinical and unreported cases. Incidence is useful for studying causes, while prevalence is useful when estimating health services needs. Therefore we assess the incidence of road traffic injuries under different conditions when looking for ways to prevent them, while we assess the prevalence of long-term disability from road traffic injuries when planning rehabilitation services.

While incidence measures events, **prevalence** is a measure of disease state; it counts all existing cases in a fixed time period (including those that began before the observation period), divided by the population size. It is, therefore, influenced by both the incidence and by the duration of illness, which is linked to survival. The time period for calculating prevalence is commonly a single point in time (point prevalence). Alternatively, prevalence can be calculated for a given period such as a year (period prevalence). Prevalence is generally the measure of choice in assessing the burden of a chronic disease because new cases may be quite rare, and existing cases can last a long time, requiring care and causing significant disability.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Illustrative Materials

Hospitals: where prevalence is high and incidence is low

Patients in the cardiac care unit have a prevalence of cardiac disease of nearly 100%?the few misdiagnoses may reduce the prevalence slightly. However, the incidence of cardiac disease in these patients is extremely low (unless we include recurrences); only those people who have been misdiagnosed are at risk of getting cardiac disease?all the others already have it.

Here Be Dragons

Morbid statistics

Morbidity statistics derived from hospital discharge data or from physician billing data are often used to define trends or compare regions. However, these statistics often reflect the amount of service being offered more than the patient's need for service. For instance, high rates of coronary artery stenting in one region compared to others probably means that access to cardiology services in that region is better.

Mortality rates

Like incidence, mortality is an event. There are various forms of mortality rate, beginning with those that refer to particular age groups, such as child mortality rates, and proceeding to the overall crude mortality (or death) rate. Because the death of a child represents the most significant loss of potential life, and also because child health is sensitive to variations in the social environment and in health care quality, there are several indicators of child mortality.

Infant mortality rate (IMR)

The **infant mortality rate** is the total number of deaths in a given year of children less than one year old, divided by the number of live births in the same year, multiplied by 1,000. Because it measures the health of children directly and the health of mothers?a significant proportion of the adult population?indirectly, the IMR is often quoted as a useful indicator of the level of health in a community. However, because of the rarity of infant death in developed countries, it is useful only in large populations?small populations produce such small numbers of infant deaths that chance variation obscures any real variations.

$$\text{Infant mortality rate} = \frac{\text{Deaths among children < 1 year old}}{\text{Number of live births in the same year}} \times 1{,}000$$

Perinatal mortality rate (PMR)

In most industrially developed nations, this is defined (for a given year) as:

$$\text{Perinatal mortality rate} = \frac{\text{Fetal deaths (> 28 weeks) + deaths in 1st week after birth}}{\text{Fetal deaths (> 28 weeks) + live births}} \times 1{,}000$$

Neonatal mortality rate (NMR)

$$\text{Neonatal mortality rate} = \frac{\text{Deaths of infants < 28 days old}}{\text{Live births}} \times 1{,}000$$

All-age crude mortality rate

This gives an estimate of the rate at which members of a population die during a specified period. The numerator is the number of people dying during the period; the denominator is the size of the population, usually at the middle of the period (mid-year population).

$$\text{Crude mortality rate} = \frac{\text{Number of deaths during a specified period}}{\text{Number of persons multiplied by the period of observation}} \times 10n$$

(*Notes:* The ?10n' simply means that the rate may be multiplied by 1,000, or even 100,000 for rare diseases, to bring the rate to a convenient whole number. The time period is commonly a year, but the example above quoted three year averages for Goosefoot. This was done because the population is rather small, and so rates might fluctuate somewhat from year to year; the 3-year average gives a more stable picture.)

The crude death rate in Canada is between 7 and 8 per 1,000 per year, but will increase slightly as the population ages. Figure 6.1 shows the ten leading causes of death in Canada in 2005 and the rate per 100,000 for each.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

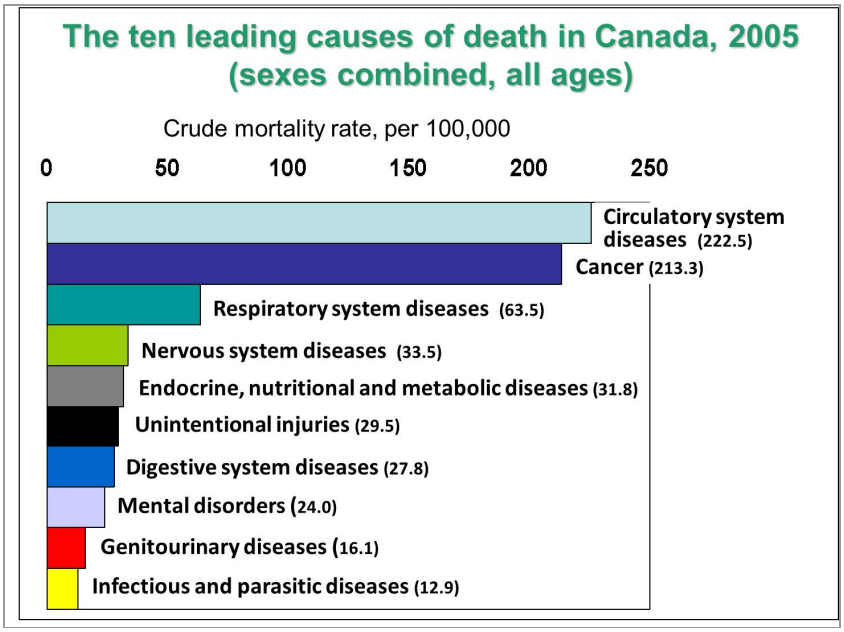


Figure 6.1 The ten leading causes of death for Canada, 2006. (Source: Statistics Canada)

The previous discussion referred to ?crude rates' in the whole population, but it is often desirable to focus on segments of the population (such as men or women), usually for purposes of comparison between them. Note that the calculations can apply to any health indicator, but here we use mortality rates as an illustration. Specific rates are based on crude mortality rates, but focus attention on a more homogeneous group defined by age, or sex, or race, etc. Rates may also be made specific for more than one characteristic of the population, such as the age-, sex-, and race-specific death rates from colon cancer.

Standardized mortality rate

Standardization is used when comparing mortality in two populations that differ in terms of characteristics known to influence mortality, and whose effect one wishes to temporarily remove. It is commonly used to remove the effect of differences in age, but can be used to reduce the effect of other confounding variables. For example, Victoria, B.C., has a higher proportion of elderly people than Whitehorse, Yukon, so a comparison of crude mortality rates will show a higher rate in Victoria than in Whitehorse simply because of Victoria's older population. To make a fair comparison of the risk of death in the two places, this **confounding** effect of the age difference can be removed by calculating age-standardized rates.

Illustrative Materials

Canadians are living longer but are also getting healthier, and this is revealed by the age-standardized mortality rate. While the crude death rate has been static at around 7 to 8 per 1,000 since 1970, the age-standardized rate has dropped from nearly 10/1000 in 1970 to just over 6/1000 in 2004 (standardization based on the 1991 population).

How would you interpret this observation?

Standardization is either indirect or direct. *Indirect standardization* is expressed as a standardized mortality ratio (SMR), which indicates the change in risk of dying in the population under study compared to a reference population (such as Canada as a whole). It is the ratio of the deaths observed in the population to the number that would be expected if this population had the same structure as the reference population. An SMR of 100 signifies that deaths are at the expected level, an SMR of 110 indicates a death rate 10% higher than expected (see Figure 6.2 for an illustration).

Direct standardization is expressed as a rate: *x* number of deaths per *y* number of individuals, but this rate is fictitious and is only used for comparison purposes. See Nerd's corner to find out how to calculate SMRs using direct and indirect standardization.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

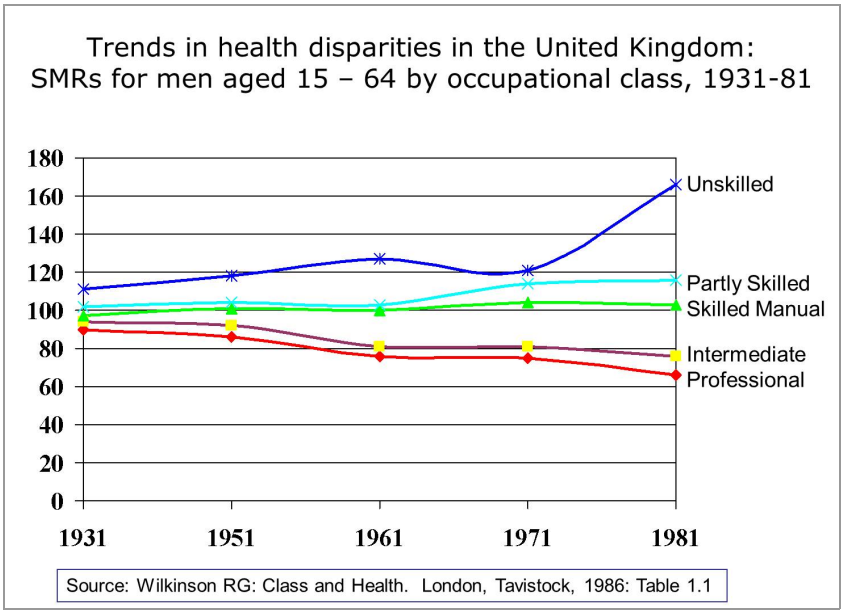


Figure 6.2: An example of indirect standardization. Note how this highlights the widening disparities in mortality rates between occupational classes in Britain during the 1970s.

Further Reading

For more on standardization, visit the Pan American Health Organization website at: http://www.paho.org/english/sha/be_v23n3-standardzation.htm

For examples of direct standardization, type "Canada surveillance online" into a search engine to find the Public Health Agency of Canada surveillance web page.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

Calculating age-standardized rates and ratios

Using data on mortality by age-group in Goosefoot and the broader Weenigo region as an example:

Direct standardization

The age-standardized mortality rate (ASMR) is calculated in 4 steps:

1. Select a reference population (usually the country as a whole) and find out from the census how many people there are in each age group (usually 5, 10 or 20-year age groups) and place them in one column of a spreadsheet.

2. Calculate age-specific death rates in Goosefoot and the broader region and put them in nearby columns.

3. For each age group within these locations, multiply these rates (e.g., 111 per 100,000 for children aged 0-14 in Goosefoot) by the number of people in that age group in the standard population. This will give a large number that indicates how many people would be expected to die in that place if their population was the same size as the reference population.

4. To get the overall ASMR, add up the number of expected deaths for each age group, first for Goosefoot and then for the Weenigo region. Then divide each by the total number of people in the standard population to get the age-standardized mortality rate for each place:

Direct Method		Goosefoot				Weenigo Region			
Age group	Canadian population (2006)	population	# deaths	crude rate / 100,000	expected deaths	population	# deaths	crude rate / 100,000	expected deaths
0-14	5,579,835	897	1	111	6221	124,813	95	76	4247
15-64	21,679,805	10,412	55	528	114521	910,602	3,506	385	83472
65+	4,335,255	3,625	76	2097	90891	180,453	6,228	3451	149623
total	31,594,895	14,934	132	884	211632	1,215,868	9,829	808	237342
		Age-standardized rate			670				751

Here, we see that although the crude death rate in Goosefoot was higher than that for the Weenigo region (884 vs. 808), the age-standardized rate is lower. This arises because there are more elderly people in Goosefoot; correcting for this indicates that Goosefoot is actually relatively healthy: their higher crude death rate was due to their being relatively old.

Note that the ASMRs are artificial, and have no meaning in isolation: they have meaning only when compared to the crude death rate in the standard population or to the ASMRs from other groups, calculated using the same weights. Another way to think about the ASMR is to view it as a weighted average of the age-specific rates for a place, with the weights being the proportion of the reference population that falls within each age-group.

Indirect standardization

This time we begin from the age-specific mortality rates from the standard population (here taken from Statistics Canada data).

1. Multiply them by the number of people in Goosefoot in each age group, and then do the same for Weenigo, to show the number of deaths that would be expected to occur if the national age-specific death rates applied.

2. Add the expected figures up to obtain the total expected deaths in each place.

3. Divide the total expected number of deaths into the observed number of deaths to obtain the age-standardized mortality ratio for each place. A value of 1 indicates that that location is experiencing the same age-adjusted mortality as the standard population. (This is commonly expressed as a percentage to give a base SMR of 100).

Indirect Method	Canada (2006)		Goosefoot			Weenigo Region		
Age group	population	mortality rate/100,000	population	observed # deaths	expected # deaths	population	observed # deaths	expected # deaths
0-14	5,579,835	47.22	897	1	0.42	124,813	95	58.94
15-64	21,679,805	230.60	10,412	55	24.01	910,602	3,506	2,099.86
65+	4,335,255	4,211.59	3,625	76	152.67	180,453	6,228	7,599.93
total	31,594,895		14,934	132	177.10	1,215,868	9,829	9,758.74
				SMR	0.75		SMR	1.01

Again, the figures indicate that Goosefoot is relatively healthy compared to the region. Note: SMRs can be compared to 1, but cannot be compared to each other, since they all use different weights.

Case Study

Dr Rao has been pondering the meaning of the Goosefoot standardized mortality rates for a while now, and wonders why the standardized rate is so much lower than the crude rate (670 versus 884). It is his wife who, in the end, suggests that maybe this is because there are mostly older people in Goosefoot: most of the younger folks have moved to Weenigo to find work. She points out that of course this will mean more deaths per thousand than in a younger population.

Dr Rao begins to feel somewhat relieved: standardizing the rates gives a more comparable result, and in fact, when the effect of age is removed Goosefoot is actually doing better than the region as a whole (670 versus 751 deaths per 100,000).

Life expectancy

Life expectancy at birth is an estimate of the expected number of years to be lived by a newborn based on age-specific mortality rates at a selected time. Life expectancy is a statistical abstraction, usually based on current age-specific death rates (after all, we will have to wait a lifetime to find out

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

how long babies born today will actually live). It is used as a summary indicator of health that can be compared across countries. In Canada in 2010, life expectancy is almost 83 years for females and 78 years for males, which places us very close to Australia and just behind Japan. The impact of changing social conditions on life expectancy is illustrated by the case of Russia during the 1990s: life expectancy can fall surprisingly quickly if conditions deteriorate, as shown in Figure 6.3.

Illustrative Materials

Life expectancy in post Soviet era Russia

The social disruption of post-Soviet Russia was reflected in rapidly rising death rates, so that between 1990 and 1994 the life expectancy for women declined from 74.4 years to 71.2, while that for men dropped by six years from 63.8 to 57.7 (Figure 6.3). Cardiovascular disease and injuries together accounted for 65% of the decline.[3]

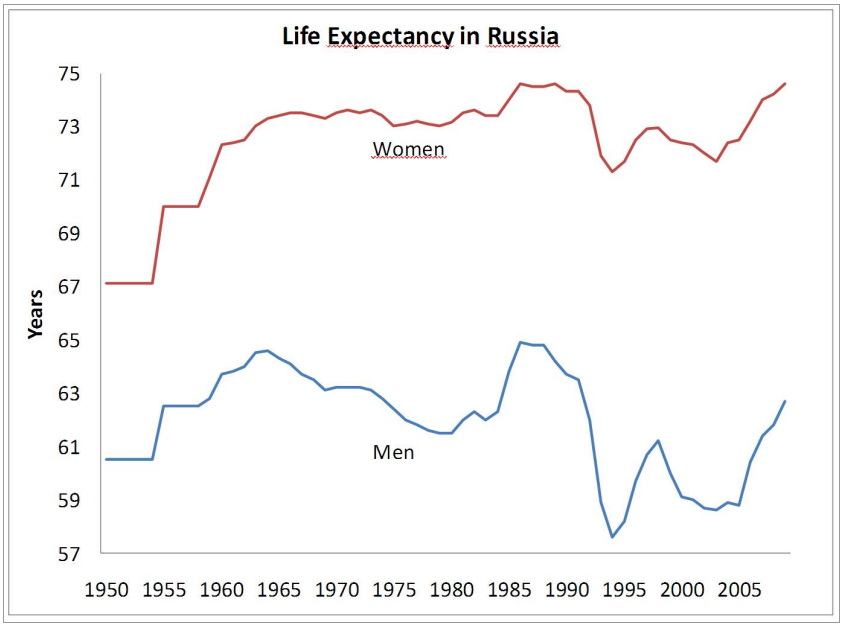


Figure 6.3: Time trends in life expectancy in Russia, 1950 ? 2007.

Source: http://en.wikipedia.org/wiki/File:Russian_male_and_female_life_expectancy.PNG accessed July 2010.

Case Study

Dr Rao ponders the curious figures for life expectancy in Goosefoot:

	Goosefoot	Weenigo region
Life expectancy at birth, men (years)	76.2	77.5
Remaining life expectancy at 65, men (years)	20.0	17.9
Life expectancy at birth, women (years)	80.7	82.4
Remaining life expectancy at 65, women (years)	20.1	21.7

He knows, of course, that women live longer than men, and the Goosefoot men live less long than those in the region as a whole. But he is surprised that the men who survive to 65 then live longer than the women. "They must be tough old men!" he says to himself, and thinks of the miners he knows. Those who do not have chronic lung disease are, indeed, strong outdoors people who are active hunting and fishing. "Perhaps if they do survive to that age they may live longer than men in the city," he muses. The tendency for longer survival among the hardy few is termed the healthy survivor effect.

Potential years of life lost

The impact disease has on a population is an important factor in guiding decisions about which diseases to try and prevent. An obvious measure of impact is the number of deaths a disease causes. In Canada, this measure identifies heart disease, cancer, and stroke as the diseases with the greatest impact. However, these conditions tend to kill people who are reaching the end of their expected life span, so preventing them might have little effect on extending overall life expectancy. Preventing premature deaths would add more years of life (and perhaps also more productive years) to individuals and to society. Premature death can be defined in terms of deaths occurring before the average life expectancy for a person of that sex, or an arbitrary value, such as75 years. With this idea, people who die prematurely lose years of potential life, and one way to judge diseases in terms of their impact on a society is to rank them in terms of the total **potential years of life lost** (PYLL) due to each. (You will sometimes see the abbreviation YPLL for ?years of potential life lost': same thing.) Thus, a person who dies from a myocardial infarction at age 55 loses 20 years of

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

potential life. These values can be summed over a population to indicate the impact (in terms of potential years of life lost) due to different causes. Priorities based on the PYLL will be different to priorities based on simple mortality rates. Figure 6.4 shows that if PYLL is used as a measure of impact, cancer and circulatory disease remain the major priorities, but injuries (unintentional and intentional) occupy third and fourth places: compare this with Figure 6.1. Indeed, taken together, suicides and unintentional injuries count more years of potential life lost than circulatory disease.

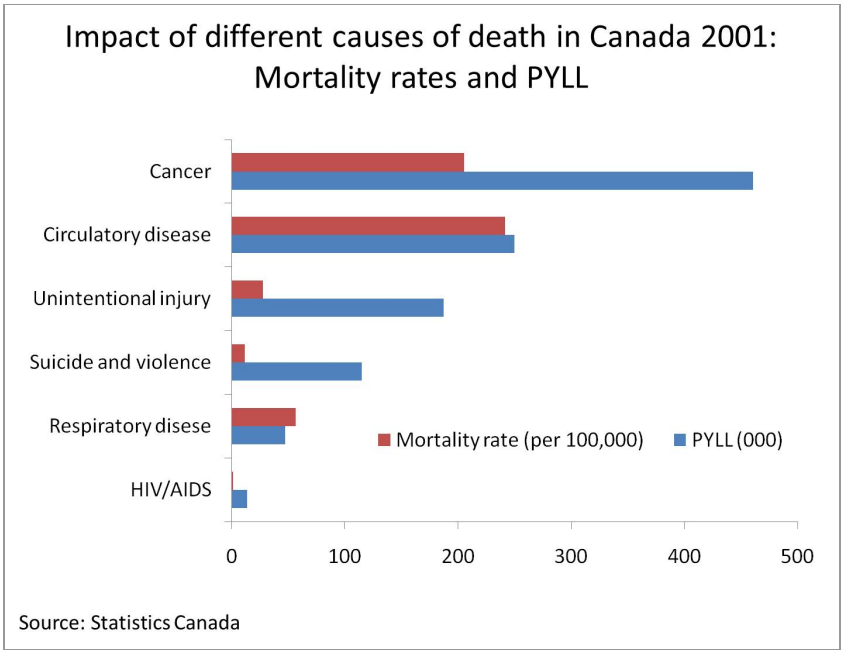


Figure 6.4: Different indicators of disease impact on mortality. Note, the PYLL values in the chart must be multiplied by 1,000 to indicate the years of life lost.

Kaplan-Meier, or survival, curves

In cohort studies, outcomes may be expressed as symptom-free survival (how long before symptoms return after a treatment) or survival (how long before death after a diagnosis), hence the term **survival curves** (see **Survival curve** in Glossary). Kaplan and Meier developed statistical tests for evaluating the difference between two survival curves. This kind of survival analysis is common in the clinical literature, as it has a number of advantages. It gives a full picture of the clinical course of a disease in terms of survival rates at specified intervals after diagnosis and/or treatment. Figure 6.5 shows a hypothetical example. Although the outcomes of the two treatments at one year are similar, Treatment A prolongs survival over the first few months after treatment. Survival curves can also be mathematically adjusted to correct for loss to follow-up, and for different follow-up periods.

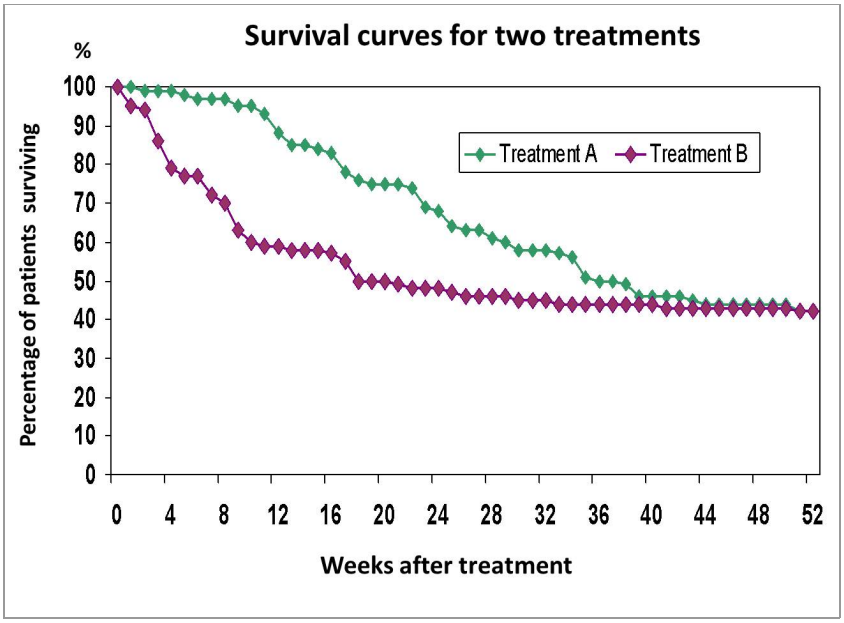


Figure 6.5: Example of a cohort study analyzed with survival curves

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

Survival curves and incidence rates

In Figure 6.5, note how the survival curves are related to incidence?of death, in this case. The incidence proportion, measured over the full 53 weeks is the same for both treatments. However, the incidence rate per person-week of follow-up is lower for Treatment A than for Treatment B, because fewer patients on Treatment A died during the first six months of follow-up.

Health measurement scales

There are evident limitations to morbidity and mortality indicators. They only apply to serious conditions, so are largely irrelevant to most people. Furthermore, a diagnosis that becomes a morbidity statistic says little about a person's actual level of function, and morbidity indicators cannot cover positive aspects of health. This led to the development of a range of more subjective indicators of health which are termed health measurement scales. Measuring health, however, is inherently challenging, because health is an abstract concept. Unlike morbidity, it is not defined in terms of specific indicators that can be used as measurement metrics.

Nerd's Corner

Measuring abstract concepts

Because health is an abstract concept, it cannot be measured directly using a mechanical scale, as weight or length is measured. Instead, *indicators* of health have to be chosen, and some form of numerical rating applied to quantify or ?scale' these. For example, if we define health in terms of physical, mental and social well-being, indicators of each of these will be selected and a scoring system devised for rating a person on each indicator. If desired, a second scoring system can then be used to represent the relative importance of the physical, mental, and social components in forming an aggregate rating. This allows the resulting health measure to be presented as a single number, called a health index, which is generally scaled to run from 0 (death) to 1 (perfect health). Alternatively, component scores may be presented separately, to produce a health profile.

Measurement scales have been developed for most common diagnoses, and these are termed disease-specific scales. Some disease-specific scales rate the severity of symptoms in a particular organ system (e.g., vision loss, breathlessness, limb weakness); others focus on a diagnosis, such as anxiety or depression scales. Other measurements are broader in scope, covering syndromes (e.g., emotional well-being scales) or overall health and health-related quality of life. These are termed generic scales, as they can apply to any type of disease and to anyone; a common example is the Short-Form-36 Health Survey, a 36-item summary of functional health. On an even broader level, some measures seek to capture the well-being of populations, such as the Canadian Index of Wellbeing, of which health forms a significant component.

The applications of health measures fall into three broad categories. *Diagnostic instruments* collect a wide variety of information from self-reports and clinical ratings, and process these using algorithms to suggest a diagnosis. There are many in psychiatry, such as the Composite International Diagnostic Interview.[4] *Prognostic measures* include screening tests and sometimes information on risk factors, and these may be combined to estimate future health states; the Lifescan Health Risk Appraisal is one example. *Evaluative measures* record change in health status over time and are used to record the outcomes of care. This category forms the largest group of instruments, and includes *generic* and *disease-specific* outcome measures.

Nerd's Corner

Objective and subjective indicators

A health indicator can be recorded mechanically, as in a treadmill test, or they may derive from expert judgment, as in a physician's assessment of a symptom. Alternatively, they may be recorded via self-report, as in a patient's description of her pain. Mechanical measures collect data objectively in that they involve little or no judgment in the collection of information, although judgment may still be required in the subsequent interpretation. With subjective measures, human judgment (by clinician, patient, or both) is involved in the assessment and in its interpretation. Subjective health measurements hold several advantages: they describe the quality rather than merely the quantity of function; they cover topics such as pain, suffering, and depression, which cannot readily be recorded by physical measurements or laboratory tests; and subjective measures do not require invasive procedures or expensive equipment. The great majority of subjective health measures collect information via questionnaires: many have been extensively tested and are commonly used as outcome measures in clinical trials. Drug trials are now required to include quality of life scales, in addition to symptom- or disease-specific scales, in order to record possible adverse side effects of treatment, such as nausea, sleeplessness, etc.

Combined indicators

Because objective and subjective indicators each have advantages, they are sometimes combined. For example, in deciding whether or not to undergo chemotherapy or surgery, a cancer patient will wish to balance the expected gain in life expectancy against a judgment of the quality of the prolonged life (considering side effects of treatment, pain, residual disability). At a societal level, this helps to address the question of whether extending life expectancy (e.g., by life-saving therapies) may also increase the number of sick and disabled people in society. This question led to the development of combined mortality and quality of life indicators such as quality-adjusted life years (QALYs). (See squaring the morbidity curve, Figure 8.2).

Adjusted life years: QALYs, DALYs, and HALYs

Quality-Adjusted Life Years (QALYs) extend the idea of life expectancy by incorporating an indicator of the quality of life among survivors. Rather

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

than count every year of life lived as though they were equivalent, this statistic adjusts the value of years lived in a state of ill-health downwards: they are counted as being worth less than a year of healthy life. In evaluating a therapy, QALYs count the average number of additional years of life gained from an intervention, multiplied by a judgment of the quality of life in each of those years. For example, a person might be placed on hypertensive therapy for 30 years, which prolongs his life by 10 years but at a slightly reduced quality level, owing to growing activity restrictions. A subjective weight is given to indicate the quality or utility of a year of life with that reduced quality (say, a value of 0.9 compared to a healthy year valued at 1.0). In addition, the need for continued drug therapy over the 30 years slightly reduces his quality of life by, say, 0.03. Hence, the QALYs gained from the therapy would be 10 years x 0.9 ? 30 years x 0.03 = 8.1 years.

The idea of QALYs can also be used to indicate the burden of an illness. For example, a year lived following a disabling stroke may be judged to be worth 0.8 normal years. Imagine Bob and Bill, twins. Bob gets a stroke at 55 and dies from the sequelae ten years later, on his 65th birthday. On the other hand, Bill was completely healthy but died after being knocked over by a bus while jogging on his 65th birthday, so that (somewhat tragically) they had exactly the same lifespan. However, if the last 10 years of Bob's life are quality adjusted, he lost the equivalent of 2 years of it from his stroke (10*0.8 = 2), so his quality adjusted life was 2 years shorter than Bill's. By summing QALY statistics such as this value of 2 years over a population, the societal burden of stroke can be estimated.

Nerd's Corner

QALY weights

The numerical weights assigned to represent the severity of disabilities are known as utility scores. They range from 0 (death) to 1.0, which represents the best possible health state. Utilities are obtained via studies in which patients, professionals or members of the public use numerical rating methods to express their preferences for alternative outcomes, including their judgment of the severity of various levels of impairment. Common rating methods include the ?standard gamble' and the ?time trade-off'.

The ?standard gamble' involves asking subjects to choose between (i) living in the state being rated (which is less than ideal) for the rest of one's life, versus (ii) taking a gamble on a treatment (such as surgery) that has the probability *p* of producing a cure, but also carries the risk 1-*p* of operative mortality. In making the judgment, the risk of death is varied until the person making the rating has no clear preference for option (i) or (ii). This shows how great a risk of operative mortality he or she would tolerate to avoid remaining in the condition described in the first option. This risk is used as an indicator of the perceived ?utility' (i.e. severity) of that condition. In principle, the more severe the rater's assessment of the condition, the greater the risk of dying in the operation (perhaps five, even ten percent) they would accept to escape the condition. The ?time tradeoff' offers an alternative way to present the standard gamble. As before, it asks raters to imagine that they are suffering from the condition whose severity is to be rated. They are asked to choose between remaining in that state for the rest of their natural lifespan (e.g., 30 years for a 40 year-old person), or returning to perfect health for fewer years. The number of years of life expectancy they would sacrifice to regain full health indicates how severely they rate the condition. The utility for the person with 30 years of life expectancy would be given as Utility = (30 - Years traded)/30.

An example of utility scaling methods used is given by the Canadian Health Utilities Index. For example, being unable to see at all receives a utility score of 0.61; being cognitively impaired?as with Alzheimer's disease?receives a score of 0.42.[5] Note that these utility judgments are subjective and may vary from population to population (a fascinating topic in itself).

Some measurement procedures allow patients themselves to supply the utility weights. This may be helpful for clinicians in helping a patient to decide whether or not to undergo a therapy that carries a risk of side-effects. An instrument of this type is the Quality-Adjusted Time without Symptoms and Toxicity.[1]

Disability-Adjusted Life Years (DALYs) and Health-Adjusted Life Years (HALYs) work in a very similar manner to QALYs. DALYs focus on the negative impact of disabilities in forming a weighting for adjusting life years, and HALYs base their valuation on the positive impact of good health. The approach of QALYs, DALYs, and HALYs can also be used to adjust estimates of life expectancy, taking account of quality of life, disability, and health respectively?the last giving rise to the acronym HALE, for Health Adjusted Life Expectancy.

Further Reading

Several guidebooks describe how to choose and interpret health measurements:

1. Spilker B, ed. Quality of life assessment in clinical trials. New York, NY: Raven Press, 1990

2. McDowell I. Measuring health: a guide to rating scales and questionnaires. New York, NY: OxfordUniversity Press, 2006

3. Bowling A. Measuring disease: a review of disease-specific quality of life measurement scales. Buckingham, England: Open University Press, 1995.

A good technical guide on the process of measuring health is:

1. Streiner DL, Norman GR. Health measurement scales: a practical guide to their development and use. New York, NY: Oxford University Press, 1995.

Reliability and validity of health measures

All health indicators, measurements, and clinical tests contain some element of error. There are three chief sources of measurement error: in the thing being measured (my weight tends to fluctuate, so it's difficult to get an accurate picture of it); in the observer (if you ask me my weight on a Monday, I may knock a pound off if I binged on my mother-in-law's cooking over the weekend obviously the extra pound doesn't reflect my true weight!); or in the recording device (the clinic's weigh scale has been acting up?we really should get it fixed).

As with sampling, both random and systematic errors may occur. Random errors are like noise in the system: they have an inconsistent effect. If large numbers of observations are made, random errors should average to zero, because some readings overestimate and some underestimate. They can occur for lots of reasons: a buzzing fly distracted Dr. Rao when he took Julie's blood pressure; you can't really recall how bad your pain

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

was last Tuesday, and so on. Random errors are detected by testing the **reliability** of a measurement.

Systematic errors fall in a particular direction and are likely due to a specific cause. Errors that fall in one direction (I do tend to exaggerate my athletic prowess?) bias or distort a measurement and are considered part of its validity. These distinctions are illustrated in Figure 6.6, using the metaphor of a shooting target: a wide dispersion of bullets indicates unreliability, whereas off-centre shooting indicates bias or poor validity.

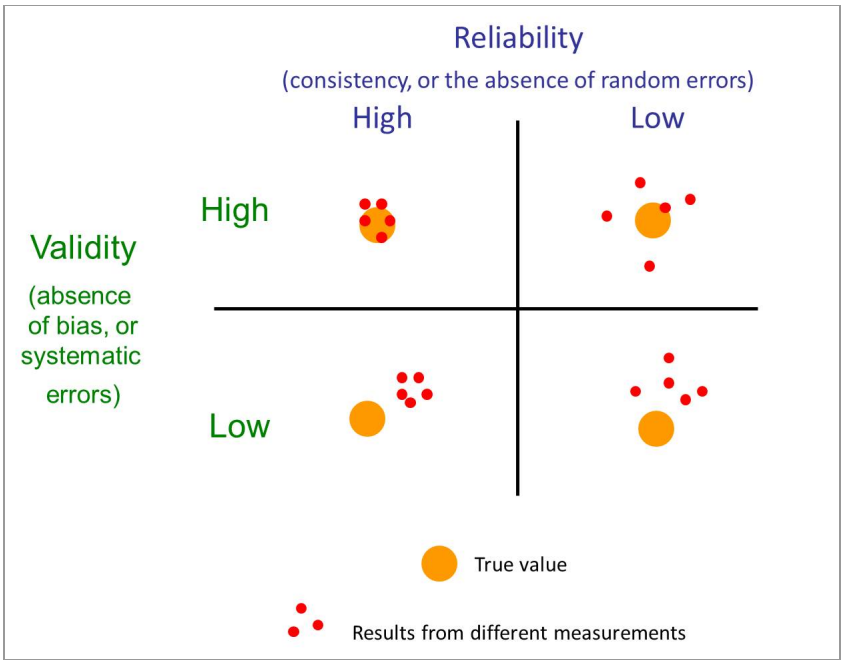


Figure 6.6 Metaphor for potential errors in the reliability and validity of a measurement, compared to the true value being estimated.

Reliability

Reliability refers to dependability or consistency. Your patient, Jim, is unpredictable: sometimes he comes to his appointment on time and sometimes he's late, but once or twice he was actually early. Jim is not very reliable. Jack, on the other hand, arrives exactly 10 minutes early every time. Even though he comes at the wrong time, Jack is reliable. A reliable measure can be very reproducible, but it can still be wrong. If so, it is reliable but not valid (top left cell in Figure 6.6).

Validity

An introductory definition of validity is: Does the test measure what we are intending to measure? A slightly more abstract definition is: How closely do the results of a measurement correspond to the true state of the phenomenon being measured? A more sophisticated definition is: What does a given score on this test mean? This last interpretation of validity fits under a more general conception in terms of: How can we interpret test results?

There is no single approach to measuring validity: the approach varies according to the purpose of the measurement and the sources of measurement error you wish to detect. In medicine, the commonest way to assess validity is to compare the measurement with a more extensive clinical or pathological examination of the patient. This is called criterion validation (see **Criterion validity** in Glossary), because it compares the measurement to a full work-up that is considered a "gold standard" criterion. For example, results of fecal occult blood testing as a screen for colon cancer can be compared to colonoscopy for a sample of people, including some with and some without, the disease. This is used where the measurement offers a brief and simple way to assess the patient's condition, and our question is: How well does this simple method predict the results of a full and detailed (and also expensive, perhaps invasive) examination?

Interpreting tests on individuals

Table 6.2 outlines a standard 2 x 2 table as the basis for calculating the criterion validity of a test. A population of N patients has been tested for a given disease with the new test (shown in the rows), and each person has also been given a full "gold standard" diagnostic work-up, shown in the columns. (This is the theoretical "gold standard" that we are assuming is correct; of course in reality gold standards may not be as golden as we would like). Several statistics can be calculated to show the validity of the screening test.

Table 6.2: Relationship of test results to the presence of disease

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Test result	Disease Present	Absent	Totals
Positive	a True Positives (TP)	b False Positives (FP)	a+b
Negative	c False negatives (FN)	d True negatives (TN)	c+d
Totals	a+c	b+d	N

Sensitivity

Sensitivity summarizes how well the test detects disease. It is the probability that a person who has the disease will be identified by the test as having the disease (the term makes sense: the test is sensitive to the disease, so can detect it). Using the notation in the table:

a / (a+c), or TP / (TP + FN).

The inverse of sensitivity is the false negative rate(c/a+c), which expresses the likelihood of missing cases of disease. A test with a low sensitivity will produce a large number of false negative results.

Some mnemonics may help you: Se**N**sitivity is inversely associated with the false **N**egative rate of a test (high sensitivity = few false negatives). Low se**N**sitivity leads to low **N**egative predictive value.

Specificity

Specificity measures how well the test identifies those who do not have this disease:

d / (b+d), or TN / (TN + FP).

Specificity is the inverse of the false positive rate (b/b+d): the likelihood of people without the disease being labelled as having it. A highly specific test will produce very few false positive results. Again, the name makes sense: a specific test produces few false positives, so detects only this disease, i.e. is specific to it.

Specificity is clinically important as a false positive result can cause worry, lead to the expense of unnecessary further investigation, and perhaps even treatment that was unnecessary.

Some mnemonics to help you: S**P**ecificity is inversely associated with the rate of false **P**ositives. Again, the term is intuitive: a specific screening test is one that detects only the disease it is specifically meant to detect; hence, it will not give people with other conditions false positive scores. Low s**P**ecificity leads to low **P**ositive predictive value.

Positive and negative predictive values

Sensitivity and specificity are inherent properties of a test, and are useful in describing its expected performance. But they can only be measured if the actual disease status of individuals undergoing the test is known. Naturally, when we apply a test in normal clinical practice we do not know who has the disease; we are using the test to help find out. Therefore, we are more interested in what a negative or positive test result means for the patient. For this, we use the predictive values.

The positive predictive value (PPV) shows what fraction of patients who receive a positive test result actually have the disease:

a/(a+b), or TP/(TP+FP).

You can see from Table 6.2 that a test with low specificity (i.e. lots of false positives) will have a low PPV. Correspondingly, the negative predictive value (NPV) shows how many people who receive a negative score really do not have the condition:

d/(c+d), or TN/ (TN+FN).

If the test has low sensitivity, its NPV will be reduced.

Predictive values and prevalence

In an important further complication of interpreting test scores, predictive values vary according to the prevalence of the disease in the population in which the test is used. Clinicians must bear this in mind when interpreting a test result: you must treat the patient, not the test result! The reasons are illustrated in Figure 6.7, which contrasts the performance of the same test in high and low prevalence settings.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

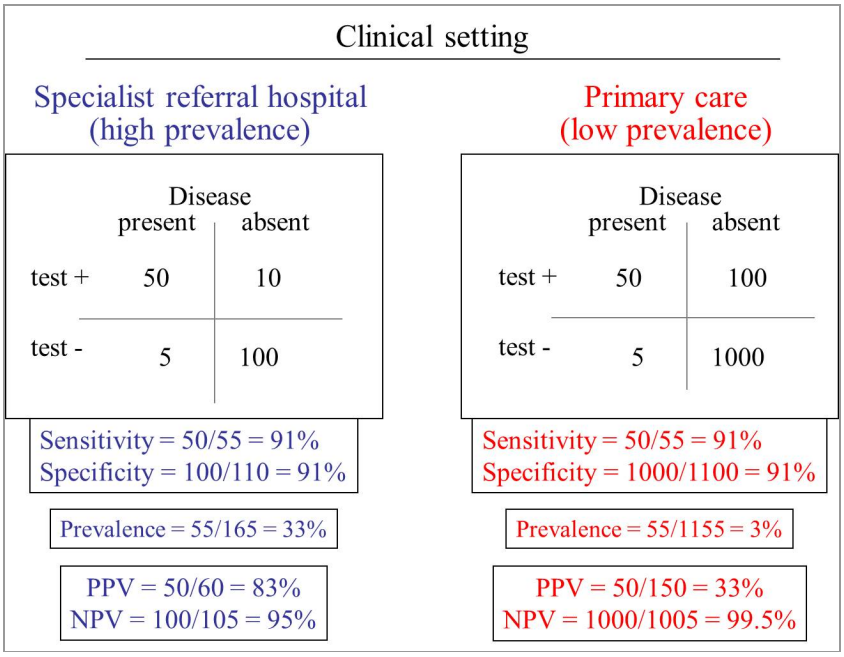


Figure 6.7: The impact of prevalence on predictive values of a diagnostic test, comparing a high prevalence setting (specialist referral hospital) with a low prevalence setting (primary care)

Tests are often validated in hospital settings, where the prevalence of the disease being tested for is high, similar to that shown in the left column of the figure. However, the test may then be used in primary care settings, where the disease prevalence is lower, as shown in the right column. Note that the sensitivity and specificity of the test remain the same in both settings, but the predictive value of a positive test is very different. This is simply because there are many fewer cases to be identified and many more non-cases in the primary care setting. So, if specificity is not extremely high, the number of false positives can exceed the true positives. At the same time, lower prevalence means that a negative test result is more accurate: you can reassure your patient with a negative score that he is very unlikely to have the disease (you will, of course, remind him to come back for re-evaluation if his symptoms continue: he may be one of the few patients with a false negative result).

In summary, interpreting test results requires insight into the population on which you are applying the test. It can be misleading to apply screening or diagnostic tests in low-prevalence settings: you may find many false positive results. One strategy to improve the positive predictive value of a test is to change from screening everyone (universal screening) to screening selectively. For example, test only people with a high risk of the illness? those with risk factors, a family history, or symptoms that suggest the disease, among whom the disease will be more prevalent. Population screening programmes can be problematic. For instance, in breast cancer screening programmes, the positive predictive value of a positive mammogram is only around 10%, so for every 100 women who are recalled for further investigation after an abnormal screening mammogram, 90 will not have cancer.[6]

Taking this one step further, the more tests you administer to a patient, for example in an annual physical exam, the more likely you are to get a false positive (and therefore misleading) score on one of them. Hence, tests should be chosen carefully and applied in a particular sequence to rule in, or rule out, a specific diagnosis that you have in mind. If the goal is to rule alternative diagnoses out, then several tests can be run together, to increase sensitivity for detecting rival diagnoses. If the goal is to rule a diagnosis in, the tests can be administered serially, stopping when a positive result is obtained. Each test should be chosen based on the conclusion you draw from the results of the previous test. Remember that unnecessary tests are not only costly to the healthcare system, but also ethically questionable if they expose a patient to unnecessary risks such as radiation from an X-ray or unnecessary treatment after a false positive result.

Ruling in, and ruling out, a diagnosis

Clinicians frequently use test results to rule in, or to rule out, a possible diagnosis, but the logic of this is also often misunderstood. To rule a diagnosis in, you need a test that is high in specificity; to rule a diagnosis out you need a test that has high sensitivity. This may sound counter-intuitive, so we will explore it further, beginning with ruling a diagnosis out. A perfectly sensitive test will identify all cases of a disease, so if you get a negative result on a sensitive test, you are fairly sure that the patient does *not* have this disease. The mnemonic is ?**SnNout**': **sen**sitive test, **N**egative result rules **out**. Conversely, to rule a diagnosis in, you need a positive result on a specific test, because a specific test would only identify this type of disease. The mnemonic is ?**SpPin**' ? **sp**ecific test + **p**ositive score rules **in**. Note, unfortunately, that if the patient gets a negative score on this test, they may still have the disease (i.e., they had a false negative result). The use of tests to enhance the likelihood of a diagnosis introduces the idea of likelihood ratios.

Here Be Dragons

You may occasionally hear someone argue that you need a sensitive test to rule in a diagnosis in, but this is false.

The reason is that the sensitive test will indeed identify most of the true cases of the disease, but a highly sensitive test will often have lower specificity. This means that there may be a number of *false* positives mixed in, so the sensitive test cannot rule a disease in.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Likelihood Ratios

A **likelihood ratio** combines sensitivity and specificity into a single figure that indicates by how much the test result will reduce the uncertainty of a given diagnosis. The **likelihood ratio** is the probability that a given test result would occur in a person with the target disorder divided by the probability that the same result would occur in a person without the disorder.

A positive likelihood ratio (or LR+) indicates how much more likely a person with the disease is to have a positive test result than a person without the disease.

LR+= sensitivity / (1 ? specificity),

A negative likelihood ratio (or LR?) indicates how much more likely a person without the disease is to have a negative test result, compared to a person with the disease.

LR? = (1-sensitivity) / specificity,

the ratio of false negatives to true negatives. Referring to Table 6.2, the formula is

[1-(a/a+c)] / (d/b+d).

In other terms, the LR+ expresses how much a positive test increases the odds that a patient has the disease; an LR? indicates how much a negative test decreases the odds of having it.

In applying this to a clinical situation, prevalence is again taken into consideration, and the use of a nomogram (Figure 6.8) removes the need for calculation. The scale on the left of the nomogram shows the pre-test probability of disease, the central column shows the likelihood ratios, and the right-hand scale shows the post-test probability. To use nomogram, you need to know the likelihood ratio for the test and the patient's pre-test probability for having a disease. The pre-test probability can be difficult to estimate (see Nerd's corner), but a first guess is based on the incidence of the condition in the setting where you are practising. This can then be modified upwards or downwards by your initial clinical impression and history-taking for this particular patient. Using the nomogram, draw a straight line from the pre-test probability through the likelihood ratio; this will cut the right-hand scale at the level of the post-test probability that the patient has the disease. Post-test probability means the likelihood that the patient has this disease, taking into account the initial probability and the test score. Of course, if your initial observations led you to feel virtually certain the patient has the disease, a positive score on the test will not add very much new information; but a negative score would.

Using the example in Figure 6.7, test sensitivity and specificity are both 0.91 so the LR+ is 0.91/(1-0.91) = 10.1. Draw a line through the pre-test probability on the left of the diagram, through 10.1 in the central column, and then read off the post-test probability on the right-hand column. For the hospital setting, the prevalence was 33%, while it was 3% for the primary care setting; this gives us a rough estimate of pre-test probability. So in the hospital setting, a positive test would mean that a patient's post-test probability of having the disease is over 80% (blue line), whereas in the primary care setting it would be around 20% (red line). In both instances, the test result has substantially increased the clinical probability of the patient having the disease. In the hospital setting, you are now pretty certain of the diagnosis. In the primary care setting, as long as the situation is not urgent, you might want to increase your certainty before launching into possibly harmful treatment. In general, tests with positive LRs higher than about 5 are useful in ruling in a disease.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

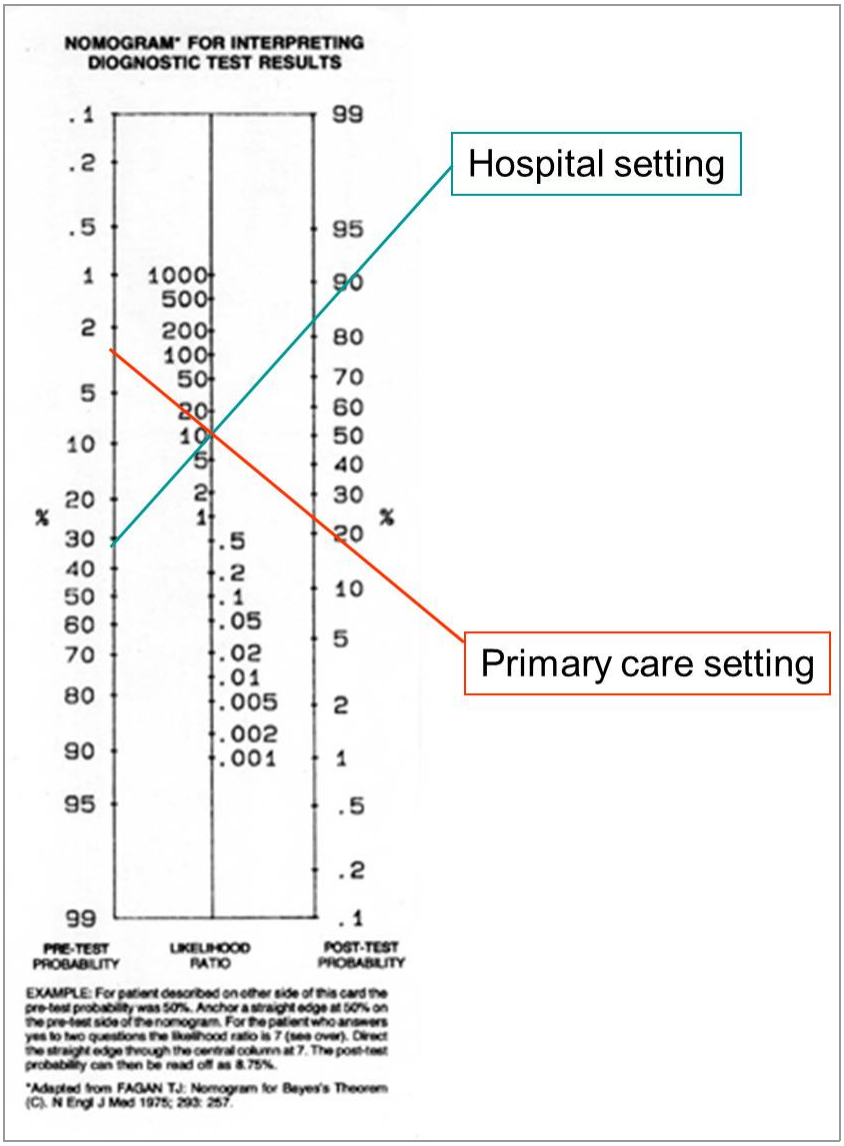


Figure 6.8: Nomogram for interpreting diagnostic test results (Sackett et al)[7]
The lines illustrate the results for LR+ described in the text.

Nerd's Corner

Pretest probability

The estimation of pretest probability takes us back to the distinction introduced in Chapter 2 between health determinants and risk factors. Determinants set the incidence rate in a population, and incidence offers a first approximation to the pretest probability of disease for any individual from that population. But we cannot apply population data directly to the individual (who is unique and unlikely to perfectly match the population average), so consideration of individual risk factors can be used to modify the crude estimate of pretest probability upwards or downwards. For example, incidence for 35 year old males might be x per thousand, but a clinician sees an individual 35 year-old man who is overweight and smokes, so his risk might be estimated as 2x per thousand. Furthermore, the pattern of signs and symptoms could raise the risk even more, to 4x per thousand, for instance. More often assessment of risk imprecisely described as high or low clinical suspicion and a diagnostic test is applied to either confirm or rule out the putative diagnosis. At this stage, a test can be used to "confirm" clinical suspicion (or high pre-test probability) However, a positive confirmatory test often indicates only a higher level of probability, but a level at which doubt must be suspended until evidence is found to the contrary. In a similar way, when a test has ruled out the disease the clinician can inform the patient that he is at very low risk, but nonetheless should report worrying symptoms. In other words, clinicians must always review and revise their diagnoses.

The likelihood ratio for a negative test result (LR-) is (1-sensitivity)/specificity, or false negatives over true negatives. It gives a result below 1: values smaller than 0.2 or so are useful in ruling out a disease. In our example, the LR- is 0.099. In the hospital setting, a patient who receives a negative test result would have a post-test probability of having the disease of around 4% (down from 33% before the test was administered). The primary care patient with a negative test would have a post-test probability of having the disease of about 0.2%?he almost certainly does not have it (a 1 in 500 chance).

Establishing cut-points: what is a normal value?

Making binary decisions is necessary in medicine, such as whether to prescribe treatment or not, to operate or not, and to reassure the patient that he does not have a disease. However, most biological measurements do not provide binary categories, but instead produce a continuous range of values, as with blood pressure, blood cholesterol, glucose, creatinine and bone density. Hence, a **cut-point** on each of these scales has to be chosen to divide the ?normal' from the ?abnormal' results. Even with qualitative assessments, such as X-rays or histology slides, decisions must be

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

made among a range of findings, which vary from definitely abnormal to definitely normal.

Defining ?normal' is not as simple as it might seem. In a superficial sense, it is defined in terms of the average, or most common, presentation for a person of that type. Unfortunately, this does not necessarily imply that it is healthy: on average, Canadians are overweight. In addition, abnormality occurs at both ends of a continuum?being underweight and being overweight are both unhealthy. Therefore, in place of the average, normal could be defined in terms of a range, perhaps defined by percentiles or by standard deviations, on the continuum being measured, such as body weight. This seems to move the notion of normal towards healthy, but setting the margins of the distribution is challenging. We cannot justify defining the normal range in terms of, say, two standard deviations above or below the mean, as this would vary from measure to measure.

A more promising approach returns us to the theme of evidence-based medicine and defines normal in terms of a range of scores above or below which treatment would be beneficial. This approach recalls the definition of need for care that was introduced in Chapter 1: abnormal is the threshold beyond which a person would benefit from treatment. An implication of this approach is that an evolution in treatments would modify the range of what is considered normal. This may be illustrated by early cognitive decline: new therapies are becoming available to treat cognitive loss earlier than before, so new layers of cognitive impairment are being defined among people who would previously have been considered ?normal', or at least tolerated as being ?a bit senile'. In a similar way, cut-points for defining hypertension have changed. Pre-hypertension was redefined in 2003 as a systolic blood pressure of 120-139 mmHg or a diastolic pressure of 80-89 mmHg.[8] By altering cut-points, more people are classified as having the disease and, therefore, become eligible for treatment. The reason for altering the cut-points has usually been because a clinical trial has shown better outcomes for this new group of patients, although the improvement may be small. Not surprisingly, the change finds favour with the drug companies that make and sell the treatments.

Finally, we may link the themes of random versus systematic error to the sensitivity and specificity of a test. In some tests, imperfect sensitivity may be due to random errors; in others it may be due to a systematic error, such as the cut-point being set too high for a particular type of patient. If the error is random and the test is repeated, the diseased and non-diseased might be reclassified.[9] In such a situation it might be appropriate to repeat slightly elevated tests to confirm that the value is stable (see Regression to the mean, below). The decision to label someone must be made taking the whole picture into account, not just a single test result. There is an inherent tension between wishing to intervene early (when treatment may prevent further deterioration) and falsely labelling a person as a patient. Such balances represent part of the art of medicine.

Nerd's Corner

Regression to the mean

Random variation in a measurement either increases or decreases the value of any single measurement on a person, compared to the mean of all values you could record for that person. Therefore, if you repeat a measurement that gave a high or a low reading, the chances are that the new value will shift towards the mean for that person. The clinical relevance of this is that an extreme result on a diagnostic test that is unstable should not cause too much concern as it may well not be replicated when the test is repeated.

Self-test questions

Self-Testing

1. What are the leading causes of death in Canada? How does the ranking of these causes change if you were to use Potential Years of Life Lost?
Leading causes of death are ischemic heart disease, followed by cancers of the trachea, bronchus and lung, with cerebrovascular disease in third place. Ranked by PYLL, for men injuries are in first place, followed by all cancers and then cardiovascular disease in third place. For women, Cancer leads, with injuries and CVD virtually tied for second place.

References

1. McDowell I. Measuring health: a guide to rating scales and questionnaires. New York (NY): Oxford University Press; 2006.

2. Morgenstern H. Ecologic studies in epidemiology: concepts, principles, and methods. Annu Rev Public Health 1995;16:61-81.

3. Notzon F, Komarov Y, Ermakov S, Sempos D, Marks J, Sempos E. Causes of declining life expectancy in Russia. JAMA 1997;279(10):793-800.

4. World Health Organization. The Composite International Diagnostic Interview, version 1.1: researcher's manual. 1994; Available at: http://www.hcp.med.harvard.edu/wmhcidi/. Accessed July 28, 2011.

5. McMaster University. Health utilities group/health utilities index and quality of life. 2000; Available at: http://www.fhs.mcmaster.ca/hug/index.htm. Accessed April, 2010.

6. Wright C, Mueller C. Screening mammography and public health policy: the need for perspective. Lancet 1995;346(8966):29-32.

7. Sackett D, Haynes R, Tugwell P. Clinical epidemiology: a basic science for clinical medicine. Philadelphia (PA): Lippincott, Williams & Wilkins; 1991.

8. Chobanian A, Bakris G, Black H, et al. The seventh report of the joint national committee on prevention, detection, evaluation, and treatment of

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

high blood pressure: the JNC 7 report. JAMA 2003;289:2560-2572.

9. Vickers A, Basch E, Kattan M. Against diagnosis. Ann Intern Med 2008;149(3):200-203.

Chapter Objectives

After reading this chapter, you will know how to:

- Access and collect health information to describe the health of a population:
- Describe the types of data used in creating a community health needs assessment.
- Be aware of important sources of clinical or population-level health data and recognize the advantages and disadvantages of each of them.
- Understand **surveillance** systems and the role of physicians and public health in reporting and responding to disease.
- Understand the importance of disease surveillance in maintaining population health and be aware of approaches to surveillance.
- Analyze population health data using appropriate measures:
- Interpret and present the analysis of health status indicators:
- Demonstrate an ability to use practice-based health information systems to monitor the health of patients and to identify unmet health needs.
- Apply the principles of epidemiology by accurately discussing the implications of the measures.
- Understand the appropriate use of different graphical presentations of data.
- Be familiar with economic evaluations such as cost-benefit / cost effectiveness analyses (see **Cost-benefit analysis** and **Cost-effectiveness analysis** in Glossary), as well as issues involved with resource allocation.
- Outline the principles of and approaches to cost containment and economic evaluation.
- Know the defining characteristics of an outbreak and how to recognize one when it occurs.
- Demonstrate skills in effective outbreak management including infection control when the outbreak is due to an infectious agent.

Case Study

Dr. Li, medical internist at Weenigo Hospital, is writing up Catherine Richards's discharge notes. Mrs. Richards had a worsening cough and was admitted from the nursing home with signs of bronchopneumonia and mild cardiac failure secondary to the pneumonia. Initial work-up also showed that Mrs Richards's diabetes was out of control and blood cultures grew pneumococcus. She responded well to therapy and went back to the nursing home a week after admission. Dr. Li is tired and is wondering what she should enter in the box marked "principal diagnosis" and just how much what she writes really matters. She wonders if she should make reference to Mrs Richards' stroke, which caused her immobility that is likely to have contributed to the development of bronchopneumonia. She would like to put ? metabolic syndrome' as one of the underlying causes for the admission, but the last time she used the term in a discharge summary the archivist warned her that there was no code for it. She is also wondering if she should report the invasive pneumococcal infection to the public health department

Surveillance

Evidence-based planning of any form of health service or preventive programme requires the initial collection of information on the types and distribution of health problems in the population. This is the task of surveillance, which refers to the systematic and ongoing collection and analysis of population-level health information in order to guide the design of public health and preventive interventions. There are two main types of surveillance. Long-term, passive monitoring of general health trends and health determinants provides key information on the health status of populations. This has documented the current epidemic of obesity and changes in trends of certain cancers. Active, ongoing, or short-term surveillance searches for emergent diseases or outbreaks such as SARS or pandemic influenza A (H1N1), and helps society respond early to new threats. In both cases, a health state, disease, or specific agent is identified as a target for surveillance.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Definition

Surveillance is the systematic collection, analysis and timely dissemination of information on population health to those who need to know, so that action can be taken. Surveillance provides information about patterns of health and disease and changes in these patterns. This information can guide prevention and control efforts as well as contribute to planning health services and, subsequently, evaluating their impact. Because health can be influenced by personal, community and environmental factors, health surveillance data come from various sources including:

- Vital statistics data, such as births and deaths
- Environmental data on air and water quality
- Health services indicators, such as hospital discharges, while
- Census data can provide information on the population, such as income, language, and ethnic group.

Much of the routine work of regional and provincial public health authorities as well as the Public Health Agency of Canada involves surveillance: http://www.phac-aspc.gc.ca/surveillance-eng.php

Role of Clinicians in Public Health Surveillance

While public health services take overall responsibility for coordinating health surveillance, individual clinicians have a major role to play by reporting occurrences of certain contagious diseases that they see. It is primary care physicians who typically see the initial cases in what may become an epidemic. They may see more people than usual presenting with a particular condition, and patients might mention other people with the same disease. While the primary care physician may be in a position to begin a preliminary investigation, she must contact the public health service to initiate prevention and control measures (e.g. contact tracing) as necessary. The public health authorities assemble such case report data to provide a broad perspective on the extent and severity of a health threat. As well as completing mandatory notification of disease, physicians also contribute to surveillance systems by completing death certificates, by entering accurate diagnoses on hospital discharge forms and billing forms.

Types of surveillance

Passive surveillance

The ?passive' in passive surveillance refers to the role of the agency responsible for it: they wait for the reports to come to them. Reports may take the form of routinely collected data, such as hospital discharge summaries, mortality data, or physician billing data. Reports can also be in the form of reports of notifiable diseases that must, by law, be reported. The health professional providing the report has an active role in this type of surveillance. However, health professionals do not always realize the importance of the information they provide, so under-reporting can be a problem.

Notifiable diseases are those considered to be of public health importance. Legislation requires physicians and laboratories to report them to the local public health agencies once they are suspected or diagnosed. This allows these agencies to track disease occurrence and identify possible outbreaks early, in order to implement prevention and control measures in a timely manner. In addition, some provinces have legislation that requires notification of possible outbreaks, even if the disease itself is not notifiable. For example, the Quebec Public Health Law demands "any physician who suspects the presence of a threat to the health of the population must notify the appropriate public health director." In this way, public health departments should be notified of potential health threats so they can investigate and intervene quickly to reduce the risk to the population. Other information on Notifiable disease is available in Chapter 11.

Nerd's Corner

Notifiable diseases

While the concept of notifiable (or reportable) disease applies mainly to communicable diseases that pose a threat of epidemic spread, it also applies to certain non-communicable conditions. In Quebec, poisoning with certain agents such a heavy metals or carbon monoxide, is notifiable, as are certain diseases caused by non-transmissible environmental agents.

The designation of a disease as notifiable may change, reflecting the level of threat it may pose to the community. As new infectious diseases emerge they may be added to the list, especially if they are commonly transmitted by international travel. A number of contagious conditions are reportable to the WHO, such as plague, cholera, yellow fever and other diseases identified in the International Health Regulations. The WHO also monitors world trends in poliomyelitis, malaria, SARS, and influenza A. Their Global Outbreak Alert and Response network is an international collaboration of experts to provide rapid response to outbreaks of international importance. Canadian federal law only provides a framework for reporting notifiable diseases, although the Public Health Agency of Canada requires the reporting of some diseases at a national level. Notification is mandated by provincial legislation and the list of diseases varies from province to province. A brief overview of the current system is available here. You can consult National Notifiable Diseases Online to see which diseases are notifiable and how frequently they occur.

Hospital and Billing Data

Hospital discharge summaries can provide useful information on patterns of disease and on the therapies being used. However, because availability of services greatly influences their use, comparisons of these data between places and over time are of limited value. Similarly, physician billing data can be used, but new methods of physician remuneration and inaccurate and missing diagnoses limit the usefulness of this data source for surveillance.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Vital Statistics: Births and deaths

Recording births and deaths is mandatory in most countries and provides basic vital statistics. Physicians are responsible for entering the causes of death on the death notification. The causes are coded according to the International Classification of Diseases so countries can compare death rates and report on the evolution of diseases. Again, the accuracy of this information depends on the accuracy with which the original reporting physician recorded the cause of death because they are rarely confirmed by autopsy.

The coding rules of the International Classification of Diseases specify that the reporting physician distinguish between the underlying cause of death and the precipitating cause(s). The underlying cause is "the disease, injury or pathological condition that initiates the chain of events leading to death.[1]" Data compilers follow a protocol which governs the decision about what disease to record as the main cause of death. For example, a patient with lung cancer who dies of pneumonia is likely to have his death attributed to lung cancer even if the pneumonia was unrelated to the cancer. Death records are routinely used to compile national information on trends in diseases such as cancer.

Case Study

Dr. Li had been unsure of how to record the principal diagnosis for Catherine Richards. She looks in the ICD to see whether it lists metabolic syndrome and the 2007 edition does not. However, it does list obesity under code E66, so she uses that instead. She wonders if there will be a revision to the ICD soon. She notices from the WHO web site that it is inviting collaborators on developing the 11th edition. http://www.who.int/classifications/icd/ICDRevision/en/index.html

Dr Li also asks her secretary, Mrs Bibi, to check about reporting the pneumococcal disease. After contacting the local public health department, her secretary comes back with the news that invasive pneumococcal infection is notifiable. Mind you, says Mrs Bibi, she had been chatting to Roxanne, a secretary in the infectious disease laboratory and Roxanne said that the laboratory passes on information about a number of diseases including pneumococcal infection to the public health unit.

Active surveillance

Active surveillance means that those responsible for it play a more active role in data gathering. This form of surveillance is more resource intensive and is usually done for specific purposes. For example, the Canadian Paediatric Society routinely sends letters to every paediatrician asking them to report on cases of rare conditions, such as acute flaccid paralysis to assess the success of polio vaccination. It also asks for information on cerebral oedema in diabetic ketoacidosis, which allowed the condition to be characterized and guidelines for the management of diabetic ketoacidosis to be developed. The Society then reports the data to the Public Health Agency of Canada.

Health surveys

Surveys, such as the Canadian Community Health Survey (CCHS) and the national census can also be viewed as active surveillance. The CCHS was initiated in 2001. Every two years it gathers data on general health and health habits from a random population sample. In the intervening years, it collects data on specific health topics from a smaller sample. Surveys can also target particular groups, such as injection drug users or people with a particular diagnosis, to document changes in patterns of behaviour that may affect disease or transmission. This is termed second-generation surveillance. The WHO gives the example of regular recording of information on HIV risk behaviours, using this to warn of or explain changes in levels of infection.

The census

Information on the population denominators that are necessary for interpreting most surveillance information comes from the census. The first national Canadian decennial census was carried out in 1871. In keeping with international norms, there has been a census every 10 years since then, always in the years ending in 1. Since1956 there has been an additional census in the years ending in 6. Both censuses cover the entire population and collect basic demographic data?about 8 questions. In addition, more detailed information is collected from a random 20% sample of the population, covering a range of demographic, social and economic topics, but not including health (about 50 questions). In the summer of 2010, the Conservative government discontinued the mandatory long-form census, replacing it with a voluntary form linked to the National Household Survey. This move was been widely criticized by a wide range of stakeholders, including French-language groups; Aboriginal and other ethnic groups; economists, city planners and the public health sector. All these sectors rely on the census to provide accurate and reliable information on demographics and socioeconomic status for program planning, advocacy and service delivery. One fear is that a voluntary long-form census will cause minority groups to be under-represented and introduce selection bias into the results.

Sentinel surveillance

Sentinel surveillance is a form of passive and active surveillance in which selected clinicians gather data and pass them on to those responsible for the surveillance. The notion of sentinel refers to keeping watch for particular diseases of interest. For example, the Canadian Primary Care Sentinel Surveillance Network links selected family health teams via an electronic record surveillance system. This can be used both to report rare events (such as side-effects they observe from influenza immunizations) and to help improve the quality of care. If the sample of physicians is carefully designed, estimates can be made of the population incidence of the event of interest without the need to survey the entire population.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Illustrative Materials

Influenza surveillance

FluWatch is Canada's national surveillance system for influenza. It provides a national picture of influenza activity throughout the year to health professionals and the public.

FluWatch

Offers early detection of influenza in Canada

Provides up-to-date information on influenza activity in Canada and abroad to professionals and the public

Monitors circulating strains of influenza virus, including new sub-types and antiviral resistance

Contributes virological surveillance information to the WHO to guide the development of vaccines for following flu season.

FluWatch information comes from several sources:

Laboratory reports of influenza and other respiratory viruses collected by sentinel laboratories across Canada

Strain identification and antiviral resistance for circulating influenza viruses from the National Microbiology Laboratory

Primary care consultations for influenza-like illnesses (ILI) from sentinel practitioners across Canada

Regional influenza activity levels as assigned by provincial and territorial FluWatch representatives

Paediatric hospital admissions and mortality associated with influenza through the Immunization Monitoring Program ACTive (IMPACT).

More information is available from the Public Health Agency web site at: http://www.phac-aspc.gc.ca/fluwatch/index-eng.php. Google also publishes information on trends in influenza activity: http://creativeherb.com/google-flu-trends-canadian-edition/

Nerd's Corner

Analyzing surveillance reports

When surveillance shows changes in patterns of disease, we should ask:

1. *Is the change real?*

Could an apparent rapid increase have happened because the initial number of cases was small? If a disease is rare, chance fluctuations can cause large proportionate variations. Grouping the data across times or places can sometimes resolve this problem; however, different groupings can sometimes lead to different interpretations of the data.

2. *Could the change be due to changed reporting practices? Are there variations in the precision of the data?*

Were all cases identified? Not every person suffering from a disease contacts health services, so they may not appear in health statistics. Screening programmes may increase the apparent prevalence of a disease because they identify cases that were previously undiagnosed. Did something change reporting practices? For example, the recent increase in reported Chlamydia cases may result in part from an emphasis on routine screening, and especially since the advent of non-invasive urine testing (see L.K. Altman, Sex diseases still rising; Chlamydia is leader, New York Times November 14, 2007)

3. *Might the change be due to a shift in definition of the disease?*

Are the same diagnostic criteria being used? If criteria change, apparent prevalence of the disease may change. For example, changes in disease definitions from the ninth to tenth version of the International Classification of Disease led to a drop in the number of deaths attributed to falls and pneumonia and a rise in the number attributed to cancer. Surveillance professionals generally adjust for the coding change, and you may need to take account of the change if you compare surveillance reports written before and after 2001. Similarly, changes in the diagnostic criteria of autism led to mistaken reports of its increasing incidence.

A bit of history: The International Classification of Disease originated in 1891 when the International Epidemiological Society started to classify causes of death. By 1900, 26 countries had adopted the classification. Since then, most countries have adopted it and the classification has been revised to reflect the development of medical science. After 1946, the WHO took over responsibility for coordinating revisions to the list. The tenth revision was published in 1994 and adopted in Canada for coding causes of death in 2001.

Patterns of disease development in a population: the epidemic curve

When our efforts to prevent disease fail and an outbreak develops, the resulting distribution of cases may take various forms that are called **epidemic curves** (see **Epidemic curve** in Glossary). These show the distribution of incident cases over time, tracing the evolution of a population outbreak, which is equivalent to the idea of the natural history of disease for an individual case. The natural history of a population outbreak is most evident in infectious disease, but also occurs in situations such as a chemical spill leading to cases of respiratory disease or, on a much longer time-scale, in non-communicable, chronic diseases. The shape of the resulting epidemic curve can be used to propose hypotheses on the nature of the disease and its mode of transmission. The curve can provide information on the pattern of spread over time, the magnitude of the outbreak (the number of cases), the likely incubation period for the condition, and can reveal outliers (in time, perhaps in place).

To characterize different types of outbreak, the Centers for Disease Control classify epidemic curves based on the suspected type of exposure.[2] In a *common source outbreak* people are exposed, in a group, to a single noxious influence. The source may occur for a brief time or it may persist,

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

depending on the circumstances. When the exposure is very brief, most people get sick at one incubation period following exposure, and this is called a *point source outbreak*?an example would be *Staphylococcus aureus* food intoxication from a single meal of tainted food at a wedding (how embarrassing!) This produces a single curve that wanes quickly, as long as there is no person-to-person spread (Figure 7.1).

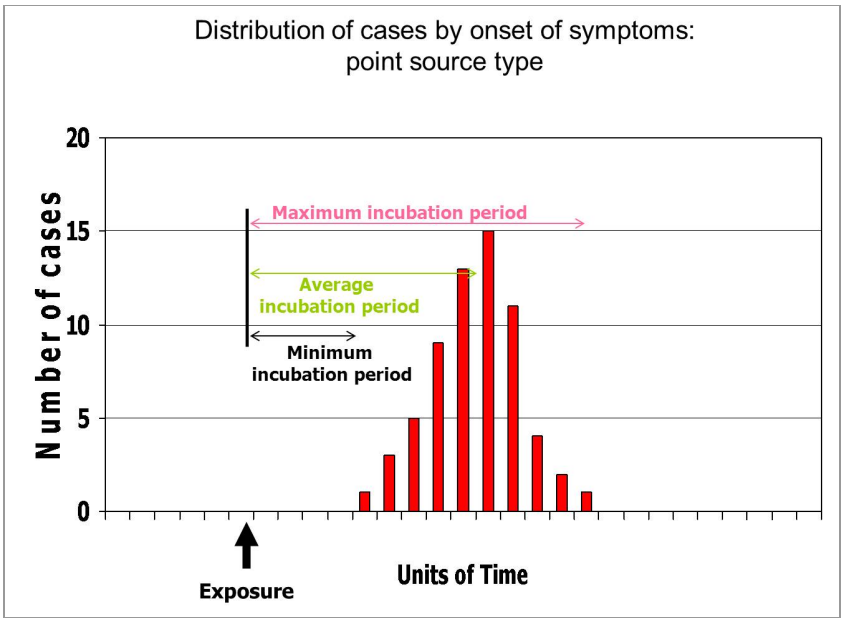


Figure 7.1: Point source exposure epidemic curve

The distribution of cases over time reflects the differential incubation period for different individuals and the time it takes them to seek help and, thereby, be included in the data collection. Perhaps some people ate more of the tainted food and became ill sooner than others, or they were more susceptible, or sought care more quickly.

Continuous source: Sometimes a common source exposure can be prolonged, as with exposure to a contaminated water supply, or via a restaurant that has a faulty refrigeration system that does not get fixed. Here, cases arise over an extended period but still originate from a common or single source. The resulting epidemic curve becomes longer and flatter, indicating the longer duration of the source and the variation in incubation periods between people: see Figure 7.2. The curve ends when the source of the contamination is corrected or when all susceptible people develop immunity. The relative flatness of the curve suggests that the infection comes from a common source and there is no person-to-person spread; otherwise the number of cases would grow over time as one person infects others.

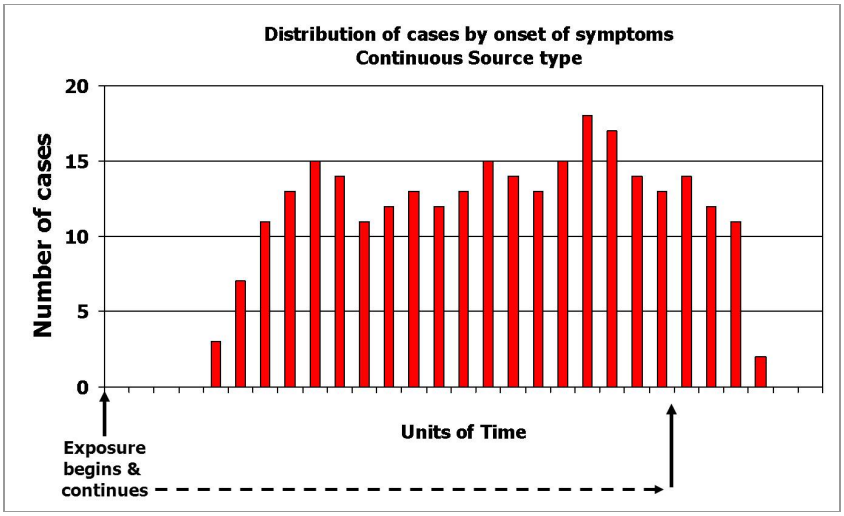


Figure 7.2: Continuous source epidemic curve

Intermittent exposure: Figure 7.3 shows an irregular pattern of cases that reflects the timing and extent of repeated exposures. It is not immediately clear whether this is a common source, such as an industrial contaminant emitted at intervals, or arises from varied sources, such as a series of outbreaks of food poisoning occurring at different summer camps for children. The gaps between the outbreaks could initially suggest person-to-person transmission followed by an incubation period, but the successive peaks do not become larger and merge as they would if the outbreaks were due to infectious spread, with one person infecting several. Hence, the epidemic curve in Figure 7.3 suggests a non-transmissible condition.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

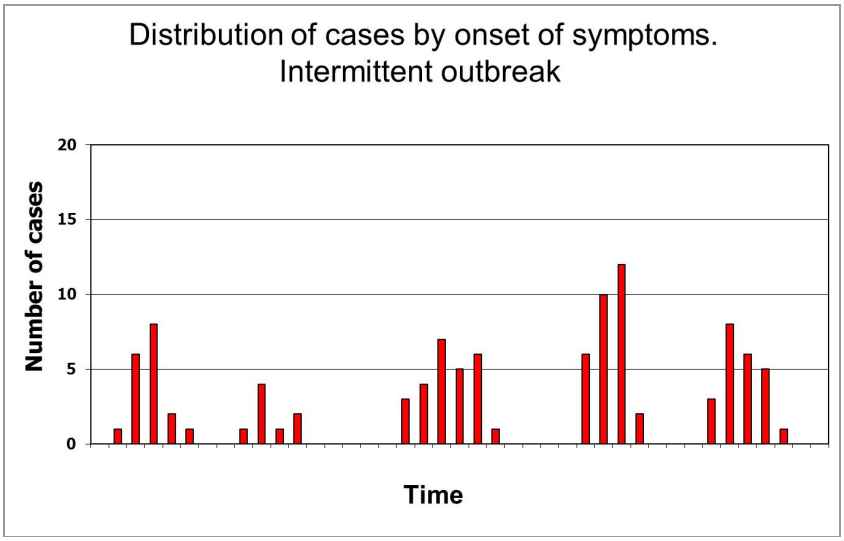


Figure 7.3: Intermittent outbreak epidemic curve

Index case with limited spread. Person to person spread is illustrated in Figure 7.4, which shows the typical pattern arising when a single index case (for example, a traveller returning from abroad) infects other people after an incubation period. This is called a point source with secondary transmission. The outbreak wanes when the infected people no longer transmit the infection to other susceptible people, perhaps because of successful control measures (isolation or quarantine).

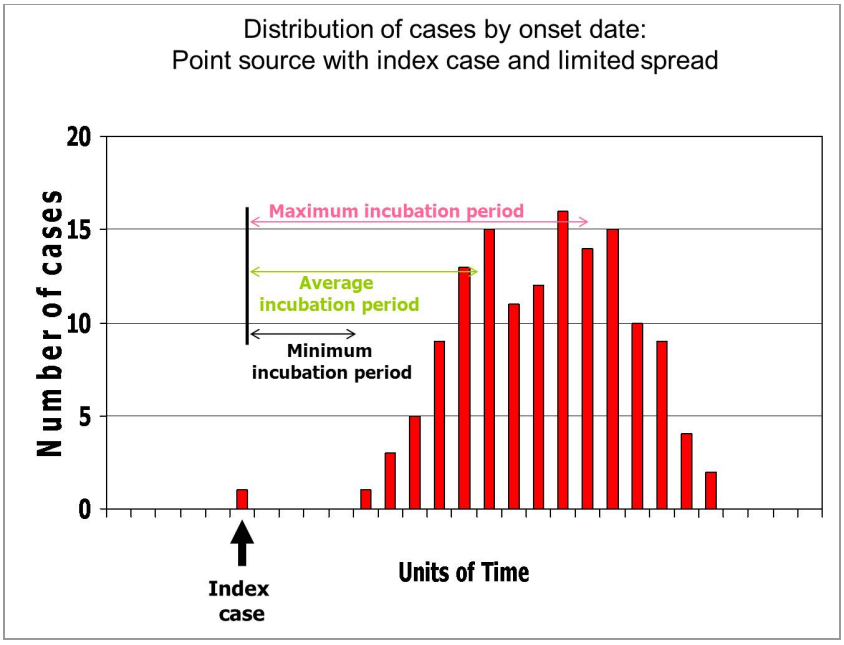


Figure 7.4. Secondary transmission from a single index case

Propagated Spread. This begins like an infection from an index case (Figure 7.4), but the secondary cases of the disease then act as sources to infect new people who, in turn, serve as sources for yet other cases. Figure 7.5 illustrates how this produces successively taller peaks, initially separated by one incubation period, but the peaks then tend to merge into waves with increasing numbers of cases in each generation (i.e., secondary and tertiary cases) and the epidemic continues until the remaining numbers of susceptible individuals declines or until intervention measures take effect. This pattern occurs with diseases such as measles that spread from person to person.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

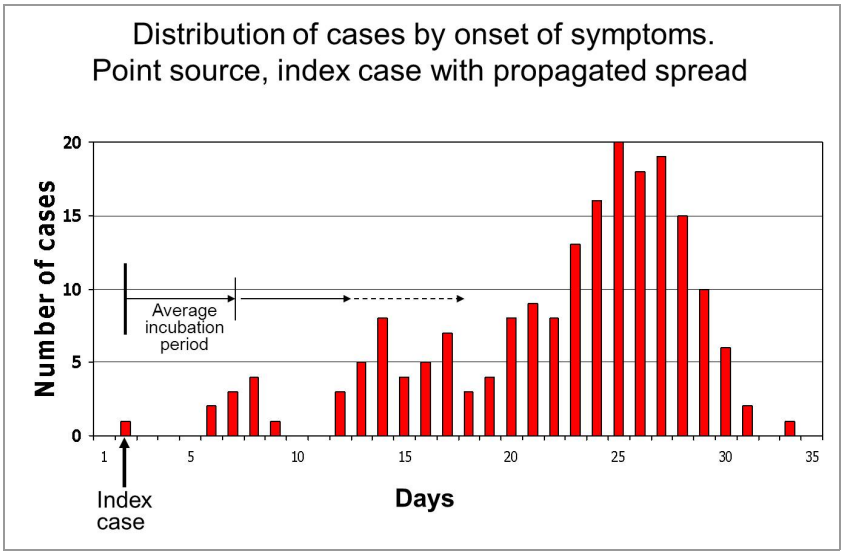


Figure 7.5: Index case with propagated spread epidemic curve

Health services research

The rising costs of medical advances mean that publicly funded systems have difficulty paying for every new treatment that becomes available. The costs mean that managers and providers of health services have to make choices about which programmes to fund. These decisions relate to whether the programmes are required, whether they are beneficial, and whether they may be improved upon. Research that informs such decisions is known as **Health Services Research**, which can be divided into needs assessment, economic evaluation, and overall programme evaluation.

Needs assessment

A key question faced by those who plan health services concerns how many and what types of health services the population needs (see **Need for care** in Glossary). This question has traditionally been answered by examining utilization data and historical patterns of care, which are used to predict future demand taking forecasts of population growth into account. An advantage of this approach is that these data are usually readily available and can be compared to other jurisdictions in order to estimate the appropriate level of service provision. However, utilization data generally reflect supply and historical demand much more than they reflect actual need for services.[3,4]

Services can also be planned on the basis of epidemiological needs assessments.[5-7] This method is based on the idea of the capacity to benefit from interventions (preventive as well as curative). It begins from the prevalence of a given disease and its risk factors in the whole population, not just those who consult. It then incorporates evidence of the effectiveness of available treatments for the disease to estimate the numbers of required interventions. These estimates can be compared with the current level of service provision to identify procedures that may be over- or under-supplied. An example of assessing the need for stroke services is presented in the Illustrative Materials Box

Illustrative Materials

How many hospital beds will we need for stroke victims?

A study estimating the requirement for stroke services in Eastern Ontario was undertaken in five steps:

- 1. Identify stroke risk factors through a combination of systematic search strategies and consultations with expert clinicians;
- 2. Estimate the frequency of strokes and of the identified risk factors in Eastern Ontario;
- 3. Identify effective health services targeting each condition or risk factor from systematic reviews and published practice guidelines;
- 4. From this, determine the number and type of health services required;
- 5. Compare these estimates with the actual provision of services.

The study showed several mismatches between the estimated need and actual provision of stroke services. The approach offers an effective way of making health care planning decisions.[8]

Economic evaluation

Limited resources mean that choices need to be made among health services, and health economics (see Economic analysis in Glossary) uses information on the costs and outcomes of care to guide these choices. Health economics is a branch of economics that considers costs, benefits, resource allocation, use, inputs, outputs and outcomes of all forms of health care.[1] Costs are the value of resources used to deliver a service. These will include items such as staff wages, buildings, equipment, maintenance, and supplies. Costs are usually measured in terms of monetary value. Outcomes or benefits are the results of a treatment or intervention, and can include symptom relief, survival rates, or improvement in quality of life, as reviewed in Chapter 6. In economic evaluations, outcomes are often measured in terms of QALYs or DALYs (see **Quality-adjusted life years** or **Disability-adjusted life years** in Glossary).The key economic assumption is that a health service should deliver the greatest possible benefit per unit of cost. As there are many ways to estimate benefits or outputs, there are four main types of economic evaluation, which are similar in the way they assess cost inputs but differ in the way they assess outputs.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

1. *Cost minimization analysis*. This is the simplest approach, applicable when the benefits of two interventions are the same, so favouring the cheaper option. An example would be the choice between a name brand and a generic drug.

2. *Cost benefit analysis*. This assesses benefit in terms of monetary value (i.e., dollars). An example would be a company's fitness programme that has a cost per participant, but that brings financial benefit to the company in terms of reduced sick leave and increased productivity.

3. *Cost effectiveness analysis*. This assesses output in terms of improved health outcomes, such as symptom control or survival. An example is a study that compared the standard method in vitro fertilization with a novel one. The study considered the different costs associated with the approaches and used the number of pregnancies resulting in live births as the outcome.[9]

4. ***Cost-utility analysis***. This is a variant of cost-effectiveness, where the measure of outcome is adjusted to include a judgment of utility, via QALYs or DALYs or the health-adjusted life expectancy (HALE). A significant advantage is that this approach allows for comparisons across different procedures and their related outcomes. Note that many published papers purporting to be cost-effectiveness analyses are actually cost-utility analyses.

Links

The cost of prevention

Is an ounce of prevention really worth a pound of cure? A challenging discussion of the cost effectiveness of prevention programmes is available from the Canadian Health Services Research Foundation website at: http://www.chsrf.ca/MIGRATED/PDF/MYTHBUSTERS/myth9_e.pdf

Programme evaluation

A health programme is a planned series of activities; the plan indicates the target, financing, roles and responsibilities, and intended outcomes of the programme. The national breast cancer screening programme, the cardiac rehabilitation clinics in the Centre hospitalier universitaire de Sherbrooke or at the Ottawa Heart Institute, the annual flu vaccination drive in Dr Rao's clinic are all examples of health programmes. Because of the universal emphasis on saving costs, evaluation has to be built into all programmes to demonstrate that public funds are being spent responsibly. Programme evaluation is the "systematic gathering, analysis, and reporting of data about a programme to assist in decision making". This definition is comparable to the definition of research in general, but the key difference is that it refers to a specific programme and that the aim is to aid decision making. An evaluation ought to address the question: Does the programme achieve its stated goals and objectives? (Goals give a broad statement of what a programme hopes to achieve, while objective are more precise targets that form steps on the path to achieving the overall goal). The answer may not be straightforward because the programme goals and objectives may not be stated in a way that makes them readily measurable. The procedures of quantitative and qualitative research described in Chapter 5 apply to programme evaluation.

Programmes can be evaluated in terms of their structure, process, and outcome, as proposed by Donabedian in 1966:[10]

- *Structure:* the adequacy of facilities, equipment, staffing, and/or administration of the programme. (Should Dr Rao hire another nurse to help Nurse Jennings on ?vaccination night'?)
- *Process:* what the programme does. For example, in a hospital this could cover the adequacy of history taking, physical examination, diagnosis, therapy, or secondary prevention. From a public health perspective, process might cover the number of screening tests performed, the number of radio spots aired, etc. (How should the clinic be set up on ?vaccination night', where should the vaccines be given, who should administer the pre-vaccine screening questionnaire and consent??)
- *Outcome:* the results or outputs of the programme. Patient attitudes and satisfaction, their level of rehabilitation or residual disability, or the death rates observed among them. (What proportion of the patients in the target group was immunized on ?vaccination night', will they come again next year, how many patients got influenza this year compared to last year?)

Possible patient outcomes of health services

Fletcher et al. offered a slightly different version of the five Ds that were introduced in Chapter 6, and these are often used as a framework for measuring patient outcomes in health services research:[11]

Dissatisfaction	Emotional reaction to disease and its care, including sadness or anger.
Disease	Symptoms, physical signs, and laboratory abnormalities.
Discomfort	Symptoms of pain, nausea, dyspnoea, itching, and tinnitus.
Disability	Impaired ability to undertake usual activities at home, work, or recreation.
Death	A poor outcome, if it occurs prematurely or if it is painful.

There are many ways to evaluate a health programme; one example is given in the Nerd's corner box "Steps in evaluating a programme".

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Steps in Evaluating a Programme

Stage 1: The Logic Model.

The first step is to work from the "logic model" for the programme, which should have been developed during the planning process for a complex programme. The logic model outlines how an intervention (such as a public health or prevention programme) is expected to produce results. The purpose is to ensure that all people involved are aware of how their activities contribute to the overall goal: it gives the big picture. It also helps guide evaluation of the programme by identifying measurable intermediate goals that track progress toward the overall goal and thereby identify where failures occurred if the end goal is not met. The elements in a typical model include a description of the situation or problem to be addressed, the inputs or resources required, the activities to be undertaken, the outputs, and the anticipated outcomes.

Here is an example of an imaginary logic model for the dental health programme described in Chapter 4.[12]

- *Situation*: The oral health of children in Glasgow had been documented as being among the poorest in Western Europe
- *Inputs*: A collaboration between the dentists, primary care physicians and city authorities
- *Activities*: Creating healthy public policy; Supportive environments; Developing personal skills; Strengthening community action; Reorienting health services (details are given in Chapter 4)
- *Outputs*: numbers of nursery schools contacted; numbers of children receiving dental care; improvements in dental health indicators
- *Target systems*: nursery schools, families with children under 5; community centres
- *Outcomes*:

- *Short term*: participation by mothers in oral health training programmes; reduced consumption of caries-inducing foods; changes in foods served in nursery schools
- *Medium-term*: improved oral health profile
- *Longer-term*: reduction in the percentage of adolescents and adults with dental problems.

A useful resource is the University of Toronto's Centre for Health Promotion "Logic Models Workbook" http://www.thcu.ca/infoandresources/publications/logicmodel.wkbk.v6.1.full.aug27.pdf (Accessed October, 2010).

Stage 2: The Evaluation

The Public Health Agency of Canada lists five key steps in evaluating a programme. Again, these may be applied to the dental health promotion programme.[12]

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

PHAC Evaluation Guide	Illustration, as applied to dental health promotion programme
1. Identify the purpose of the evaluation: ▫ State the purpose of the evaluation ▫ Develop a logic model for the programme, indicating how the structure and process are expected to achieve the outcome (see Nerd's corner box) ▫ List the stakeholders ▫ Develop evaluation questions and check the feasibility of undertaking an objective evaluation	"To assess the dental health outcomes of an oral health promotion programme by secondary analyses of routine caries datasets for Glasgow 5-year-olds over the interval from 1997-98 to 2003-04.[12]" The logic model is shown in the box below. Stakeholders included the City of Glasgow Health department; University of Glasgow Dental Hospital; Scottish Office Department of Health; several community groups. The feasibility of accessing dental records was confirmed
2. Select appropriate methods: ▫ State expectations ▫ Formulate a data collection plan ▫ Develop a logistics plan and run a feasibility check	A self-report survey appeared infeasible; there was need for expert assessment of oral health to record outcomes. "In Glasgow, cross-sectional caries surveys of randomly selected 5-year-old children in primary school reception classes are carried out routinely as part of a national programme." Permission was granted to retrieve anonymous data from previous surveys.
3. Develop tools: ▫ Review existing measures ▫ Select questions and response categories ▫ Plan a quality assessment for the data collection	The surveys "are conducted according to the British Association for the Study of Community Dentistry standardized criteria" "?and involve annual national training and calibration exercises immediately preceding each survey?"
4. Gather and analyze data: ▫ Data collection and pre-test ▫ Quantitative and/or qualitative analysis	(This study used secondary data from an existing data set) An index of area socioeconomic deprivation was added to the oral health data, linked by postcode. Quantitative analyses used changes in mean oral health scores over time, comparing experimental city districts to districts in which the programme was not implemented.
5. Make decisions: ▫ Interpretation of results ▫ Formulate an action plan ▫ Produce a report.	"This paper provides evidence of positive and reproducible outcomes following targeted community-based oral health promotion activities in SES-challenged communities." "The earlier in infant life that caries risk factors are ameliorated, the greater the impact on caries incidence for the individual child and caries prevalence at the community level.[12]"

Case Study

Dr. Li's Case Study Revisited:

1. What is the importance of documenting diagnoses in hospital discharge summaries?

Public health practitioners use hospital discharge information and billing codes to monitor the trends in incidence, prevalence and clinical severity of various acute and chronic diseases. This hospital data provides valuable information for program planning and evaluation and enables international comparisons of disease rates.

2. How are diseases classified internationally?

Through the ICD-10 classification (International Classification of Disease, 10th revision).

3. Which types of infections need to be reported to the public health department? Why is this important?

While the concept of notifiable (or reportable) disease applies mainly to communicable diseases that pose a threat of epidemic spread, it is also applies to certain non-communicable conditions. The Public Health Agency of Canada publishes criteria for nationally notifiable conditions: http://www.phac-aspc.gc.ca/publicat/ccdr-rmtc/06vol32/dr3219a-eng.php#app2

Self-test questions

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Self-Testing

1. Identify the differences between active, passive and sentinel surveillance, and provide an example of each.
Passive surveillance involves reviewing routinely collected information (sometimes collected for other purposes, such as physician billing records) to record trends in disease. The information is then passed on to public health authorities so that action may be taken if required. An example would be routine reporting of cases of notifiable diseases.
Active surveillance is often used under more urgent circumstances, and implies setting up a system for collecting data on a disease specifically in order to track its development. This is often initiated by public health authorities who are concerned about a particular public health threat. An example would be reporting cases of H1N1 influenza during a suspected outbreak.
Sentinel surveillance is a combination of both: it involves establishing a carefully selected network of physicians or public health units who routinely report information on diseases of interest. An example would be a network of primary care physicians who are chosen according to where they practice in order to study and contrast patterns of sexually transmitted infections in rural and urban areas.
2. A systematic review has concluded that a new programme to reduce hypertension is extremely effective. In your capacity as advisor to the Minister of Health, how would you determine the need for such a programme and how would you evaluate its impact in terms of outcome and cost?
Turning an effective programme into a population-wide policy implies presenting a political argument that would explain the justification for diverting funds away from another programme (opportunity cost).
An example of the steps to be followed to formulate an effective political and economic argument for a new hypertension policy are as follows:
First, you would have to demonstrate the current costs of hypertension (in terms, for example, of numbers of people affected, or their subsequent morbidity and mortality, the costs of treating conditions that arise secondary to hypertension, etc).
Second, summarize the current approach to managing hypertension: how well does the current system work; how badly broken is it and how much of an advantage would the new programme be?
Next, you would have to assess the feasibility of implementing the proposed new programme on a population-wide basis (do you have the facilities, staff, equipment for implementing the programme; could it be implemented in all areas, or would it create disparities in access?)
Finally, you need to anticipate the costs of implementing the programme, and compare these with the benefits that would accrue: will an early intervention for hypertension actually reduce costs of subsequent treatment? Ideally, this cost-benefit analysis should be compared to the cost-benefit of existing ways of managing hypertension and of any other programmes that could be abandoned to release funds for this new programme.
Many of these steps would involve modelling the programme impact and costs, based on estimation procedures that rely on previous surveys (for example, covering the prevalence and geographic distribution of hypertension).

References

1. Last JM. A dictionary of public health. New York: Oxford University Press; 2007.

2. Centers for Disease Control. Interpreting an epidemic curve. In: Epidemiology in the classroom: steps in an outbreak investigation. Available at: http://www.cdc.gov/excite/classroom/outbreak/steps.htm. Accessed July, 2010.

3. Morgan M, Mays NH, WW. Can hospital use be a measure of need for health care? J Epidemiol Comm Health 1987;41(4):269-274.

4. Labelle R, Stoddart G, Rice T. A re-examination of the meaning and importance of supplier-induced demand. J Health Econ 1994;13(3):347-368.

5. Acheson R. The definition and identification of need for health care. J Epidemiol Community Health 1978;32:10-15.

6. Stevens AR, J. Introduction. In: Stevens A, Raftery J, editors. Health care needs assessment: the epidemiologically based needs assessment reviews Oxford: Radcliffe Medical Press; 1994.

7. Frankel S. Health needs, healthcare requirements, and the myth of infinite demand. Lancet 1991;337:1588-1590.

8. Hunter D, Grant H, Purdue M, Spasoff R, Dorland J, Bains N. An epidemiologically-based needs assessment for stroke services. Chronic Diseases in Canada 2004;25(3/4):138-145.

9. Polinder S, Heijnen E, Macklon N, Habbema J, Fauser B, Eijkemans M. Cost-effectiveness of a mild compared with a standard strategy for IVF: a randomized comparison using cumulative term live birth as the primary endpoint. Human Reproduction 2008;23(2):316-323.

10. Donabedian A. Evaluating the quality of medical care. In: Mainland D, editor. Health Services Research: Milbank Memorial Fund; 1967.

11. Fletcher R, Fletcher S. Clinical epidemiology: the essentials. (4th ed) ed. Baltimore (MD): Lippincott, Williams & Wilkins; 2005.

12. Blair Y, Macpherson L, McCall D, McMahon A. Dental health of 5-year olds following community-based oral health promotion in Glasgow, UK. Int J Paediatric Dentistry 2006;16:388-398.

AFMC Primer on Population Health
Part 3 - Practice: Improving Health

AFMC Primer on Population Health
Chapter 8 Illness Prevention and Health Promotion

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Chapter Objectives

After reading this chapter you will be able to:

- Appreciate the role that physicians can play in promoting health and preventing diseases at the individual and community level (e.g., prevention of low birth weight, immunization, obesity prevention, smoking cessation, cancer **screening**, etc.);
- Understand how public policy can influence population-wide patterns of behaviour and affect the health of a population;
- Define the concept of levels of prevention at the individual (clinical) and population levels, as well as formulate preventive measures into their clinical management strategies;
- community-level interventions to promote health and prevent disease;
- Describe the advantages and disadvantages of identifying and treating individuals versus implementing population-level approaches to prevention;
- behaviour change, including predisposing, enabling, and re-enforcing factors; and
- Identify the potential community, social, physical, and environmental factors that might promote healthy behaviours, as well as ways to assist communities in addressing these factors.

Case Study

Dr. Rao is seeing David Richards. Four weeks ago David went to the city emergency room with a cough. He was given antibiotics. The ER doctor suspected underlying asthma, so she referred David back to Dr. Rao for follow-up. With a detailed history, knowing the family, and having examined the respiratory system, Dr. Rao is fairly certain that David doesn't have asthma. However, he notes that David smokes tobacco and marijuana and regularly drinks more than 5 units of alcohol in one sitting. He also pays little attention to safety. He doesn't use his car seat belt. Although he works in construction, he doesn't have steel-capped boots and doesn't bother with ear protectors on noisy jobs, and has, at least once, been dismissed from a work site for not wearing a helmet. David has a steady girlfriend, in spite of his occasional "one night stands." He uses condoms only if his partner insists. Having discovered all this, Dr. Rao decides he needs to do something about helping David to reduce his health risks.

Approaches to health improvement

There are two main overlapping approaches to maintaining and improving health: the first is to identify individuals at high risk and intervene to reduce their risk, and the second is to reduce the average risk level for the whole population. Intervening with individuals at high risk is generally the domain of clinical medicine, although public health authorities coordinate certain clinically implemented programmes in order to achieve population health objectives. For example, in Canada, the breast cancer screening programme and the childhood vaccination programme are undertaken in individual encounters, but have population level objectives. Other individual interventions, such as hypertension and diabetes screening and treatment, or counselling on tobacco and alcohol, are generally not organized to the extent that they have population level objectives.

The second way of improving health is to intervene at the population level; sometimes this is without the consent or even knowledge of people in the population. An example of a population level intervention is the addition of certain nutrients to foods to reduce disease. For instance, iodine can be added to salt to reduce goitre, vitamin D can be added to milk to replace what is lost in the skimming process, B vitamins are added to flour and bread to replace the vitamins removed with the bran. These population-wide interventions work by shifting the entire distribution of exposure to reduce the overall level of risk in the population. Although some individuals retain a risk higher than the population mean, the overall burden of disease is reduced.

Population-wide interventions also target health determinants in an attempt to improve overall health, rather than to prevent specific diseases. For instance, if income redistribution policies succeed in reducing poverty, they will improve health and reduce the burden of all diseases associated with poverty. Improving built environments can address a number of health determinants and risk factors at once. For instance, neighbourhoods that offer pedestrians safe walkways and services within walking distance encourage active transport and, as a result, address risk factors for the metabolic syndrome and arthritis. Adequately lit buildings reduce the risk of accidents. These broad kinds of interventions are difficult to bring about because they require collaboration across a number of sectors in society, each one of which has different roles and objectives.

Intervening in individuals or in populations?

Geoffrey Rose discussed the relative merits of treating individuals and treating populations.[1] His ideas stem from the empirical observation that many cases of disease arise in people who are not in a high-risk group. Moreover, the number of cases arising from the population at average risk is often greater than the number occurring in the population at high risk simply because there are so many more people in the average-risk population. Figure 8.1 uses Canadian data to show that 61% of cases of diabetes mellitus will come from people at low or moderate risk, whereas only 39% will arise from those at high risk. Rose illustrated a similar point using heart disease, hypertension, and trisomy 21 as examples, but there are many other diseases that show similar patterns.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

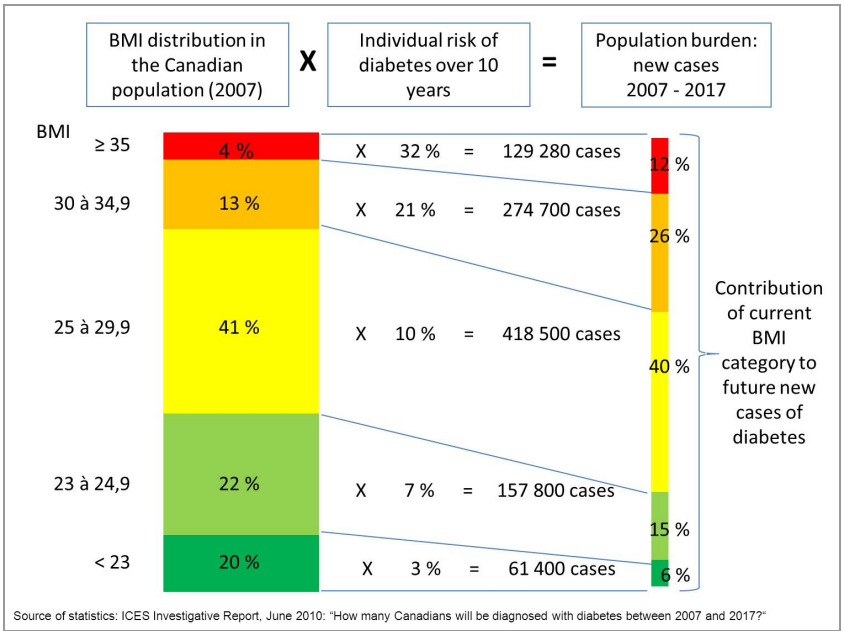


Figure 8.1 Comparison of the number of incident cases of diabetes by category of body mass index, Canada 2007-2017.[2]

Rose proposed that preventing disease by trying to shift the entire population distribution of a risk factor can be more efficient than focusing interventions solely on people at high risk. For instance, estimates suggest that in North America a 14% decrease in the number of cerebrovascular accidents could be achieved either by decreasing the average blood pressure by 2 mm Hg or by successfully treating everyone with a diastolic pressure of 95 mm Hg or greater.[3] However, not all diseases are amenable to this type of population strategy. For instance, an analysis of the potential effects of cholesterol reduction to prevent coronary heart disease suggested that a high risk strategy is more likely to be effective than a population strategy.[4] The optimum preventive strategy depends on the disease to be prevented, the distribution of its risk factors in the population and the likelihood of achieving the desired reduction in the risk factor.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Illustrative Materials

Trisomy 21 high risk and average risk

Comparison of the effects of total eradication of risk in a high-risk group versus a small reduction of risk in the entire population

Table 8.1 shows figures that Rose used to illustrate his arguments;[1] they refer to the risk of giving birth to a child with trisomy 21 by maternal age group in 1979.[5] Women aged 35 and over are considered to be at high risk for giving birth to a baby with trisomy 21. If screening aimed at this group were 100% effective and prevented all trisomy 21 births, this would reduce the total number by 6 (see births for women aged 35+ in the column ?Number of trisomy 21 births').

Table 8.1: Distribution of trisomy 21 infants by the mother's age.

Mother's age	Total births	**Trisomy 21 per 1,000 births**	Number of trisomy 21 births
Under 30	111,429	0.7	78
30 to 34	12,308	1.3	16
35 to 39	1,351	3.7	5
40 to 44	73	13.1	0.95
45 and over	1	34.6	0.05
Total			100

By contrast, Table 8.2 illustrates the hypothetical impact of a population-wide intervention that reduced the risk by just 0.1 per thousand in all age groups:

Table 8.2: Impact of reducing risk of trisomy 21 births by 0.1 per thousand across all maternal ages

Mother's age	Total births	Trisomy 21 per 1,000 births	Number of trisomy 21 births
Under 30	111,429	0.6	67
30 to 34	12,308	1.2	15
35 to 39	1,351	3.6	5
40 to 44	73	13.0	0.94
45 and over	1	34.5	0.05
Total			87

Lowering the risk of trisomy 21 by just 0.1 per thousand pregnancies in all age groups would reduce the total number of trisomy 21 births to 87, a decrease of 13?a better result than achieved by the high risk strategy.

Individual and population approaches each have advantages and disadvantages, as outlined in Table 8.3. In practice, as indicated in the **Ottawa Charter for Health Promotion** (see Chapter 4), health improvement requires different strategies to reach different levels?individuals, communities, and populations?simultaneously. In fact, it should draw on the ecological model of health.

Table 8.3: High-risk versus population strategies (adapted from Rose[1])

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

	Individual-centred ?high-risk' strategy	Population ?average-risk' strategy
Advantages	Intervention is appropriate to the individual. People who learn that they are at high risk are likely to change their behaviour to reduce their risk (as predicted by the Health Belief Model). Physicians feel justified in reducing risk factors in high-risk patients. Arguably cost-effective use of resources, which are directed to individuals most in need of them. Favourable benefit to risk ratio; high-risk individuals are likely to gain more benefit from the intervention for the same likelihood of harm as lower-risk individuals.	Intervention aimed at roots of problem reduces illness in the whole population, including those at low or average risk. Tackles condition in its early stages when interventions may be more effective. A small change in the level of a risk factor in a population can improve the health of a large number of people Behaviourally and socially appropriate?as non-smoking becomes more "normal," smokers smoke less and are more likely to attempt to stop smoking.
Disadvantages	Difficulties and costs of identifying high risk groups and individuals. Dividing line between average and high-risk is often arbitrary, and many ?average-risk' people can still be at significant risk. Reaches those most at risk but has little impact on the disease burden in society, because most cases of disease occur in people at low or moderate risk. Palliative and temporary?the determinants are not addressed, so there will always be individuals who need the intervention. Behaviourally inappropriate?a change of behaviour sufficient to reduce risk significantly may put the individual outside the norms of his or her social circle.	The small benefit to most individuals (see Prevention paradox) can be outweighed by the risk of the intervention, even if this is also small. Inefficient (and perhaps morally questionable): it demands change by a large number of people who would not have developed the disease at all. There is little motivation for low-risk individuals to change behaviour. There is a danger of increasing inequity in health (see Health inequity).[6] Unless specifically designed strategies are used, vulnerable people in most need of change are least likely to do so. Intervening in apparently healthy people is ethically more sensitive that intervening in people with problems. It can be seen as close to social engineering, which is unacceptable in a liberal society.

Nerd's Corner

The case against Rose

Since Geoffrey Rose published his article "Sick individuals and sick populations" in 1985, followed by his book *The Strategy of Preventive Medicine* in 1992, his approach has been criticized. Some critics argue against the ethics of the population approach to prevention, declaring that for many diseases evidence of causation is weak and information on risk factors comes from observational studies. Proactive preventive actions should be based on solid evidence, but controlled trials on entire populations would be difficult. If a scientifically secure basis is not required before intervening on the lives of all people, the possibilities of abuse are considerable.[7]

An additional argument holds that Rose took an unduly negative view of the high-risk intervention approach. As our understanding of risk factors increases, it should be possible to target interventions more precisely to those at risk, making the high risk approach at least as effective as a population approach in preventing specific diseases. Rose's demonstration focused on the use of a single marker to identify high-risk people, producing too many false positives and false negatives to allow accurate identification of a person's risk. More recent risk algorithms, some of which include genetic markers, allow more accurate identification of individuals at high risk, including many who would not have been identified using only a single risk factor. Use of these algorithms may well prove more effective in lowering disease mortality rates than the use of population strategies.[4]

In practice, health improvement strategies should consider both individual and population approaches. The two are generally not mutually exclusive and will usually complement each other.

Links

For an easy to read discussion on population and individual risk, see Chapter 2 of the 2002 World Health Report:
http://www.who.int/whr/2002/chapter2/en/index4.html

The goal of prevention

At the population level, disease prevention serves to compress morbidity. Over time, individuals are exposed to various noxious factors so that by middle age their health begins to decline, and continues to do so at varying rates until death. This is illustrated using survival curves and is shown in

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Figure 8.2. As shown in the diagram, preventing?or at least delaying?potentially fatal diseases shifts the survival curve up and to the right, so the curve becomes square. An obvious concern is that extending the life of as many people as possible to an advanced age may mean that the life years gained are lived in a state of poor health. Empirically, however, this turns out not to be the case, as the actions that extend life also seem to improve health, so that the combination of disease prevention and health promotion tends to raise the disability-free survival curve also. Squaring the curve represents the presumed ideal of postponing the decline in health associated with age so that people enjoy a good quality of life and die as healthy as possible as late as possible; it implies that interventions should aim to improve both quality and length of life, not merely to prolong life at all costs. In practice this is reflected, for instance, in the decision not to recommend routine that breast cancer screening continue into old age, because the life years gained would not outweigh the costs and potential harm of screening.

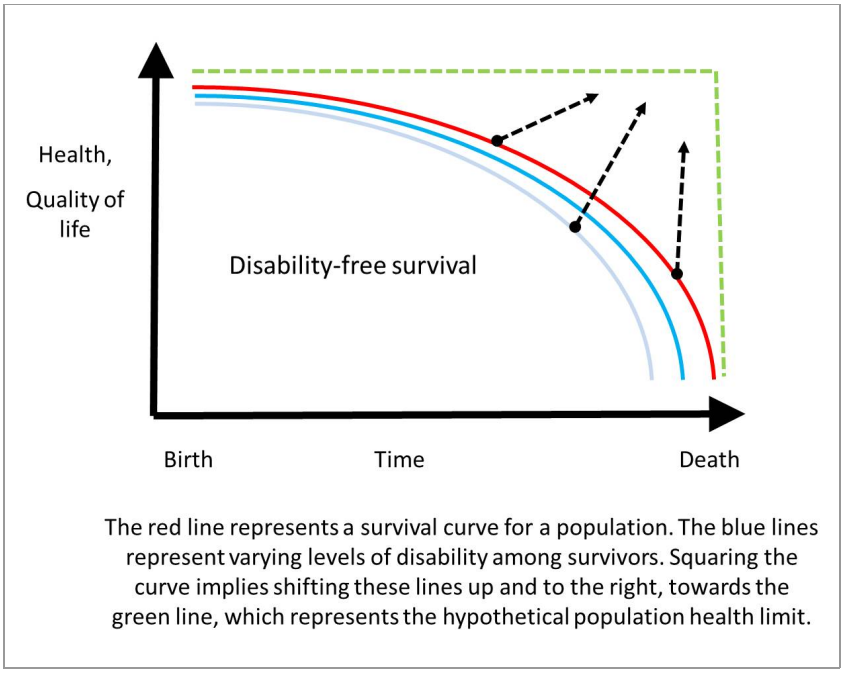


Figure 8.2: Squaring mortality, compressing morbidity curves (Adapted from Fries[8])

The Canadian Task Force on Preventive Health Care

To improve patients' health, the physician, in consultation with the patient, is responsible for choosing effective health interventions. The methods of evidence-based medicine are useful in guiding physicians in the choice of effective interventions. However, the amount of evidence is overwhelming, so a number of professional and academic organizations have reviewed it to produce evidence-based guidelines for disease management and prevention. Previous generations of North Americans believed that an annual physical examination was essential for maintaining health. The annual check-up was promoted by insurance companies in the 1920s and later by the American Medical Association in the 1930s and so became part of the collective medical creed. It consisted of a head-to-toe physical examination and the use of whatever tests were available: blood count, urine glucose and protein, chest X-ray, and, since the 1950s, ECG, CT-scans, and MRIs.

When Medicare began in Canada and everyone could have a government-funded annual physical examination, provincial ministers of health realized that it could be very expensive, that its components varied considerably, and that some were of questionable value. In 1976, the conference of Deputy Ministers of Health established a Task Force on the Periodic Health Examination, chaired by Dr. Walter O. Spitzer, then newly appointed to McGill University, to determine what services should be provided in the annual physical examination. Spitzer brought together epidemiologists, family physicians, general paediatricians, internists, and a psychiatrist. Experts in particular topics assisted in assembling evidence for the effectiveness of preventive procedures, which was then assessed by those without particular expertise in the topic. The Task Force established a set of rules for searching the available evidence, assessing its quality, and expressing judgements in simple terms. A Task Force declaration that not enough evidence was available became a stimulus to research. In this way the Task Force laid the foundations of what has now become known as evidence-based medicine.

The central recommendation of the Task Force's first report in 1976 was to abandon the annual check-up altogether. Because of this, it changed its name to the Canadian Task Force on Preventive Health Care. The report also recommended when and how to use a series of preventive interventions, including counselling on various topics, screening, chemoprevention (e.g., vitamin and micronutrient supplementation), and vaccination. All the recommendations were graded according to the strength of the evidence supporting them and the balance of risk and benefit.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

The Canadian Task Force on Preventive Health Care

The Canadian Task Force report was widely respected around the world and was copied a few years later in the U.S. The Canadian and U.S. groups worked together, producing a series of revisions and new topic reports. Even though the two groups studied the same evidence, their recommendations sometimes differed, reflecting differences in culture and context (see Chapter 14, section Influences on health policy making). As new evidence accumulated, the first report was updated continually until a second full report appeared in 1994. It was presented in a large red book, fondly known as ?the red brick'. The second report continued to be updated until 2006. Lack of funding caused the Task Force members to resign in 2005, but it was reformed in 2010. Its website is http://www.canadiantaskforce.ca/

A major obstacle facing the Canadian Task Force is its lack of influence on the health care system, so that, while its guidelines are appreciated, their implementation is hampered by lack of incentives and the difficulty of changing established practice patterns and professional roles. In recent years a number of changes in the system have improved the conditions for implementation. In particular, developments in physician remuneration mean that physicians are more likely to be rewarded for practising prevention. Also, the development of the multidisciplinary approach to care is likely to favour preventive practice.

Groups in other countries, similar to the Canadian Task Force, have continued to function including

- the US Preventive Services Task Force
- the National Institute for Clinical Excellence in the UK
- the Royal Australian College of General Practitioners "Red Book"

Similarly, in Canada and other countries many professional organizations are using the methods of evidence-based medicine to produce Clinical Practice Guidelines (CPG) on a range of topics.

Prevention in clinical practice

Clinicians are a resource for individual patients and for their practice population and community as a whole. As a resource for individual patients, they maintain and improve health by applying appropriate interventions that

- promote health
- prevent disease
- reduce the duration or severity of illness
- reduce disability and handicap.

Links

Levels of prevention

See chapter 4 for discussion of levels of prevention and how they fit into the natural history of disease.

As a resource for their practice population, clinicians can

- implement office systems that improve the practice of prevention
- implement continuous quality improvement (see chapter 13)
- make sure that their practice setting promotes health (e.g., provide bicycle stands, make sure the location is easily accessible on foot or by public transport and is accessible to people with physical handicaps, provide appropriate waiting-room reading material, provide no soft drinks or junk food outlets or vending machines)
- advocate for the health of their practice population. For specialists, this may include ensuring that the consultations they provide are accessible and oriented towards the needs of the population. Similarly, clinicians involved in teaching should ensure that their teaching takes account of population needs. For all physicians, it includes the promotion of equity and high quality preventive care.

When considering prevention in a clinical setting, we tend to think only of the interaction between a clinician and a patient. However, this interaction does not occur in isolation, but lies at the centre of a set of influences and constraints on the patient and clinician (see Figure 8.3). Some influences and constraints are more fully described in Chapter 2. To improve clinical preventive practice, the influence of many elements of the system should be recognized and their modifiability assessed.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

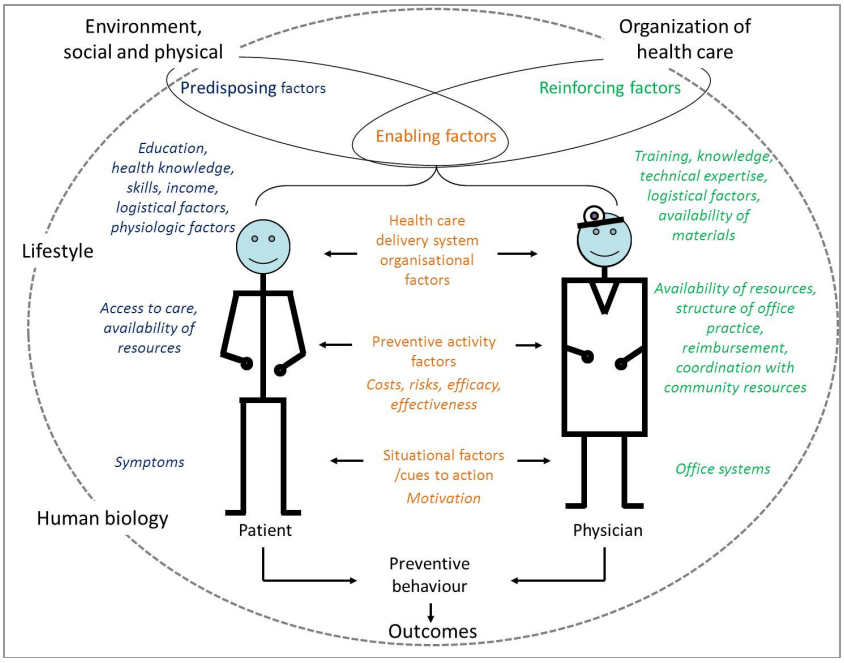


Figure 8.3: Influences on clinical practice (adapted from Walsh & McPhee[9])

Office systems for improving prevention in practice

The numerous influences on prevention mean that there is no single solution to improving preventive practice (see Figure 8.3).[10] Positive attitudes of clinicians and their staff?and knowledge of the evidence for the effectiveness of prevention?are essential. A range of information tools and prompts for patients and clinicians can help.[11] The following are some types of tools that can improve the practice of prevention.

- Preventive care flow sheets: kept in the patient's chart, these track the interventions and their outcomes.
- Reminders for patients: mailed to patient when an intervention is due.
- Visual prompts in office: for example, posters showing recommended interventions by age group.
- Patient information and patient-held preventive records: similar to care flow sheets, they give the patient the ownership of care.
- Chart reminders: can include coloured stickers indicating that patient has a specific risk behaviour or is at risk for a specific illness. The placement of stickers must respect patient confidentiality.
- Prevention prescriptions: these set out the objective to be achieved (negotiated with the patient), date for follow-up, and information about the changes to be made. A prescription reminds the patient what he has agreed to do.
- Health risk appraisal: questionnaire on health risks. Should be used in context; may be given to patients as they await their appointment.
- Computerized tracking systems: similar to flow sheets and chart reminders. Can be used to generate prevention prescriptions and patient-specific information.

Links

For more information on putting prevention into practice, please see *10 Steps: Implementation Guide. Put Prevention Into Practice*: http://www.ncbi.nlm.nih.gov/bookshelf/br.fcgi?book=hsarchive&part=A8916

Readers of French can also look at Description, impact et conditions d'efficacité des stratégies visant l'intégration de la prévention dans les pratiques cliniques : revue de la littérature http://msssa4.msss.gouv.qc.ca/fr/document/publication.nsf/ff52dbec0b2ed788852566de004c8584/d93931b92d211699852573d2005375e8?OpenDocument

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

Precede-proceed model for planning and evaluation[12]

A way of deciding where to start putting prevention into practice and how to evaluate actions is the Precede-Proceed model developed by Lawrence Green.

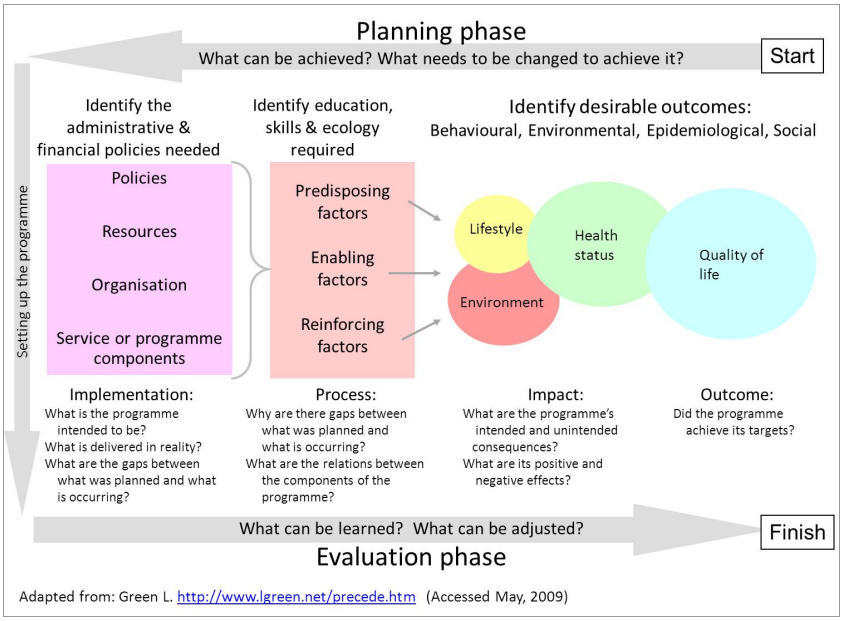


Figure 8.4: Precede-Proceed model for planning community health promotion programmes

In re-orienting clinical practice, the Precede-Proceed model is a valuable resource for a clinician planning a project. It lists a range of factors that should be considered in programme planning, based on empirical evidence describing the community that the physician serves. The idea is that the planning phase begins at the right of the model by assessing quality of life, the health status that influences it, and lifestyle and environmental factors that influence the health status. This guides the development of objectives for improving health status and quality of life of patients living in this community (blue arrow). The predisposing, enabling, and reinforcing factors that promote or inhibit the achievement of these objectives can be listed as targets that should be addressed by a health programme, and from this the policies, resources, and organization of services and programmes required to achieve these targets can be assessed. In this way a complete portrait of the problem is drawn up along with a plan for implementing solutions. Once a programme has been set up, it can be evaluated. For this, the steps of the evaluation phase are followed in the opposite direction to see if the planned action took place, if targets were achieved, and if the expected outcomes followed (green arrow).

The Precede-Proceed model builds on the explanatory models of health behaviour described in Chapter 2 and illustrates the broad range of factors that can account for preventable ill health in individuals. Note that many of these factors are outside the control of the patient and the clinician, which is why a community-level intervention becomes so important.

Case Study

Dr. Rao puts a sticker on David's chart to remind himself that David should be counselled on his risk factors. He asks David to see Nurse Jennings, who enquires further about his use of tobacco and marijuana. David is reluctant to change his smoking habit: he doesn't feel that it is doing him any harm. Nurse Jennings suggests that the cough and suspicion of asthma may be related to smoking and gives David a leaflet about smoking that, in particular, addresses young people's concerns. Similarly, David feels his marijuana intake is not a problem: he thinks he can drive after a joint, which he can't after alcohol. Most of his friends smoke marijuana without apparent problems.

Changing behaviour

Medical intervention usually requires that the patient change his behaviour. Most often, the change is a simple one such as taking a pill every day, but sometimes it means changing a lifestyle pattern, which can be much more difficult. Current behaviour has often been shaped by years of living a certain way within a certain social circle and changing it can have untoward consequences. For instance, quitting smoking means that the patient gives up the social contact with other smokers; reducing alcohol intake can reduce the pleasure of the Hockey Night in Canada session with friends; changing diet may mean no more convenient stops at fast food outlets.

There are many theories that attempt to explain how health behaviour is shaped (see Chapter 2). Prochaska and DiClemente combined a number of them to develop their transtheoretical model of the stages through which behaviour changes, and this is particularly useful in clinical encounters.[13] The model was originally formulated to describe a person's level of readiness for smoking cessation and describes the stages through which a patient typically passes in adopting a new behaviour and identifies factors that may motivate or impede such progression. Using the transtheoretical model, a clinician can diagnose which stage a patient has reached and, drawing on knowledge of the various forces at play during each stage, he can tailor his intervention to provide "stage-appropriate" support for behaviour change.

The Stages of Change model states that at any time, for any behaviour, a person is in one of the following stages of behaviour change:

1. *Precontemplation*: People in this stage are not intending to change their behaviour (e.g., David Richards may have no interest in quitting

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

smoking). For practical purposes, the time horizon is defined as not planning to change within the next six months.

2. *Contemplation*: The individual has not begun to change their behaviour, but is thinking about it and intends to do so within the next six months (e.g., David has now talked to Dr. Rao about stopping smoking).
3. *Preparation*: The individual has not begun to change his behaviour, but intends to do so in the next thirty days (e.g., has set a quit date).
4. *Action*: The individual has changed his/her behaviour within the past 6 months (e.g., has quit).
5. *Maintenance*: The individual has practiced this new behaviour for 6 months.
6. *Relapse:* Patients often find it hard to maintain the new behaviour and relapse. This may lead them either to abandon the idea of changing, so revert to precontemplation, or else stimulate them to try again, and so re-enter the contemplation or even the preparation stage, illustrated in Figure 8.5.

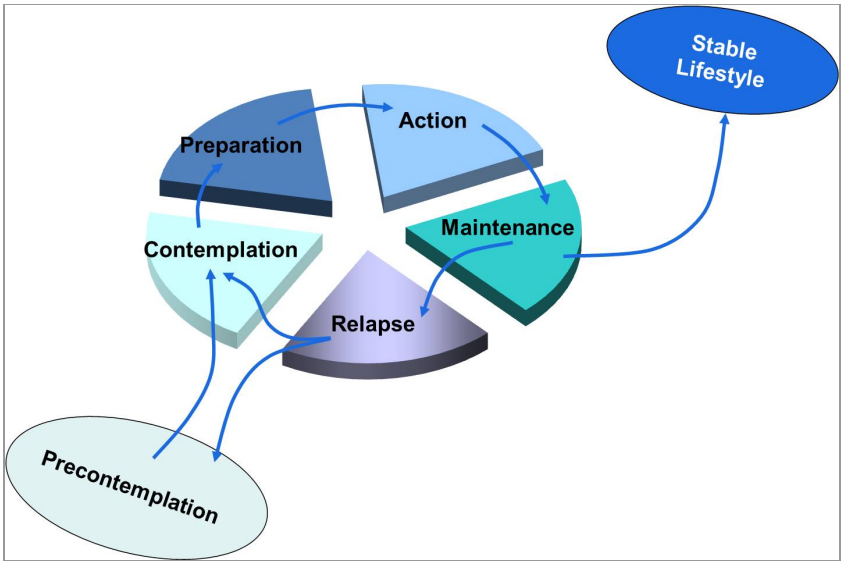


Figure 8.5: Illustration of the Transtheoretical Model of Behaviour Change.

The clinician's role will vary across these stages, and this is illustrated in Table 8.4.

Table 8.4: Stages of change and how to manage them as applied to exercise (adapted from Zimmerman, Olsen, Bosworth[14])

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

	Stage	Explanation	Therapeutic Goal	Tips for the clinician
1.	Pre-contemplation	The patient doesn't see the need to change: "My uncle was obese and hated exercise. He lived until he was 90.	To encourage patient to start thinking about changing	▫ Encourage the patient to consider change ▫ Personalize health information: "A person like you has a one in five chance of developing diabetes." ▫ Highlight any symptom that might be related to the behaviour. Express concern, but avoid scare tactics ▫ Get the patient on a positive train of thought: "What would you see as the advantage of changing?"
2.	Contemplation	Thinking about change, weighing the costs and benefits: "I know I should, but it would take too much time."	To encourage patient to evaluate the benefits and barriers in a balanced fashion	▫ Ask patient to list benefits of, and barriers to, change. Ask for solutions?one barrier at a time. Explore ambivalence ▫ Encourage the patient to consider planning for a change. Push gently?rather than "you should try this," say "some people find this helpful; you might too."
3.	Preparation	Making small changes: "I've bought a pair of jogging shoes, but. . ."	To encourage patient to develop a plan for change	▫ Encourage action. Put the patient in touch with programmes that can help them ▫ Ask about setting a precise date for change. Help invent strategies for handling challenges the patient will face.
4.	Action	Taking definitive action: "I went jogging this morning."	To encourage patient to put the plan into action	▫ The physician's role is to supply encouragement and support. Congratulate and encourage small successes. Ask about problems encountered and review solutions. Schedule follow-up visits
5.	Maintenance	Maintaining new behaviour: "I've been exercising regularly for about 6 months now."	To help new behaviour to become ?normal.'	▫ Continued encouragement ▫ Ask about what the patient will do if tempted to relapse
6.	Relapse	"The weather got bad and I just couldn't continue."	To re-engage the patient in the process	▫ Ask about lessons learned from previous change. Reformulate failure into partial success. Remind patient that relapse is a normal part of the process of change

Understanding the process of behaviour change and the influences on it avoids "blaming the patient" (see Chapter 1) for not adhering to health recommendations. Clinicians who feel frustrated that a patient is not changing should ask themselves if their diagnosis of the patient's stage of change was correct, and if their intervention was stage-appropriate. Note that counselling can be an opportunity to learn from patients. The clinician can ask patients about the tricks they used to make and maintain the change and then pass these ideas on to other patients in similar situations.

Since its original publication, the transtheoretical model has been refined. Figure 8.6 compares the original to a later adaptation by Weinstein.[15]

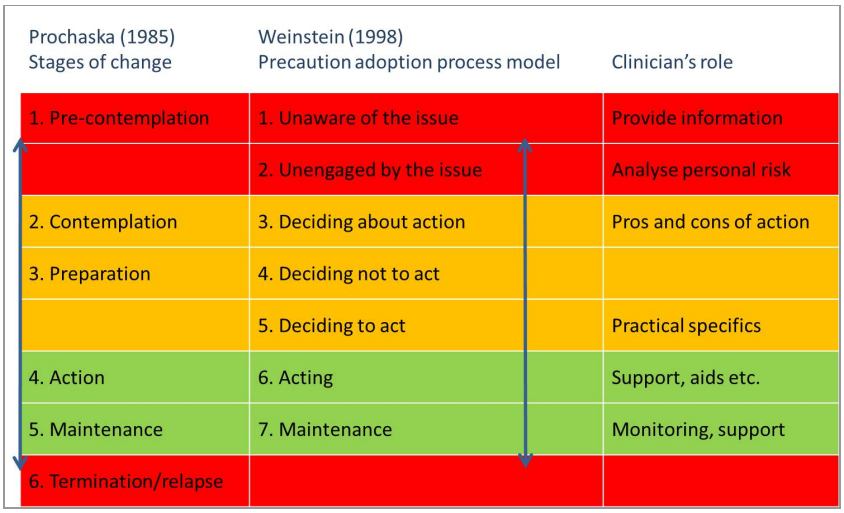


Figure 8.6: Two models of behaviour change and the clinician's role.

The arrows indicate that patients can progress through the stages in either direction. The Physician's role is to help the patient progress to ?Maintenance'.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Case Study

Nurse Jennings gives David tips on the safe use of marijuana and alcohol, mentioning that the effects of marijuana can last for a day or more depending on the dose and that during this time driving can be impaired. She also talks about safer sexual practices. She asks about exercise. David mentions that he used to enjoy basketball in high school. Nurse Jennings encourages him to take it up again. She remembers that another patient said that the local YMCA was a good starting point for getting into sports. She gives David some information leaflets on it. Later, during lunch break, Dr. Rao and Nurse Jennings discuss if there is anything that can be done to help patients get as physically active as they should be to protect their health. Nurse Jennings says that it's best when people can incorporate activity into their usual everyday tasks rather than having to make time for it. She asks Dr Rao if it would be possible to put a sign on the elevator door encouraging people to use the stairs and indicating where they are. Dr Rao makes a mental note to use the stairs more often himself and wonders what else the clinic can do to encourage healthy behaviour.

Nerd's Corner

Difference between prevention and treatment?

Chapter 4 introduced the stages of prevention, but made the point that these are arbitrary distinctions that only have heuristic value. Chapter 1 showed that the definition of ?disease' is also arbitrary. Many experts likewise believe that the distinction between disease and risk factors is artificial: it depends on the point in the disease's natural history that is being considered. Tobacco addiction, hypertension, and diabetes can be considered diseases in themselves, but they are, in turn, modifiable risk factors for chronic lung disease, stroke and renal disease, respectively. Even surgery for subarachnoid haemorrhage does not cure the haemorrhage or its effects; it merely attempts to prevent another.

The WHO's definition of prevention implies that all clinical interventions should prevent something: the occurrence, progression or duration of disease, or its resulting disability or handicap.[16] The risks and benefits of intervening versus not intervening should be assessed and discussed with the patient in every case. Perhaps the main difference between ?prevention' and ?treatment' is the fact that preventive interventions are applied to patients who do not have the targeted disease and who have often not solicited the intervention. Intervening in an ostensibly healthy patient, especially one who has not asked for it, adds to the importance of the clinician taking care to minimize harm, maximize benefit, and ensure informed consent.

Harm reduction

Sometimes changing risky behaviour proves extremely difficult: addiction offers a common example. In this instance, the physician can consider a harm reduction approach. Harm reduction is based on the idea that risk-taking is a natural part of our world. It can be difficult?even impossible?to stop a risky behaviour, but its harmful effects can be reduced. The approach is usually associated with the management of substance abuse, but it can be useful in other fields. Providing clean injection equipment to intravenous drug users to prevent sepsis and transmission of blood borne disease is an example. Providing methadone maintenance for opiate users to reduce the risks associated with illicit drug use is another. A population-level example is Operation Red Nose which, by providing people with a lift home after holiday parties, aims to reduce the road traffic accidents associated with alcohol instead of aiming to stop people drinking.

Definition

Harm reduction refers to reducing the negative consequences of risk behaviours, rather than trying to eliminate the behaviours themselves (e.g., smoking only outdoors rather than not smoking). It can be defined as a set of practical strategies that help people to reduce harm associated with engaging in risk-taking behaviour (adapted from Harm Reduction Coalition, United States, 2000; definition supported by Canadian AIDS Society, 2000).

Harm reduction has evoked significant debate and is often dismissed by people who oppose the behaviours. For example, many people oppose placing condom vending machines in high school washrooms on the grounds that they condone, or even encourage sexual activity. Similarly, providing safe injection sites and clean injection equipment may be construed as condoning illegal drug use. In particular, provision of clean equipment in prisons is seen as an admission that efforts to keep drugs out of prison are a failure.

From a health perspective, however, if actions aimed at eradicating a behaviour have repeatedly failed, it is logical to implement evidence-based approaches that reduce the harm associated with it. This does not mean that attempts to reduce the risky behaviours themselves should be abandoned. Recent comprehensive programmes view harm reduction as one of four essential components in reducing the burden of unhealthy behaviours: prevention, treatment, harm reduction, and enforcement.[17]

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Illustrative Materials

Politics and harm reduction

In 2003, North America's first and only supervised injection facility, Insite, was established in Vancouver's Downtown Eastside, one of Canada's poorest neighbourhoods. Insite was designed to increase access to health care and addiction services, to reduce overdose-related deaths and reduce the transmission of blood-borne infections.[18] There are an estimated 12,000 injection drug users in Vancouver, one third of whom live in the Downtown Eastside.

Insite users bring their own drugs to an injection booth where they are provided with clean injection equipment, including needles. Nurses, trained to respond to overdoses, supervise the injections. Insite also provides onsite addictions counselling and referral to detoxification facilities and other forms of community support, as well as providing a point of access to the health care system for drug users who are not otherwise well connected.[18]

The facility has been successful in encouraging its patrons to use detoxification services. It has reduced the number of deaths from overdose and it is associated with a decrease in high-risk behaviour, including needle sharing, though there has been no documented decrease in transmission of HIV or HCV. Insite has also led to improved public order, with decreased rates of public injecting and less injection-related litter. Its critics argue that safer injection facilities will lead to an increase in drug-related crime and promote drug use in the community, but these claims are not supported by the current evidence.[19,20]

Despite the growing body of evidence supporting safer injecting facilities, the federal government has remained opposed to them. Insite operates under a federal exemption of the national drug policy. In 2006, when the pilot phase of Insite was due for renewal, the federal government threatened to withdraw this exemption and close down the facility. However, local community-based organizations, with the support of the scientific and medical community, took the federal government to court, stating that federal drug policy was inconsistent with the Charter of Rights and Freedoms. On May 27th 2008, a provincial judge found in their favour and Insite remained open. Since this time, the BC Court of appeal has upheld this ruling and a Canada Supreme Court ruling is pending.[21]

To read about Vancouver's safe injection Insite, visit the Vancouver Coastal Health website at: http://supervisedinjection.vch.ca/

Further Reading

Harm reduction and Vancouver's four pillars

You can see more about Vancouver's "four pillars" programme for drug abuse at the Four Pillars Drug Strategy website: http://vancouver.ca/fourpillars/

For more on Harm Reduction see:

http://www.medicine.uottawa.ca/sim/data/Harm_Reduction_e.htm

http://www.medicine.uottawa.ca/sim/data/Harm_Reduction_f.htm

Nerd's Corner

Using a market model to reduce substance abuse

Supply reduction (e.g., destroying crops from which illicit drugs are derived or interrupting the shipment of drugs). This approach has been tried for years, but it is expensive, tends to cause an escalation of violence and may not be cost-effective. Furthermore, underlying determinants such as poverty and corruption in the supplying countries continue to drive the supply side of the equation.

Price increases are effective in reducing the use of legal substances, such as cigarettes. It is somewhat analogous to supply reduction in that both reduce access to the substance. Selling the substance only in large quantities also reduces access, by increasing the asking price?for instance, in Canada cigarettes are sold only in quantities of 20 or more. Pricing is particularly effective in limiting consumption among youth, who generally have less disposable income. Note that tobacco manufacturers have bypassed legislation on the quantity of cigarettes that can be sold by promoting the use of cigarillos to young people. As these are not technically cigarettes, they can be sold singly. The price of illegal substances is outside legal control, although supply reduction activity drives prices up. Patterns of consumption of illegal substances are generally dictated by their street price.

Demand reduction focuses on the other side of the equation: discouraging people from wanting illicit drugs. This is usually attempted via education and information, community programmes, or legal penalties. Research suggests that these approaches may work best for those who depend least on drugs and is less effective in the high risk population.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Harm reduction advice for people who use recreational drugs[22]

For users of any drug:

- Don't mix substances
- Drink plenty of water but not to excess
- Use drugs only in a safe place with people you trust
- Ask people to intervene if they notice anything wrong
- Eat well, get enough sleep, and exercise regularly
- Know the effects of the substances you take
- Know where you can go for help.

Additional advice for users of intravenous drugs:

- Make sure you are up to date with hepatitis A and B vaccinations
- Try non-intravenous routes of administration (by mouth, sniffing etc.)
- Reduce frequency or quantity of use
- Avoid overdose: inject slowly, pay attention to the quality and purity of the drug
- Use safe injection practices and don't share equipment
- Bring used needles to a drop box.

Health promotion

Health promotion, also known as primordial prevention aims to help people increase control over and improve their health. Although it may have an effect on specific diseases, it has a more general aim: to enhance health in order to develop the person's resistance to the adverse influences of physical and social environments. It can be carried out directly on an individual level or indirectly at a population level. Population health promotion seeks to influence the determinants of health, using community action, lobbying, or the publication of information to induce change in public policies, social and physical environments, and health services. Such changes filter down to the individual level, creating contexts that support the development of personal skills and health habits. In addition, health promoters often use interventions, such as health information campaigns, that target individuals directly. The integrated model of population health promotion (Figure 8.7) demonstrates the main dimensions that need to be considered when planning a health promotion intervention. It asks questions such as: Who needs the intervention? What should the intervention target? How should the objectives be achieved?

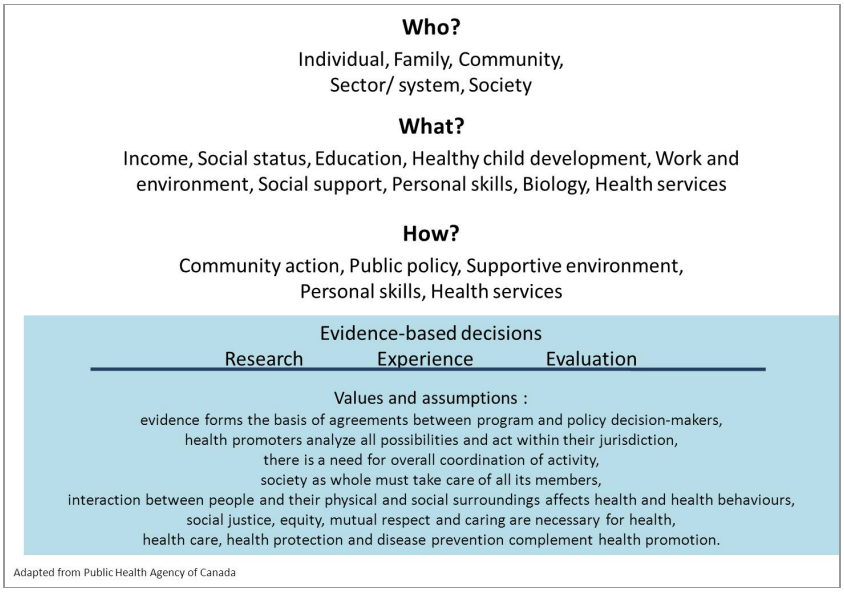


Figure 8.7: The three dimensions of health promotion[23]

Who needs the intervention?

This is not simply a matter of who has a problem; it asks at what level, from individuals to society, intervention is likely to be most effective and efficient. For example, should school boards, parents, or their children be targeted in encouraging good adolescent nutrition? For some health

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

problems, intervention at a population level is not feasible: the causative factors may be unclear, the disease may be extremely rare, or the intervention's cost-benefit ratio may be unfavourable. For other problems, population or community-level interventions are possible and likely to improve the situation more than individually targeted interventions. For instance, legislation on tobacco succeeded in reducing the level of tobacco use more than could have been hoped for by counselling individuals. In most cases, interventions aimed at more than one level are necessary and can mutually reinforce each other. For instance, to reduce the prevalence of the metabolic syndrome, clinicians should advise on diet and exercise, and communities should provide access to healthy foods, and the infrastructure for facilitating exercise, such as local exercise facilities or incentives for active transport. Meanwhile, societies should assess the balance between the economic good (jobs, cheap food) produced by the food industry and the health of the population.

Definition

Active Transport is the use of a mode of transport that involves physical effort on the part of the user. It includes walking, running, and cycling. It also includes using the stairs instead of the elevator or escalator. Public transport is more active than private automotive transport, because it generally entails walking to and from the transport stop and the destination.

What should the intervention target?

Clinicians are ideally placed to deal with the health of individuals. They are trained to deal with pathological processes and to advise on personal health, health practices, and ways of coping with illness. However, societal and environmental determinants also greatly affect the health of individuals. Outside of offering health services, physicians may need to raise awareness of particular health problems and lobby for solutions: thereby taking on the health advocate role in the CanMeds framework. To choose priorities for disease prevention, it can be helpful to consider individual and population approaches for the different levels of prevention.

How can health be achieved?

To achieve population health objectives, different strategies must be employed. For instance, mass publicity campaigns alone have little impact in changing smoking rates. Different types of laws designed to reduce smoking are also required: laws that ban smoking in public and that reduce the visibility of tobacco products lower its social acceptability; those that raise the price of tobacco products are effective in reducing smoking among youth. Prolonged effort in lobbying and awareness-raising is necessary to get such laws passed.

Similarly, when dealing with individual patients, clinicians must use different strategies to help patients achieve the best health possible. For instance, in the management of obesity, in addition to assessing the patient's stage of change and discussing the problem and a target weight with the patient, the clinician could suggest that the patient join a weight loss programme, that the patient's family be enlisted to provide heart healthy meals, and that the patient begin a graduated exercise programme that incorporates exercise into the daily routine by using active transport.

Self-test questions

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Self-Testing

1. Name the stages of change and suggest how to recognize each one according to what patients say.

Stage	A patient might say
Pre-contemplation: The patient doesn't believe that he is at risk and sees no need to change.	"My uncle never (always) did and lived until he was 90."
Contemplation: The patient is aware of the risk and how to avoid it, but still lacks motivation to change. Typically weighing the costs and benefits.	"I know I should, but ... "
Preparation: The patient is beginning to make small changes, the plans are more concrete.	"I've bought a pair of jogging shoes (supply of chewing gum, a bus pass, recipe book...), but... "
Action: Concrete plan with definite date in the near future.	"I'm starting tomorrow."
Maintenance: The patient has made the change and is maintaining it, although it might not be fully integrated as "usual" behaviour.	"I've been sticking to it for a month now."
Relapse. After behaving differently for a time, the patient falls into his previous behaviour patterns.	"I was doing well, but... "

2. What are the major influences that determine if an intervention will be applied?
In clinical situation, some factors influence the clinician, others influence the patient and yet others influence both.
Patient factors: personal factors, such as the patient's education, health knowledge, and income.
Physician factors: the physician's training and technical expertise.
Health care delivery system factors: the costs, risks, effectiveness and acceptability of the intervention itself.
Situational factors: cues to action during the consultation influence intervention.
(See Figure 8.3 adapted from Walsh and McPhee 1992)

3. What are the advantages and disadvantages of reducing population risk as opposed to targeting high-risk individuals?
Advantages of a population strategy
Population strategies aim at the roots of the problem. As such, they reduce illness in the whole population, including those at low- or average risk. In a number of conditions, the population at low- or average-risk produce more cases of illness than the population at high risk (for example, see Figure 8.1), so that a population strategy may prevent more cases than strategies aimed at high risk individuals.
A small change in the level of a risk factor in a population can result in a large change in outcomes.
By emphasising the upstream, situational causes of the problem rather than individual factors, population strategies reduce the likelihood that individuals will feel blamed for their behaviour, and they provide a more sustainable outcome. Population strategies aid in making the desired behaviour "normal". This encourages its adoption. They also eliminate the need to screen and identify a "high risk" group. This means that people at low or average risk are not falsely reassured and people at high risk are not needlessly made anxious.

Disadvantages of a population strategy
The small benefit to most individuals can be outweighed by the risk of the intervention, even if this is also small. As an example of the risk of intervention, recommendations about reducing fat in the diet have led some mothers to feed their baby low-fat milk in the mistaken belief that this is good for the child, who actually needs full-fat milk.
The lack of obvious personal benefit can reduce acceptability to the individuals who make up the population, particularly among those who consider themselves at low risk.
There is a danger of increasing health inequities: resources may not be directed to people most in need. Unless specifically designed strategies are used, vulnerable people in most need of change are least likely to do so. When a general population strategy is used, it is often the more educated and informed who follow the recommendations, that is those usually at least risk.
Interventions that affect apparently healthy people are open to ethical questions and they may be hard to justify politically. Interventions that affect people who have not consented to the intervention can be seen as social engineering, which is unacceptable in a liberal society.

4. What are the basic values of a health-promoting approach?
According to the Ottawa Charter (see figure 8.7) the values of a health promotion approach are the following:
- *evidence forms the basis of agreements between programme and policy decision-makers;*
- *health promoters analyze all possibilities and act within their jurisdiction;*
- *there is a need for overall coordination of activity;*
- *society as whole must take care of all its members;*
- *interactions between people and their physical and social surroundings affect health and health behaviours;*
- *social justice, equity, mutual respect and caring are necessary for health;*
- *health care, health protection and disease prevention complement health promotion.*

5. Give an example of a primary preventive intervention recommended in clinical practice.
Consultation of guidelines such as those produced by the US task force on preventive care produce a number of such interventions, such as counselling against the use of tobacco, counselling for the use of seat-belts, or prescribing folic acid to women wishing to become pregnant. Primary prevention refers to interventions that are offered before disease occurs.

6. Give an example of a primary preventive intervention that can be applied to a population.
A wide variety of primary preventive interventions have been applied to populations. Examples include passing laws on seat-belts and on bicycle and motorcycle helmets, as well as those restricting advertising, selling and consumption of tobacco products. Many municipal activities in regulation of housing, water supply and sewage, as well as programmes of food inspections also contain elements of primary prevention. Activities that promote healthy lifestyles, such as the construction of cycle paths and the offer of healthy cooking classes in impoverished areas are others.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Discussion Points

1. In your experience of clinical situations, what are the barriers to primary preventive practice?

2. In your experience of clinical situations, what increases the practice of prevention?

3. What information resources are available to clinicians to support the choice of intervention to reduce risk?

References

Further Reading

Morabia A, Costanza MC. What is this thing called preventive medicine (III)? Preventive Medicine. 2008;47:151-153.

Rose G. The strategy of preventive medicine. Oxford: Oxford University Press; 1992.

Kue Young T. Planning health: concepts and methods. Oxford: Oxford University Press; 2005.

Spasoff R. Policy choices. epidemiologic methods for health policy. New York (NY): Oxford University Press; 1999.

1. Rose G. Sick individuals and sick populations. Int J Epidemiol 1985;14:32-38.

2. Manuel DG, Rosella LCA, Tuna M, Bennett C. How many Canadians will be diagnosed with diabetes between 2007 and 2017? Assessing population risk. 2010 June(July).

3. Cook NR, Cohen J, Hebert PR, Taylor JO, Hennekens CH. Implications of small reductions in diastolic blood pressure for primary prevention. Arch Intern Med 1995 Apr 10;155(7):701-709.

4. Manuel DG, Lim J, Tanuseputro P, Anderson GM, Alter DA, Laupacis A, et al. Revisiting Rose: strategies for reducing coronary heart disease. BMJ 2006 Mar 18;332(7542):659-662.

5. Alberman E, Berry C. Prenatal diagnosis and the specialist in community medicine. Community Med 1979 May;1(2):89-96.

6. Frohlich K. Is tobacco use a disease? CMAJ Canadian Medical Association Journal 2008 Oct 21;179(9):880-882.

7. Charlton BG. A critique of Geoffrey Rose's 'population strategy' for preventive medicine. J R Soc Med 1995 Nov;88(11):607-610.

8. Fries JF. Aging, natural death, and the compression of morbidity. N Engl J Med 1980 Jul 17;303(3):130-135.

9. Walsh JM, McPhee SJ. A systems model of clinical preventive care: an analysis of factors influencing patient and physician. Health Educ Q 1992;19(2):157-175.

10. Oxman AD, Thomson MA, Davis DA, Haynes RB. No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. Canadian Medical Association Journal 1995 Nov 15;153(10):1423-1431.

11. Gemson DH, Ashford AR, Dickey LL, Raymore SH, Roberts JW, Ehrlich MH, et al. Putting prevention into practice. Impact of a multifaceted physician education program on preventive services in the inner city. Arch Intern Med 1995 Nov 13;155(20):2210-2216.

12. Green LW, Kreuter MW. Health promotion planning: an educational and ecological approach. 3rd ed. Mountain View, CA: Mayfield; 1999.

13. Prochaska JO, DiClemente CC. Stages of change in the modification of problem behaviors. Prog Behav Modif 1992;28:183-218.

14. Zimmerman GL, Olsen CG, Bosworth MF. A 'stages of change' approach to helping patients change behavior. Am Fam Physician 2000 Mar 1;61(5):1409-1416.

15. Weinstein ND, Lyon JE, Sandman PM, Cuite CL. Experimental evidence for stages of health behavior change: the precaution adoption process model applied to home radon testing. Health Psychology 1998 Sep;17(5):445-453.

16. World Health Organization. Health Promotion Glossary. 1999(September).

17. City of Vancouver. Four Pillars Drug Strategy. Available at: http://vancouver.ca/fourpillars. Accessed September, 2009.

18. Vancouver Coastal Health. Supervised Injection Site Website. Available at: http://supervisedinjection.vch.ca/. Accessed October 7, 2010.

19. Kerr T, Wood E, Montaner J, Tyndall M (editors). Findings from the Evaluation of Vancouver's Pilot Medically Supervised Safer Injecting Facitility - Insite. British Columbia Centre for Excellence in HIV/AIDS ed.; 2009.

20. Wood E, Tyndall MW, Montaner JS, Kerr T. Summary of findings from the evaluation of a pilot medically supervised safer injecting facility. CMAJ Canadian Medical Association Journal 2006 Nov 21;175(11):1399-1404.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

21. Dooling K, Rachlis M. Vancouver's supervised injection facility challenges Canada's drug laws. CMAJ 2010 Sep 21;182(13):1440-4.

22. Santé publique Santé mentale et dépendances. Prévention en pratique médicale : Alcoolisme et toxicomanies. Available at: http://www.santepub-mtl.qc.ca/mdprevention/fiches/alcool/evaluer.html, December 2009.

23. Hamilton N, Bhatti T. Population Health Promotion: An Integrated Model of Population Health and Health Promotion. Available at: http://www.phac-aspc.gc.ca/ph-sp/php-psp/index-eng.php. Accessed September, 2009.

Chapter Objectives

After reading this chapter you will be able to:

- principles of screening and be able to evaluate the utility of a proposed screening intervention, including being able to discuss the potential for lead-time bias and length bias;
- Demonstrate an understanding of ethical issues involved in resource allocation.

Case Study

Paul Richards consults Dr Rao because his back has been causing him pain recently and he wonders if he needs an X-ray. In passing, he heard about the "prostate test" on the radio and he wants to know if he should get one. After taking a history and doing an examination for back pain, Dr Rao advises Paul to continue his normal activities and, if possible, to get more exercise. He does not find any indication for imaging and explains this to Paul. He checks the clock and wonders if he has time to discuss prostate screening or if he should ask Paul to arrange another appointment to do so.

Goal of screening

Screening aims to detect disease early, before it becomes clinically manifest. In formal terms it is the application of a test or procedure to asymptomatic people in order to classify them into two groups: one at high risk of developing clinical disease, the other at low risk of developing it in order that people at high risk can be treated to prevent the effects of the disease. The assumption underlying screening is that disease diagnosed early in its development responds better to treatment than it does once it has started causing signs and symptoms, so that early diagnosis results in better quantity and quality of life for the patient. However, while screening appears to make sense, it has to be considered very carefully. It represents an intervention on a patient who, as he may not have the disease being screened for, may not benefit directly, so the clinician has a responsibility to ensure that any harm is even more unlikely. Growing awareness of this duty has resulted in a number of tests that were once promoted being subsequently dropped: screening does not always meet expectations.

Nerd's Corner

Screening failures

Here is a list of screening tests that have been discarded:

- Urine dipsticks for renal disease
- Examination for heart murmurs in children
- Wassermann test (for syphilis) before marriage
- Regular electrocardiograms
- Homocysteine for heart disease
- CA 125 for ovarian cancer
- Regular chest x-rays for lung cancer
- Breast self-examination
- Self-examination for testicular cancer
- Infrared temperature sensors at airports for SARS.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Here Be Dragons

The aim of screening

Screening generally falls into the category of secondary prevention?detection of disease at an early stage, so permitting an intervention while the disease may still be mitigated. A screening test indicates the probability of a patient having the disease, yet some chronic diseases uncovered by screening undergo spontaneous remission and others progress so slowly (as with some prostate cancers in elderly men) that they are unlikely ever to harm the patient. As a result, we consider that the goal of screening is to distinguish people who are at risk of suffering the harmful effects of a disease from people not at risk of this in order to be able to offer effective treatment only to those at risk.

A large number of people have an interest in screening. Those who carry out the screening and those who supply screening machinery have a financial interest in screening, which can bias their opinion of it. Those with the condition being screened for may wish others to avoid their experience and so advise people to be screened. Health service managers, responsible for controlling costs, may worry that the cost of screening will drain their budget and divert funds away from other pressing demands. For instance, cervical cancer screening may be a drain on the primary care budget, while any savings produced by preventing cancer will be seen in the cancer care budget. With each stakeholder using evidence to defend her point of view, the decision to implement population screening programmes is often controversial.

Definition

Screening terminology

Mass or universal screening is aimed at a population defined demographically. It can be an organised programme such as the provincial newborn screening programmes or the Canadian breast cancer screening programme. Universal screening is not always formally organised; it may occur when individual clinicians respond to clinical practice guidelines. Clinicians can set up **systematic screening** for their practice population, for instance by implementing an office system to ensure that all the practice's adult patients have regular blood pressure checks. Many clinicians practise **opportunistic screening** using any consultation as an opportunity to carry out appropriate screening interventions.

Case-finding also aims to diagnose pathological processes in the asymptomatic stage, but the population targeted is usually defined by risk factors other than demographic. For instance, after discovering that a patient has familial hypercholesterolemia, a clinician may screen others in the patient's family. Similarly, patients with diabetes mellitus should be screened for retinopathy.

The term case-finding is also used:[1]

- 1. In infectious disease control to mean a standard procedure whereby diligent efforts are made to locate and treat persons who have had close or intimate contact with a known case. This is also known as **contact tracing**,
- 2. In epidemic control to mean seeking people who have been exposed to a risk or to potentially harmful factors.

Issues in deciding whether or not to screen

The difference between screening and diagnostic testing can be defined merely in terms of differences in the risk of the disease for which a patient is being screened or tested. In screening, the patient being screening is asymptomatic and at apparently low risk of the disease being screened for. In diagnosis, the patient has symptoms or signs indicating a high likelihood of the disease being tested for. Due to the similarity of the two situations, some of the issues mentioned below for screening are also relevant to diagnostic testing.

Identifying who will develop clinical disease

Few indicators of pre-clinical disease are dichotomous?that is, few differentiate sharply between who will develop the condition and who will not. Blood sugar, cholesterol, and blood pressure are all continuous scales. Based on the test sensitivity and specificity (sometimes combined using the Receiver Operating Characteristic (ROC) curve) experts decide on a point on the scale that distinguishes those who are at high risk of developing the clinical condition from those who are at low risk. Even in cancer screening, although some findings on imaging or cytology can be categorized definitely as neoplastic or as benign, other images or specimens show intermediate changes for which even experienced readers may have differing interpretations. Furthermore, some people develop disease even though the screening results put them in the unlikely group (false negatives), and some people, apparently at high risk, never develop disease (false positives). If the cut-off point is set too far towards the high risk end of the scale, too many patients with early stage disease will be missed, so the sensitivity and the negative predictive value will be low. If the point is too far towards the low risk end of the scale, the costs of investigating and reassuring large numbers of low-risk patients will outweigh the benefits of treating patients early who otherwise would have developed a clinically significant condition, and the specificity and positive predictive value will be low.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

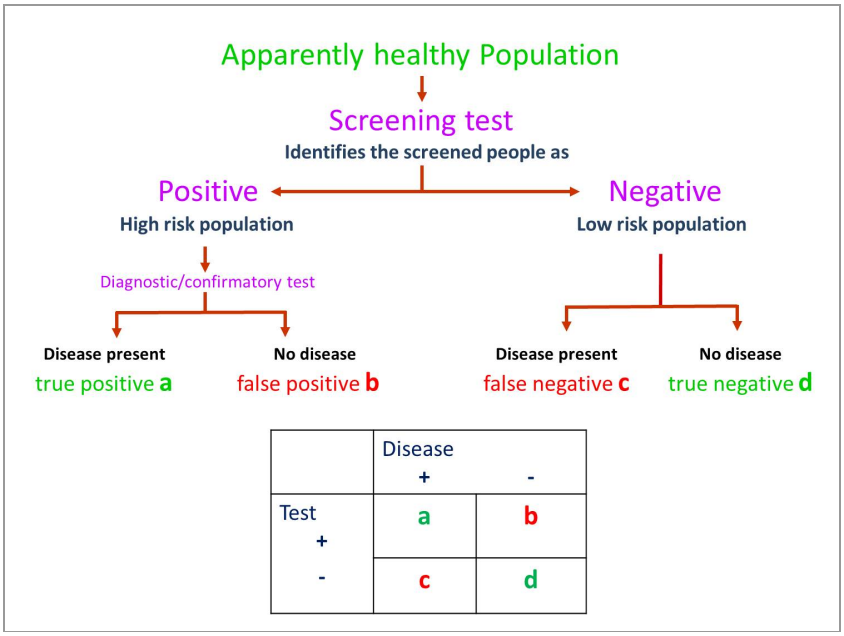


Figure 9.1: Screening test performance

No test is 100% specific and 100% sensitive, so there will always be false negatives and false positives. Furthermore, the prevalence of the illness under study affects the positive predictive value of the test. Apparently healthy populations tend to have a low prevalence of the disease being screened for; therefore, the positive predictive value of screening tests tends to be less than that of diagnostic tests.

Because it is often hard to detect disease in its early stages, only tests with excellent performance characteristics (sensitivity and specificity) can be used for screening. In order for mammography to be acceptable as a screening test, technicians and radiologists, many of whom were experienced in diagnostic mammography, needed further training to ensure that the performance characteristics of their screening matched those in breast cancer screening trials.

Overdiagnosis

Because it detects disease at a pre-symptomatic stage, that is before it causes perceivable damage or disability, screening may identify disease that would never harm the patient. While Chapter 1 presented the idea of **natural history** as though disease progresses steadily, this is not always the case. Disease may regress because of the patient's natural defences. For example, about two thirds of cases of mild cervical dysplasia regress to normal, whereas progression from mild to severe dysplasia or worse occurs at a rate of about 1% per year.[2] Slow progression may mean that the patient dies from other causes before succumbing to the cervical cancer detected by the screening. There is, as yet, no method to predict what will happen in individual cases. The old adage that "more men die with prostate cancer than die of prostate cancer" is true. The prostate specific antigen (PSA) test is only about 70% sensitive but is still useful in detecting early cases of prostate cancer.[3] However, autopsy series find that approximately 30%?40% of men over the age of fifty have prostate cancer?a proportion that rises with age?yet only about 3% of men die from it. It is not yet certain what should be done once early prostate cancer has been detected. Similarly, studies of breast cancer screening have shown that up to 30% more breast cancers are diagnosed in the screened group than in the unscreened group, probably due to the detection of cancers that would never otherwise have become apparent.[4]

Further Reading

What the PSA developer thinks

To read the opinion of Prof Richard J. Ablin, who developed the PSA test, on how the test is used, go to: http://www.nytimes.com/2010/03/10/opinion/10Ablin.html

Because of progress in imaging and screening techniques, cancers can now be diagnosed at an extremely early stage. The natural history of these very early lesions is poorly understood. In particular, it can be difficult to distinguish between inflammation and early neoplasia. Those responsible for diagnosis?pathologists and radiologists?prefer to err on the side of caution and so may overdiagnose borderline cases.

Some misconceptions about screening

Early detection means good prognosis

A patient presenting with symptoms of a very early cancer is likely to have a longer life span after diagnosis than one presenting in the very late stages of the disease. Patients diagnosed as a result of screening are also likely to do better than patients presenting with symptoms. However, both these observations can result from sampling biases known as lead-time bias and length bias, and may not provide an accurate picture of the value of catching the illness and intervening at an early stage.

Lead-time bias

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Lead time bias arises because there is a delay between when disease is detectable by screening and when it is likely to produce symptoms and be diagnosed without screening. Screening leads to an earlier diagnosis, as shown in Figure 9.2, so the patient lives longer with the diagnosis, but may not live longer overall. Hence, if the screening programme is evaluated in terms of survival time, this may give a falsely positive impression of success.

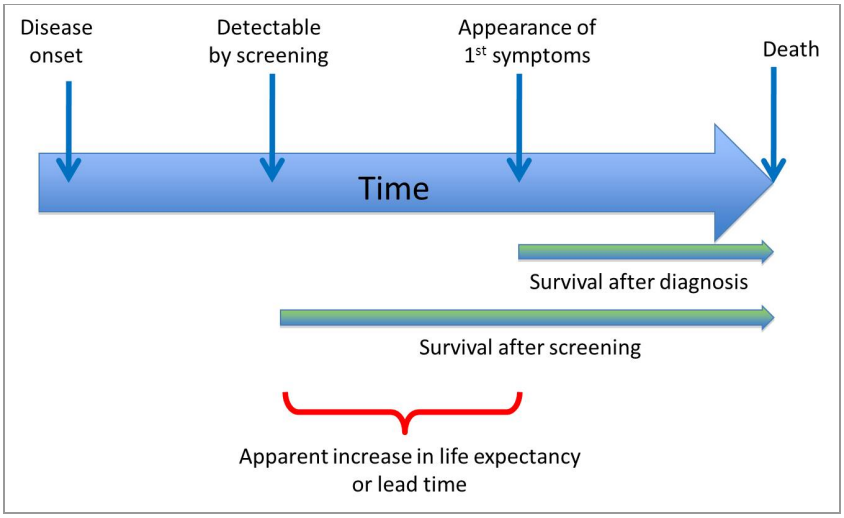


Figure 9.2: Lead time bias

Length bias

Slowly progressing variants of disease remain in a pre-symptomatic (but screen-detectable) stage for longer than rapidly progressing variants of the same disease. Hence, screening is more likely to detect slowly progressing cases. To a clinician comparing screen-detected cases with symptomatic cases that include the more rapidly progressing ones, it appears that screen-detected cases live longer, as shown in Figure 9.3.

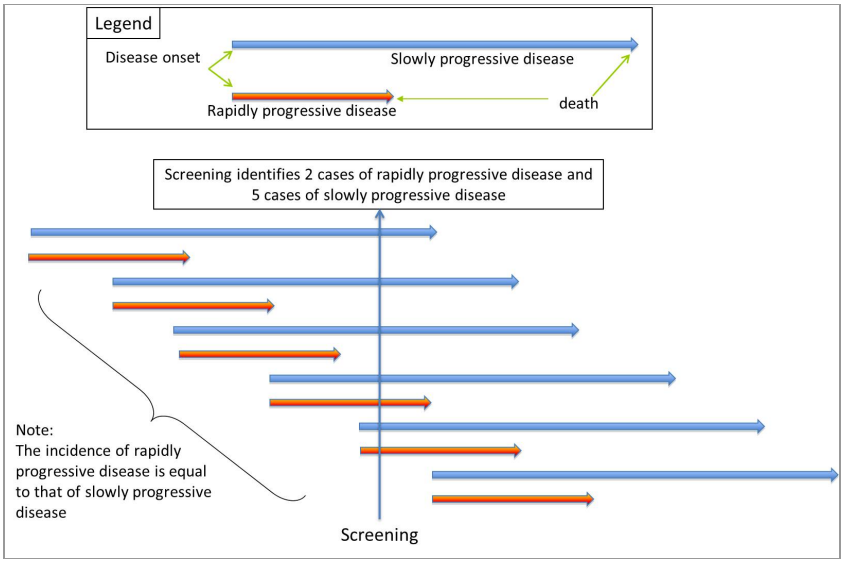


Figure 9.3: Length bias. In this population there are equal numbers of slow (blue) and rapidly-progressing (red) cases. But the screening test will identify five slower cases and only two rapidly progressing cases. Calculating mean survival from these seven cases will give an impression of longer average survival than occurs in the population.

People who are screened may live longer, but for other reasons

People who readily adopt new, supposedly healthy behaviours tend to be healthier than those who do not. Observational studies, which allow people to self select, generally show that people who have opted for screening have better outcomes than those who did not. For instance, women who have Pap smears are less likely to die from cervical cancer. However, they also tend to be better-educated and wealthier and, therefore, have a lower base-line risk of dying from cervical cancer than those who don't have Pap smears. Even after controlling for stage at diagnosis, women in higher socio-economic groups have a better prognosis than those in lower socio-economic groups. It is difficult to determine how much the favourable outcome in women who have Pap smears is due to the test and treatment, and how much is due to their other favourable health determinants.

"I wouldn't be here today if my disease had not been found early"

One of the most compelling pro-screening arguments is the personal testimony of patients who, having been screened, were diagnosed, and treated. These patients often attribute their continued survival to the fact of having been screened and sometimes become advocates for it. However, for an individual case, it is impossible to tell if screening has made a difference. Furthermore, about 70% of unscreened women diagnosed with breast cancer do not die from the cancer.[5] Interestingly enough, though, one rarely hears the statement "I wouldn't be here today if my

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

symptoms had not caused me to consult a physician."

"If I get screened, I will not get the disease"

A further danger that both the population and professionals fall into is the belief that screening is 100% effective. Although the results of screening are well known and pamphlets on screening clearly state that screening does not save every life, those who are undergoing screening tend to believe that theirs is the life that will be saved. This cognitive reaction is logical: if you do not believe it will save you, why undergo the discomfort of screening? Sadly, even the most effective screening programmes are not 100% effective. For example, screening and treatment for colorectal cancer reduces mortality by about 15%, but people whose cancers were detected by screening can still die from them. Controlled trials show that, at best, breast cancer screening reduces mortality by 30%, not 100%. In one study, the five year survival of women with screen-detected breast cancer was 94%, compared to 84% in women with cancers not detected by screening.[6] In these cases, mortality was reduced, but not prevented. Many other screening tests are much less effective. Diagnosing and treating diabetes may moderately reduce or delay diabetic complications, but the evidence suggests that it makes a difference only in hypertensive patients.[7]

The benefits of screening are often communicated in terms of population risks. For instance, breast cancer screening can lower the population death rate from breast cancer by 30%. An individual's risk is, therefore, reduced by 30%. But this is a relative risk. In terms of **absolute risk**, a 50-year-old woman who commits to biennial screening for the next twenty years only lowers her risk of dying from breast cancer from around 4% to around 3%. Although the benefit for a whole population is considerable, the benefit for the individual can be fairly small (see Chapter 8).

Links

Points to consider

For "Why whole body CT screening is not a good idea", visit the Canadian Health Services Research Foundation website:
http://www.chsrf.ca/PublicationsAndResources/Mythbusters/ArticleView/09-12-01/96a2ba7f-c114-410e-a9fe-123b74682ece.aspx

For "Why early detection is not good for everyone", visit the Canadian Health Services Research Foundation website:
http://www.chsrf.ca/PublicationsAndResources/Mythbusters/ArticleView/06-06-01/e2338423-6991-472e-ae90-adf84f1e1130.aspx

Details of policies

Even when a screening procedure is proven valuable, a series of choices must be made about how to put it into practice. Careful assessment of the epidemiology of the disease is needed to indicate which population group to screen, including the age at which to start and at which to stop screening. For example, Chinese women in Hong Kong have a much lower incidence of breast cancer than women in western countries; therefore, a mammography programme there would currently bring little benefit.[8] However, because Hong Kong lifestyles are changing, it might be of benefit in the future.[9] Regarding cervical cancer screening, many countries do not start until age 25, since below that age cancer is extremely rare, and minor changes related to Human Papillomavirus that are read as abnormal commonly regress spontaneously.

The cultural and organizational context of screening also has an effect on recommendations. For instance the epidemiology of breast cancer is similar in the UK, Australia and North America, yet the UK recommends screening for breast cancer every three years, Australia every two years, and the U.S. recommended it every year until 2009 when the recommendation was changed to every two years.[10]

Deciding on the screening interval requires knowledge of the natural history of the disease being screened for and of the **sensitivity** of the test. Too often, for convenience people are screened annually, not because this is what evidence supports. For instance, some U.S. authorities and many health professionals still recommend annual Pap smears. Interestingly enough, the best results for cervical screening come from Finland and Holland, where Pap smears are carried out every five years, but include a very high proportion of the population. The common Canadian practice of screening compliant women every year may be one of the factors that ties up resources and turns attention away from population groups most at risk from invasive cervical cancer and who are least likely to be screened, such as women living in rural areas, Aboriginal women, recent immigrants, older women, and those with lower incomes.

Finally, there is the question of who should carry out the screening. Physicians often feel a responsibility to do so, although other professionals or even screening technicians may be able to provide a reliable service at less cost. Similarly, machines capable of repetitive tasks currently performed by humans, such as reading Pap smears or mammograms, may be more effective and, in the long term, cheaper than humans.[11]

The costs of screening

Even though the cost per screening test is usually low, when the test is applied to whole populations considerable sums of money are involved. People with screen-detected abnormalities must be investigated; those who have the disease must be treated. Many of these people might not otherwise have had any investigation or treatment. According to the estimates of one researcher who looked at costs of breast cancer detection and treatment with and without screening, screening reduces the cost of treating breast cancer by 21%. However, a comparison of the overall health care costs including those not related to breast cancer, for screened versus unscreened women, showed that breast cancer screening increases total health care costs by nearly 6%.[12]

The cost of screening is not just the price of the test, which alone looks affordable. Image readers and laboratory staff may need specific training and updates. Screening requires an organized system to ensure the quality of equipment, screening, and patient management systems to ensure that patients get the follow-up they need. A system to remind patients when screening tests are due might also be needed. The costs of screening are

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

not merely financial costs to the system, but also include the physical, mental, and financial costs to patients and their families (see Table 9.1).

Table 9.1: The physical and emotional cost of screening

Costs due to:	Examples
The test itself	Colonoscopy carries a small risk of bowel perforation; mammography exposes the breast to radiation, albeit minimal.
Procedures engendered by false positive results	Diagnostic and treatment procedures carry risk. In the case of treatment for putative cancers, these can be severe. About 10% of women who undergo breast cancer screening for the first time and about 6% of women getting subsequent screening undergo further testing only to find that the screen result was falsely positive.[13]
Unnecessary anxiety after false positive results	Most patients experience anxiety on being told of a positive screening result. Some patients take considerable time to get over the anxiety. As many as 5% of Pap tests in young women are false positives.
False reassurance after negative results	After having a negative screen, whether a true or a false negative, some patients ignore subsequent symptoms of the disease.
Imposition of difficult choices	Men over 65 found to have an asymptomatic abdominal aortic aneurysm have a choice of either undergoing a very risky operation immediately or face a 70% risk of death within the next five years.
Prolongation of the period of illness	As their disease is detected earlier, patients who are screened live longer with the diagnosis. In some cases, the consequent anxiety and changes to lifestyle reduce the quality of life.
Labelling	If the disease is viewed negatively, the patient runs the risk of being viewed, or viewing himself, prejudicially. Consequences are both social and psychological. They can also be financial if the patient loses his job or can't find insurance after being diagnosed.

Ethical issues particular to screening

Beneficence and non-maleficence

The clinician who proposes interventions to patients who have not asked for them has an added responsibility to ensure the likely benefit outweighs any possible harm. Organized screening programmes rely on reaching a large proportion of the population in order to achieve the benefits expected for population health. However, screening often has no impact on an individual's risk from the disease being screened for, and the test and follow-up procedures carry an inherent, although minimal, risk of harm.

Patient autonomy

Patients can be coerced into screening by enthusiastic clinicians and by publicity campaigns that stress population results more than individual benefit. Even careful clinicians can have difficulty ensuring that patients have a realistic view of the risks and benefits of screening so that they can give properly informed consent. (see chapter 10) Clinicians also face patients asking for screening tests, such as total body scans, an annual physical examination, or a PSA test, for which evidence does not support their use in most populations. The physician must balance his duty not to harm against his duty in relation to patient autonomy of choice. There is no ready answer to this dilemma: each case must be considered on its own merits. In general, however, clinicians should probably dissuade the patient from undergoing procedures that have not been shown to be beneficial.

Given that population programmes generally need to achieve a high uptake in order to achieve their effectiveness, some programmes include call and recall systems that send reminders to patients when their screening is due. Patients may interpret the content of these letters as having had the decision made for them and, therefore, abdicate their right of informed consent.[14]

Equity

Deprived, vulnerable people at greater risk of disease are less likely to undergo screening interventions. Screening programmes can actually increase the health gap if a higher fraction of those with better health determinants attend, while vulnerable people remain untouched by the programme.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

The Inverse Care Law

An unfortunate inverse relationship between need and demand was termed "the inverse care law" by Julian Tudor Hart: those at highest need are the least likely to receive screening or other care.[15]

An unintended consequence of interventions is often that their initial uptake occurs among opinion leaders, informed people and the "worried well" so that their health improves, widening the disparities in health between them and other, less informed people. If subsequent population health efforts succeed in increasing uptake by specifically targeting the less informed people, the gap can narrow again. However, as innovations are always being proposed, this health gap produced by delayed uptake of innovations is likely to persist. A report from the Public Health Agency of Canada presented a diagram illustrating the impact of delayed adoption of innovations on producing socioeconomic disparities in health at http://www.phac-aspc.gc.ca/ph-sp/disparities/pdf06/disparities_discussion_paper_e.pdf (see Figure 3, page 21 of the report).

Case Study

Having started to discuss prostate cancer screening, Dr Rao realised that Paul had a number of deep-seated fears and false beliefs on the subject. Dr Rao asked Paul to drop into Nurse Jennings' office to get fact sheets on prostate cancer and prostate cancer screening. He also asked Paul to make another appointment when Dr Rao would have more time to continue the discussion.

Criteria for introducing a screening test

Given screening's superficial attractiveness, but underlying potential for wasteful use of resources, in 1968 the WHO published criteria for judging screening programmes. Wilson and Jungner, the experts who developed the criteria, noted that "the central idea of early disease detection and treatment is essentially simple. However, the path to its successful achievement (on the one hand, bringing to treatment those with previously undetected disease, and, on the other, avoiding harm to those persons not in need of treatment) is far from simple though sometimes it may appear deceptively easy."[16] Their criteria, therefore, went beyond just the consideration of the test characteristics to consider the context in which the test would be introduced, including disease characteristics and the organizational context of the screening.

Nerd's Corner

The original WHO criteria for screening[16]

The criteria developed by Wilson and Jungner may be helpful for clinicians in deciding whether or not to apply a test in the context of their practice:

1. The condition sought should be an important health problem.
2. There should be an accepted treatment for patients with recognized disease.
3. Facilities for diagnosis and treatment should be available.
4. There should be a recognizable latent or early symptomatic stage.
5. There should be a suitable test or examination.
6. The test should be acceptable to the population.
7. The natural history of the condition, including development from latent to declared disease, should be adequately understood.
8. There should be an agreed policy on whom to treat as patients.
9. The cost of case-finding (including diagnosis and treatment of patients diagnosed) should be economically balanced in relation to possible expenditure on medical care as a whole.
10. Case-finding should be a continuing process and not a "once and for all" project.

Since the publication of these criteria in 1968, many more options for screening have been developed. Some formal, organized screening programmes have been introduced, and others have been considered. Opinions on screening criteria have also progressed; the criteria now include the notion of screening programmes, informed choice for patients, equity in application of the programme, and evidence of costs and benefits, and of effectiveness of the programme.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

WHO criteria for screening: synthesis of screening criteria[17]

1. The screening programme should respond to a recognized need.
2. The objectives of screening should be defined at the outset.
3. There should be a defined target population.
4. There should be scientific evidence of screening programme effectiveness.
5. The programme should integrate education, testing, clinical services, and programme management.
6. There should be quality assurance, with mechanisms to minimize potential risks of screening.
7. The programme should ensure informed choice, confidentiality, and respect for autonomy.
8. The programme should promote equity and access to screening for the entire target population.
9. Programme evaluation should be planned from the outset.
10. The overall benefits of screening should outweigh the harm

Each criterion forms part of a chain, which, like any chain, is only as strong as its weakest link. If the proposed screening activity fails to meet any of the criteria, it should not be carried out.

Self-test questions

Self-Testing

1. One of your colleagues has just read an article that shows that screening for orange disease using the marmalade test can prolong survival of patients with the disease. What information do you need before deciding whether or not to support this conclusion?
You need to know what kind of evidence is presented in the article. An observational cohort study that showed that patients who were screened had better results that those who were not is subject to bias due to self-selection; patients who choose to have a test generally do better than those who don't because they tend to also have other more positive health determinants. A case series that showed that patients presenting after screening had a longer survival time than those presenting when the disease is clinically apparent is subject to lead-time and length biases; screening picks up cases earlier on in the natural history of disease with the result that, as the time of diagnosis is advanced, the period during which the patient is known to have the disease is prolonged. This makes it appear that survival has been prolonged. Furthermore, screening favours the diagnosis of indolent forms of disease, so cases picked up by screening are likely to be those which progress slowly.

Because of such problems the only acceptable evidence for screening comes from a randomised controlled trial. High level evidence in screening is particularly important because screening always prolongs the period of illness and often brings psychological, social and financial problems to people who would otherwise have enjoyed apparently good general health for a longer period.

If you are satisfied that the article describes a valid controlled trial which shows that screening results in a reasonably large clinical benefit, you then need to check a few other factors before forming an opinion. The main one concerns how much work has already been done on the subject and is if the article agrees with it. It is generally rash to accept findings that disagree with those of other valid studies. It is also rash to act on the basis of a single trial.

You would also like to know background information about orange disease, about its frequency, its likely outcomes and the effectiveness of treatment options. Is the natural history known, and is early treatment more effective than treating at a later stage? Your assessment includes consideration of both length and quality of life, as well as the risks of treatment. Secondly, is a screening test possible ? is there a preclinical phase that lends itself to early diagnosis?

Once you are satisfied that the evidence from a number of sources is solid, then you need to look at issues in implementation. (See the WHO criteria) You would like to know about costs of screening and of treatment for the patient and for the health care system, and whether or not you have at your disposal the means of carrying out the test, diagnosing patients who are found to be at high risk of orange disease and the means of treating those confirmed to have orange disease. You would like to know about the performance characteristics of the serum marmalade test: its positive and negative predictive values in the general population. You would also like to know whether or not patients find it acceptable.

Reflection Questions

1. Describe the breast cancer screening programme in your province.
2. What are your provincial guidelines for prostate cancer screening?
3. At what level of glycemia or blood sugar do your professors recommend action should be taken in healthy patients?

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

References

Further Reading

Rosen G. Preventive medicine in the United States 1900?1975. New York: Prodist, 1976.

Welch HG. Should I be tested for cancer? Maybe not and here's why. Berkeley: University of California Press, 2004.

1. Porta M. Dictionary of epidemiology. New York: Oxford University Press; 2008.

2. Holowaty P, Miller AB, Rohan T, To T. Natural history of dysplasia of the uterine cervix. J Natl Cancer Inst 1999 Feb 3;91(3):252-258.

3. Collège des médecins du Québec. The PSA test and screening for prostate cancer. 1998 February;2009(August).

4. Morrison AS, Brisson J, Khalid N. Breast cancer incidence and mortality in the breast cancer detection demonstration project. J Natl Cancer Inst 1988 Dec 7;80(19):1540-1547.

5. Cancer surveillance online. Available at: http://dsol-smed.phac-aspc.gc.ca/dsol-smed/cancer/index_e.html. Accessed December, 2009.

6. Cortesi L, Chiuri VE, Ruscelli S, Bellelli V, Negri R, Rashid I, et al. Prognosis of screen-detected breast cancers: results of a population based study. BMC Cancer 2006;6:17.

7. US Preventive Services Task Force. Screening: Type 2 Diabetes Mellitus in Adults. 2008; Available at: http://www.ahrq.gov/CLINIC/USPSTF/uspsdiab.htm. Accessed December, 2009.

8. Leung GM, Lam TH, Thach TQ, Hedley AJ. Will screening mammography in the East do more harm than good? Am J Public Health 2002 Nov;92(11):1841-1846.

9. Leung GM, Thach TQ, Chan E, Foo W, Meng O, Fielding R, et al. Short-term, medium-term, long-term, and lifetime risks of developing and dying of breast carcinoma in a Westernized Chinese population: Evidence from Hong Kong between 1976 and 2000. Cancer 2005 Feb 1;103(3):501-508.

10. US Preventive Services Task Force. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2009 W-236; Nov 17;151(10):716-726.

11. Davey E, d'Assuncao J, Irwig L, Macaskill P, Chan SF, Richards A, et al. Accuracy of reading liquid based cytology slides using the ThinPrep Imager compared with conventional cytology: prospective study. BMJ 2007 Jul 7;335(7609):31.

12. Johnston K. Modelling the future costs of breast screening. Eur J Cancer 2001 Sep;37(14):1752-1758.

13. Health Canada. Organized Breast Cancer Screening Programs in Canada, 1999 and 2000 Report. 2003; Available at: http://www.phac-aspc.gc.ca/publicat/obcsp-podcs00/pdf/obcsp00_e.pdf. Accessed August, 2009.

14. Osterlie W, Solbjor M, Skolbekken JA, Hofvind S, Saetnan AR, Forsmo S. Challenges of informed choice in organised screening. J Med Ethics 2008 Sep;34(9):e5.

15. Hart JT. The inverse care law. Lancet 1971 Feb 27;1(7696):405-412.

16. Wilson JMG, Jungner G. Principles and practice of screening for disease. 1968; Available at: http://whqlibdoc.who.int/php/WHO_PHP_34.pdf. Accessed August, 2009.

17. Andermann A, Blancquaert I, Beauchamp S, Dery V. Revisiting Wilson and Jungner in the genomic age: a review of screening criteria over the past 40 years. Bull World Health Organ 2008 Apr;86(4):317-319.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Chapter Objectives

After reading this chapter, you will be able to:

- Recognize the implications of environmental hazards at both the individual and population level;
- Respond to the patient's concerns through appropriate information gathering and treatment;
- Make appropriate recommendations for patients and exposed populations so as to minimize their health risks and maximize their overall function;
- Identify common environmental hazards and be able to classify them into the appropriate category of chemical, biological, physical and radiation;
- Identify the common hazards that are found in air, water, soil, and foods;
- Describe the steps in an environmental risk assessment and be able to critically review a simple risk assessment for a community;
- Conduct a focussed clinical assessment of exposed persons in order to determine the causal linkage between exposure and the clinical condition;
- Communicate simple environmental risk assessment information to both patients and the community;
- Describe the importance and impact of good, culturally-appropriate communication with the patient, the patient's family and, if necessary, the community as a whole with regard to risk factors and their modification; and
- Name and describe the common methods of health protection (such as agent-host-environment approach for communicable diseases, and source-path-receiver approach for occupational/environmental health).

Case Study

Julie Richards consults Dr. Rao because she is concerned about her eyes which have been red and irritated ever since she spent a day cleaning out the basement last week. She wonders if it is related to radon gas. She heard radon is a problem in the area and that it accumulates in basements. Dr. Rao is well aware that the environment can cause health problems. For Dr Rao, the notion of environment includes the natural one, the built or man-made one and the social environment. Examining his records, he notices that this is not the first time anyone in the Richards family has experienced such problems.

Seven years ago, when Paul injured his neck at work, Dr. Rao asked him about injuries sustained by other people at his work, and also about absenteeism. At the time, Dr. Rao made a note that an unusually high number of people who worked in the mine were consulting him for reasons ranging from trivial upper respiratory tract infections to myocardial infarction. A few weeks after Paul was injured the first round of lay-offs at the mine was announced. It was then that Dr Rao wondered if the work environment might be a problem.

Identifying environmental problems

Many diseases are caused by exposure to environmental hazards. Although a high index of suspicion is essential for diagnosis, once diagnosed, environmentally-induced problems often respond to simple environmental solutions. Moreover, many patients are worried about elements of their environment that have little or no effect on health, so physicians need to be able to distinguish those which are hazardous from those which are not. Environmental factors are involved in the aetiology of virtually every illness. To diagnose illness predominantly environmental in origin, physicians have to include a thorough environmental assessment in their history-taking. Clues that an illness is caused by an environmental factor include:

- The patient suspects it;
- The pattern of illness is atypical; for example, the patient is not in the usual age group, the usual risk factors are absent, the course of the illness is unusual, the symptoms do not respond to the usual treatments;
- The temporal pattern of the illness suggests it (for example, symptoms improve when the patient goes on holidays and worsen on returning home, or they worsen when the patient is at work and improve at home);
- There is no obvious other cause for the illness; or
- The signs and symptoms suggest specific toxins, such as lead or mercury poisoning.

Once an environmental cause is suspected, the physician should take a detailed environmental history to identify all hazards the patient may have been exposed to (see the box "Taking an environmental history"). All of the patient's activities and environments should be explored. The chronology of events, patient's proximity to the presumed source, and whether or not other people are affected (i.e., time, place, and person) should be explored for evidence that supports or discounts the hypothesis of an environmental cause. It might be necessary to call in a range of professions and authorities to ensure that the patient and any other people exposed are adequately protected. The public health service is often the first port of call, but, unless the problem is widespread in the community, public health has no authority or funding to act in relation to domestic environmental problems, although it is a source of advice. Even when the problem is widespread, the public health department's role is likely to be one of education and information. Some provinces require physicians to notify the public health department of particular environmental illnesses or possible outbreaks. The patient's work health and safety board might be involved in the case of a work hazard. In an emergency, as with a case of carbon monoxide poisoning, public security services (fire and police) would coordinate control efforts.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Taking an environmental history[1]

Ask about ?CH_20PD_2':

- **C**ommunity: neighbourhood sources of hazard; industry, waste storage
- **H**ome: year of construction, renovations; materials used in construction and decoration; moulds; garden and house plants; use of cleaning products, pesticides, herbicides
- **H**obbies and leisure: exposure to chemicals, heavy metals, dusts, or micro-organisms
- **O**ccupation: current and previous occupations; work with known hazards; air quality
- **P**ersonal habits: hygiene products; smoking
- **D**iet: sources of food and water; cooking methods; food fads
- **D**rugs: prescription, non-prescription, and alternative medications; health practices

If a scanning question reveals a possible hazard, ask detailed questions to find out as much as possible about the nature and level of the hazard and then check Time, Place and Person:

- *Time*: When did symptoms begin? When did exposure begin? When do symptoms get worse? When do they improve?
- *Place*: Where is the patient when symptoms get worse? Where is the likely hazard? What is the channel through which the hazard reaches the patient?
- *Person*: Does anyone else have similar symptoms? Who? When? Where?

Here Be Dragons

Take the full (figurative) tour

Ask questions about all the environments that the patient inhabits?even if the patient suspects one in particular, or if your first few scanning questions point to a specific environment. If you stop enquiring as soon as you find your first hazard, you might miss other hazards to which the patient is exposed. The real cause of the patient's problems may not be the first your history uncovers.

Case Study

Dr. Rao asks Julie a bit more about her activities around the time of her eye problems. She says she was cleaning out the basement, which was very dusty. She also used a cleaning product that she hadn't used before. Because it was cold outside she didn't open the window while she was working. The day before cleaning the basement she had gone for a walk; it had been a windy day and some dust had blown into her eyes. Right now her eyes are normal. Dr. Rao advises Julie to make sure that, if she is cleaning a space she hasn't cleaned for some time, she should make sure it is well ventilated. He also suggests that she have a look at the Canada Mortgage and Housing website for advice on home maintenance.

As this is not the first time that a patient has asked about radon gas, Dr. Rao asks his practice manager to contact the local public health department for information on radon that he could pass on to his patients.

Links

Housing maintenance

For information on health and housing maintenance issues visit the Canada Mortgage and Housing Corporation website: http://www.cmhc-schl.gc.ca/ The section ? Maintaining a Home' contains a lot of practical information.

Reducing risk

The three main stages in addressing the health effects of environmental hazards are risk assessment, risk management, and risk communication. The clinician applies these when discussing environmental disease with a patient; a public health officer applies them in responding to community wide problems.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Definition

Hazard is the inherent capability of an agent or a situation to have an adverse effect. A hazard is a factor or exposure that may adversely affect health.

Risk is the probability that such an event will occur.

Health risk = hazard x exposure x susceptibility.

Thus, for there to be a risk, a susceptible individual or population (receiver) must be exposed (path) to a hazard (source) (see Figure 10.1). If there is no exposure, there is no risk. If there is exposure, but no susceptibility, there is no risk. The same idea is used in fire prevention: use a non-flammable material (i.e., reduce susceptibility) or reduce its exposure to heat (block the path), or remove the source of heat.

Risk assessment

Risk assessment is the process of evaluating the likelihood of occurrence and probable severity of health effects due to a hazard. Various agencies assess risk situations, including occupational health agencies, environmental protection agencies, and public health authorities. Clinicians confronted with a possible environmental disease might use some of the following risk assessment steps to formulate a diagnosis and disease management plan.

Steps in risk assessment[2]

- Hazard identification: Is an environmental hazard involved? What is it?
- Risk characterization: Is the hazard likely to cause these types of symptoms in this type of patient?
- Exposure assessment: Is the patient's exposure enough to cause these symptoms?
  - Risk estimation: How much has the hazard contributed to the patient's condition?

Hazard identification

Hazard identification identifies the agent responsible for the problem, its adverse effects, the target population, and the conditions of exposure. Hazards are categorized as biological, chemical, physical, ergonomic, psychosocial, or related to safety. Table 10.1 gives examples of the various types of hazard and their possible health effects.

Table 10.1: Examples of the various types of environmental agents and their associated health effects

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Type of hazard	Examples	Health effects
Biological	Bacteria, viruses	Specific syndromes associated with different agents e.g., salmonella food poisoning, hepatitis A, infection with Methycillin resistant staphylococcus (MRSA)
	Moulds	Allergies, cancers
	Animals	Allergies, zoonoses
Chemical	Heavy metals (the risk may depend on the physical state)	Specific syndromes e.g., lead poisoning, mercury poisoning
	Benzene	Acute myeloid leukemia with prolonged exposure
	Carbon monoxide	Asphyxiation
	Asbestos	Asbestosis, carcinoma of the lung, mesothelioma
Physical	Noise	Hearing loss
	Radiation	DNA damage leading to cancers
	Ultraviolet light	Skin damage, vision loss
	Temperature extremes	Hypo- or hyperthermia
	Poor protection from kinetic energy	Falls and collisions leading to bone and soft tissue injury
Ergonomic	Poorly designed work station	Back pain
	Physically repetitive activity	Repetitive strain injuries
Psychosocial	Job stress	Non-specific physical and psychological manifestations
	Poor social support	Psychological problems

Links

Hazardous occupations

For a source of guidance and information for patients in hazardous occupations

visit the Canadian Centre for Occupational Health and Safety website: www.ccohs.ca

Hazards can be found almost everywhere in the environment, including the air, water, soil, and in food. Hazards may affect health directly or indirectly, when they produce a change which puts a person or population at risk. Table 10.2 lists some of the common hazards.

Table 10.2 Examples of hazards in the environment

Hazards in:	Examples
Air	Carbon monoxide Smog Particulate matter
Water	Fecal contamination Cryptosporidia E. Coli Blue-green algae
Soil	Heavy metals Petroleum by-products
Food	Listeria Salmonella Mercury in fish

Indoor hazards that cause health problems include household chemical products, carbon monoxide, radon, mould, lead, consumer products (cosmetics, perfumes, hygiene products). Second hand smoke is still a problem in some households.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

The environmental history (CH_2OPD_2) may indicate a hazard in the patient's environment. The clinical picture may indicate what kind of hazard to search for. Airborne irritants can cause itchy or sore eyes, runny nose or coughing. Irritants that come into direct contact with the skin can cause dermatitis. Allergens can cause numerous symptoms including dermatitis, asthma, sneezing. Asphyxiants cause different respiratory problems depending on the type of asphyxiant. For instance carbon monoxide blocks oxygen transport. Other gases can accumulate in high concentrations displacing oxygen in the air. Ingested or absorbed substances may damage specific organ systems in pathognomonic ways. For instance, severe lead poisoning causes neurological changes, abdominal pain, and anaemia, poisoning with elemental mercury typically produces tremors, among other neurological symptoms, while mesothelioma is linked to exposure to asbestos.

The effects of a hazard may be delayed, sometimes for many years. This is especially true of cancer-causing agents. Identifying this type of hazard and evaluating claims about it can be difficult. Furthermore people can be exposed to more than one hazard, each contributing to their health problem. For instance, in the case of a miner, who smokes and has chronic lung disease, it can be difficult to assess the relative contributions of his occupation and his smoking.

In a situation where there seems to be an outbreak of a disease of environmental origin, public health officials will collect initial information on the possible sources. Once a case definition is established, cases are sought and information is collected on when they were exposed, when they started showing effects, where they were during that time, and demographic characteristics (time, place and person). This information is then collated to produce a picture of the distribution of cases in time and place. (see Chapter 11 Steps in outbreak management).

Risk characterization

This step describes the potential health effects of a hazard and answers the clinician's question about whether the identified hazard could possibly cause the patient's symptoms. As far as possible, effects on molecular, biochemical, cellular and organ systems are described in this step. A chemical hazard may only cause health problems when in a specific form. Also the route of entry into the body may be an important determinant of the damage caused. Furthermore, a person's response to a hazard is mediated by the factors that influence the hazard's **toxicokinetics** and **toxicodynamics**. The individual's genetic makeup and environmental factors moderate the toxicokinetics and toxicodynamics, so susceptibility to damage by a hazard varies from one person to another.

Illustrative Materials

Differential susceptibility to malaria

Malaria is an example of a hazard from which the risk is modified by the genetic makeup of at-risk people. The sickle cell trait confers some resistance to malaria, and this could explain why the trait has persisted in African populations. One study showed that, compared to children without the trait, those with the sickle cell trait had a relative risk of all cause mortality of 0.45 (95% CI 0.24?0.84) from the ages of 2 to 16 months.[3] This is the peak age for severe malaria. There was no difference in mortality before the age of 2 months, probably because of maternal immunity, or after the age of 16 months, probably because those who survive to 16 months develop some immunity from repeated exposures to small infectious loads.

Nerd's Corner

Kinetics and dynamics[4]

Toxicokinetics = the activity or fate of toxins in the body over a period of time, including the processes of absorption, distribution, localization in tissues, biotransformation, and excretion.

Kinetics comes from the Greek word that means movement. Cinema (moving pictures) and kinesiology are derived from the same word.

Toxicodynamics = the study of the biochemical and physiological effects of toxins and the mechanisms of their actions, including the correlation of actions and effect of toxins with their chemical structure. It includes the effects of a toxin on the actions of other toxins.

Dynamics comes from the Greek word meaning "force" or "power."

The terms toxicokinetics and toxicodynamics are analogous to pharmacokinetics and pharmacodynamics.

A low dose of some substances may be beneficial for a person, while a higher dose is toxic. For instance, fat soluble vitamins, such as A and D, are essential for health, but too much of either is toxic. Also, sunlight in small doses increases the production of Vitamin D, while in high doses it can cause skin cancer. Some hazards, such as heat and noise, must reach a threshold level before damage occurs. Others, including many cancer-causing agents, are presumed to cause damage even at the lowest measurable levels. The effects of some agents, such as X-rays, are cumulative, while others, such as alcohol, allow the body to recuperate somewhat during temporary breaks in exposure.

Definition

Hormesis refers to a biphasic dose response to an environmental agent characterized by stimulation or a beneficial effect at moderate doses and an inhibitory or toxic effect at high doses.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

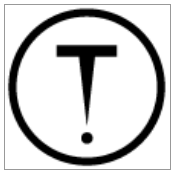
Nerd's Corner

Toxicokinetics of stress

Toxicokinetics and the varying dose-impact of different toxins have a parallel in the relationship between psychological stress and resulting strain or distress. Personality mediates the impact of psychological stress on the resulting strain, producing characteristic stress-strain responses. In toxicology, different substances have differing exposure-damage curves somewhat reminiscent of the stress-strain curves described by Young's modulus http://en.wikipedia.org/wiki/Stress-strain_curve .

For information on hazardous substances, clinicians can look to the scientific literature, check with the local public health department or call the toxicology or poison centre to find out about the characteristics of the possible hazard. If the probable source of the hazard is an industrial product, information is likely to be available via the Workplace Hazardous Materials Information System (WHMIS), which sets out the labelling requirements for hazards, see Table 10.3. The product label also indicates if a ?Material Safety Data Sheet' is available for the product. This sheet contains further details about the hazard, how to handle it safely and what to do in an emergency.

Table 10.3 the Workplace Hazardous Materials Information System symbols[5]

	Class A - Compressed Gas	Contents under pressure. Cylinder may explode or burst when heated, dropped or damaged.
	Class B - Flammable and Combustible Material	May catch fire when exposed to heat, spark or flame. May burst into flames.
	Class C - Oxidizing Material	May cause fire or explosion when in contact with wood, fuels or other combustible material.
	Class D, Division 1 Poisonous and Infectious Material: Immediate and serious toxic effects	Poisonous substance. A single exposure may be fatal or cause serious or permanent damage to health.
	Class D, Division 2 Poisonous and Infectious Material: Other toxic effects	Poisonous substance. May cause irritation. Repeated exposure may cause cancer, birth defects, or other permanent damage.
	Class D, Division 3 Poisonous and Infectious Material: Biohazardous infectious materials	May cause disease or serious illness. Drastic exposures may result in death.
	Class E - Corrosive Material	Can cause burns to eyes, skin or respiratory system.
	Class F - Dangerously Reactive Material	May react violently, causing explosion, fire or release of toxic gases, when exposed to light, heat, vibration or extreme temperatures.

Exposure assessment

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Exposure assessment is the step that quantifies the exposure of a person or population to a hazard. The levels of some hazards can be directly measured, either in the environment or in the people exposed. More often, however, exposure must be estimated from a careful history of the patient's activities, as well as an inspection of the environment.

Hazards can contaminate air, water or soil, or be part of the man-made environment (such as lead in paint, or asbestos in some older buildings). They may produce their effects by coming in contact with the skin, or they may be inhaled, ingested, absorbed transdermally, or injected parenterally. A radiation source may damage from a distance. Activities, working practices, and processes during which a person is likely to be exposed must be recorded in detail to determine how and how much the person was exposed to the hazard.

When the suspected agent is one that causes problems after a long delay, the exposure history may have to trace exposures back twenty years or more. Occupational disease can occur in retired people; in workers, it can be caused by previous employment.

Illustrative Materials

Thalidomide and phocomelia[6]

Thalidomide was developed as a sedative and anticonvulsant drug. As a sedative, it was remarkable because it was almost impossible to die from an overdose; in fact, no LD50 could be established. It was first marketed as a "harmless" sedative in Germany in 1957, and it became the drug of choice for many conditions, including morning sickness in early pregnancy. At the time, drugs were thought not to cross the placenta and, therefore, not to affect the foetus. However, by 1960, geneticists and paediatricians began to see children with unusual limb abnormalities that are now known as phocomelia (seal flippers). By 1961, it was noted that the drug could cause peripheral neuritis in people taking it. There were also reports of phocomelia in babies whose mothers had taken thalidomide in early pregnancy and scientists were beginning to suspect that thalidomide was causing the abnormalities. By the end of 1961, thalidomide had been taken off the market in the UK. In Canada, the drug continued to be sold until March 1962, although physicians were warned not to use it in pregnant women.

Recently, thalidomide has gained a new lease on life. In 2005, it was found effective in treating weight loss and cachexia associated with various cancers, as well as in slowing the growth of myeloma cells.

Risk estimation

Risk estimation quantifies the likelihood that a hazard will affect a specific person or population as well as the size or severity of the effect. In this step, the information obtained during the previous steps is summarized and collated. The epidemiological model can be used to put together the information on what the hazard is, where it comes from, how the environment allows it to come into contact with the host, and the host's susceptibility to the hazard in order to arrive at a conclusion on who is at what level of risk from what hazard. Table 10.4 shows some examples of what needs to be considered in this step.

Table 10.4: Examples of various host, agent, and environmental risk factors for selected categories of health problems

Problem type	Health problem	Host	Agent	Environment
Infectious disease	Hepatitis C virus infection among people who inject drugs	Co infection with HIV	RNA virus (Flaviviridae family)	Lack of sterile drug preparation and injection equipment (e.g., syringes)
Environmental health problem	Asthma	Genetic susceptibility	Allergen	Carpets; pets in household; inefficient ventilation
Occupational health problem	Back injuries in manufacturing facilities	Posture	Mechanical forces	Lack of equipment necessitates human lifting.

The risk for an individual and for a population should be balanced against the cost and risk of intervention. In individuals it might be a choice between the risk of asthma and the loss of a beloved pet, or, if the risk is work-related, it could be a choice between staying healthy or staying employed. In populations, the decision can affect a number of sectors. For instance, a population may be at risk from pollutants emanating from a factory. It might, however, be at greater risk from the poverty that would result if the factory were closed. Once the levels of risks due to the different hazards are documented, possibilities for risk reduction can be explored.

Nerd's Corner

Risk Assessment versus the precautionary principle[7]

The four steps of risk assessment described in the text fall into a quantitative risk assessment paradigm. They can be used to assess probable risks and benefits, allowing people to choose the option with least risk. However, the paradigm has been criticised for not taking into account the complexities and uncertainties of risks and risk assessment. The precautionary principle expresses the "better be safe than sorry" approach. It is generally used in situations where there is risk of severe, immediate, and irreversible damage to people or to the environment. A criticism of this approach is that it tends to be based on the assessment of only one option and may ignore the risks and benefits of the others.

Occupational hazards

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Most people spend a considerable part of their life at work, and if hazards are present, workers are likely to have had prolonged exposure to them. In some occupations, specific hazardous agents must be used, putting workers at risk of occupational diseases associated with them. Physicians should ask patients four simple questions: "What sort of work do you do?" "What sorts of work have you done?" "Do you think your work may be a cause of your problem?" And "What job are you going back to?"

Definition

The vocabulary of work-related health problems is governed by provincial legislation on work health and safety and workers' compensation. As a result, the precise definitions of key words may vary. However, the underlying concepts remain the same.

Work accident: causes an **occupational injury**. It must arise out of employment and in the course of employment, and includes one of the following:

- a wilful and intentional act, not being the act of the worker who suffers the accident
- a chance event or incident occasioned by a physical or natural cause
- a disablement caused by an occupational disease
- a disablement or disabling condition caused by employment.

Even though the word accident is used in legislation, according to many experts in the prevention of work injury it is to be eschewed; although most accidents are preventable, *accident* includes the notion of inevitability.

Occupational disease: a disease peculiar to or characteristic of an industrial process, trade, or occupation, or a disease that arises out of and in the course of employment.

According to Canadian federal and provincial laws, specific services must be provided for the protection of workers and for compensating those who are harmed as a result of work.

Table 10.5: Some examples of occupational disease

Condition	Agent	Example of at risk occupation
Berylliosis	Beryllium	Aerospace industry
Byssinosis	Cotton dust (numerous agents)	Cotton industry
Farmer's lung	Mould in hay	Farming
Asbestosis, Mesothelioma	Asbestos	Demolition work; ship-building
Hepatitis A	Hepatitis A virus	Sewer workers
Silicosis	Silica dust	Stone workers

Common, non-specific work-related diseases include dermatitis, asthma, and musculoskeletal disorders; other examples are shown in Table 10.5. It is also accepted that mental health can be influenced by stressful conditions at work. Injuries do not occur randomly; groups at particular risk of occupational injury include:

- Young workers lacking in work experience or safety training?particularly those in temporary or summer jobs;
- Workers obliged to work long hours at a fast pace. This could include piece-workers or those working on production lines;
- Workers given work responsibility without the authority or control necessary to meet that responsibility; and
- Workers in certain high-risk industries such as construction and forestry.

Risk Management

The source-path-receiver model, derived from energy technology, is often used in place of the epidemiological model to derive ways of controlling risk. The source is the equipment or process that is directly responsible for a noxious output. The output could be a form of energy, such as acoustic, thermal, etc., or it could be a substance, such as toxic fumes or dusts. The path is the conveying medium (air, water, etc.), while the receiver is the human being. In contrast to the epidemiological model, which allows for bidirectional influences, the source-path-receiver model is unidirectional.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

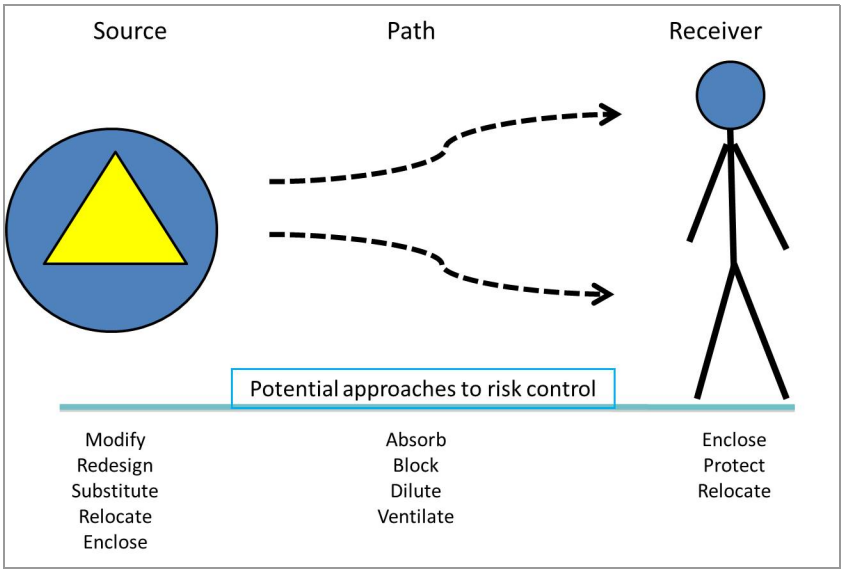


Figure 10.1: Source-path-receiver: a model to identify ways of reducing environmental risks

Links

Source-path-receiver and noise

For more on the source-path-receiver model as it relates to noise, visit Dr. Apple L. S. Chan's website : http://personal.cityu.edu.hk/~bsapplec/source-p.htm

As complete elimination of risk is generally impossible, risk management aims to reduce the risk from the hazard without causing another problem. Strategies to reduce risk can target the source of the hazard, its path or the receiver, whether the risk involves a person or a population. Strategies that address the source can be educational or advisory, technical or engineered, economic or regulatory. As illustrated in Figure 10.1, such strategies can modify or substitute the hazard (for example, replacing asbestos with other materials), or enclose it if it cannot be removed. Legislative strategies can be used to reduce the emission of hazardous pollutants.

Modifications to the path include environmental changes; for example, when smog levels are high, closing house windows prevents smog from entering the home, or, in countries where malaria is common, a film of environmentally safe soap or oil on standing water prevents breeding of malaria-carrying mosquitoes.

Finally, strategies can alter the susceptibility of individuals and populations to the risks of hazards. Examples include immunization against infectious disease, ensuring the nutritional status of disadvantaged children, and counselling for behaviour change. Barriers can also be used to reduce exposure. For example, power line workers use special tents in winter to reduce exposure to cold, and airport apron workers use ear protectors to prevent noise damage. Other safety measures should be promoted: safety harnesses, helmets and steel-capped boots for construction workers and miners.

Case Study

Shortly after counselling David on his health behaviours (See Chapter 8), Nurse Jennings happened to meet Mr White, a contractor for whom David occasionally works. She began by asking Mr White what he thought of his workers' safety habits and if there had been many injuries on his work sites?.

Illustrative Materials

An environmental hazard in a First Nation community

In October 2005, 450 people from the Kashechewan First Nation were evacuated from their community because of problems with their drinking water. The community had been under a "boil water" advisory for two years. For five years there had been continuing problems with the supply; there was persistent E. Coli contamination and the chlorine used to reduce this was aggravating skin problems. The Ontario Ministry of the Environment found several problems with the water quality monitoring and with the treatment of drinking water. There were also a number of problems with the sewage system, including the fact that the sewage system outlet was upstream of the drinking water intake, which increased the risk of drinking water contamination.[8]

This incident prompted calls for improved standards for the drinking water supply in First Nations communities. It was noted that the regulatory framework in First Nations communities did not ensure that water quality was as high as that in other Canadian communities. By 2008, no new legislation on water quality had been developed.[9]

Haddon's matrix and injury prevention

Dr. William Haddon combined the epidemiological model with a time dimension to create Haddon's matrix, that sets out the factors that determine

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

the level of risk associated with a hazard and the severity of its effects, in order to identify modifiable risk factors for injury.[10] It is commonly applied to the analysis of road traffic collisions, as shown in Table 10.6. In this example the host is the person injured (driver or other), while the agent is the equipment that determines how much energy (mechanical or thermal) is transmitted to the host. The environment refers to the physical and social environments in which the injury occurs. The time dimension distinguishes factors that operate before the injury (e.g., recent snow had made the road slippery) from those that operate at the moment of injury (e.g., the driver was wearing a seat belt). Other factors operate following the collision (e.g., bystanders knew CPR).

Table 10.6: Illustration of Haddon's matrix applied to a motor vehicle collision

	Host	Agent	Environment	
			Physical	Social
Pre-event	Driver in a hurry	Car recently serviced	Road design	Enforcement of speed limits
Event	Wearing seat belt	Air bags working	Icy patch	Number of people standing nearby who may get hurt
Post-event	Has cell phone to call for help	Tendency for car to catch on fire	Emergency vehicle access	Level of assistance provided by bystanders

The matrix is now being used in other situations, including the assessment of medical errors. An advantage of using the matrix is that it gives the full picture of the problem instead of focussing the blame on individuals; it leads to more constructive consideration of ways to reduce this type of error. Table 10.7 illustrates a situation in which potassium was used instead of saline for dissolving a drug that was then administered intravenously, causing severe cardiac arrhythmia and leading to the collapse of the patient.

Table 10.7: Illustration of Haddon's matrix applied to an incident of erroneous intravenous administration of potassium

	Host (patient)	Agent	Environment	
			Physical	Social
Pre-event	Unwell, co-morbidity	Inexperienced nurse; medications packaged in similar containers	Only one drawer in which to store both potassium and saline	Rushed doctor waiting for information on another patient; senior nurse on break
Event	Too sick to enquire about what medication he is being given	Nurse hurriedly mixes medication; monitors patient's pulse while injecting	Night-time; ward lighting low, patient's bed lamp not working.	Nurse being called to other patients; must hurry
Post-event	Immediate signs of arrhythmia	Calls cardiac team	Cardiac trolley was left at the far end of the corridor after an earlier call	Cardiac team responds quickly

As a way to enhance patient safety, completing the matrix identifies hazards as well as protective factors. The protective factors, such as nurses monitoring patients while they administer injections, calling promptly for help, and quick response of the cardiac team should be reinforced, while the risk factors?storage and packaging of solutions, poor lighting, inadequate supervision, and harassment by other team members?can be rectified. (To prevent this type of mistake, many hospitals no longer store potassium on the ward.)

Active versus passive interventions

Active interventions rely on continued compliance for their effectiveness, whereas passive interventions are applied once only. In general, passive interventions are more effective that active ones, because people do not have to remember to use them. For instance, administration of influenza vaccine (passive) confers immunity for the whole influenza season, whereas hand-washing (active) must be repeated several times a day. An airbag (passive), once installed in a car, stays there until it is needed, whereas drivers and passengers must remember and be willing to put on seat belts (active). Particularly in the workplace, people's compliance with the use of safety equipment can be poor, even in well-managed workplaces. For example, ear protectors can be hot and uncomfortable, and may mask auditory danger signals, so that reducing environmental noise levels is preferable, if possible. Safety harnesses can impede mobility, but in some situations they can be replaced by guard rails or safety cages.

Links

To read more about issues in environmental health as well as how Canadian environmental health services are organised, go to the National Collaborating Centre for Environmental Health at: http://www.ncceh.ca/

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Risk Communication

Once a risk has been identified and quantified and methods of reducing it have been found, the information has to be communicated to the people at risk to allow them to understand their risk and take steps to avoid it. Communication is the exchange of information; good communication is the exchange of information in a way that the recipient understands what the sender intends (Figure 10.2). A model of communication that was derived from a mathematical model of communication technology is useful as a basis for thinking about interpersonal communication. This model distinguishes six elements in the communication process: the message, the messenger, the encoding, the channel, the decoding, and the recipient.

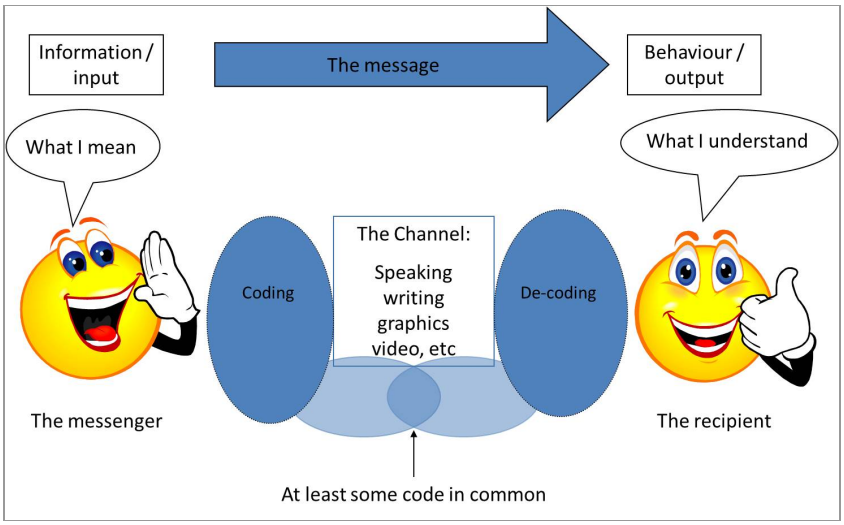


Figure 10.2: The communication process showing the relationships between the elements of communication

The message

How risk is perceived depends on a number of factors besides the actual level of risk. Factors that influence risk perception can be related to the hazard itself (the exposure) or to the possible effects of the hazard (the outcome). The psychological state of the person perceiving the risk, as well as his socio-cultural background, can modify his perception of the hazard and of its potential effects. Factors that increase the perception of danger are listed in Table 10.8.

Table 10.8: Factors that increase public perception of danger[11,12]

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Characteristics of Exposure

Involuntary

Not under personal control

Unnatural (e.g., terrorist attack)

Unfamiliar

Inequitable distribution of risk and benefit (people who suffer the consequences experience no gain from the activity)

No trust in institutions involved

Media attention

Characteristics of the Outcome

Catastrophic (instead of chronic)

Affects children or future generations

Unknown or uncertain outcome

Affects identifiable people, not statistics

Dreaded outcomes (e.g., cancer)

Immediate (vs. delayed)

Irreversible

Media attention

These characteristics can transform a minor statistical risk into a major perceived risk. For example, approximately 3,000 people were killed in the 9/11 attacks, whereas every year the U.S. typically experiences over 40,000 deaths related to road traffic incidents; many Americans were understandably distraught after 9/11, but deaths on the roads hardly attract any attention. From Table 10.8, the difference in reaction was due to the unfamiliarity of the events; they were beyond the control of even the U.S. government; they were catastrophic and attracted a great deal of media attention; they caused the simultaneous death of many identifiable people. By contrast, road traffic fatalities are familiar, under the control of road users, and don't affect many people at any one time. To use a less dramatic example, in 2007 there were 0.012 deaths per million miles flown on U.S. commercial airlines compared to 0.91 deaths per million vehicle miles in urban areas and 2.27 deaths per million vehicle miles in rural areas, yet air travel frightens us more than driving along a country road.

Table 10.8 also explains why some people worry about environmental risks yet refuse to change their own high-risk lifestyles. For example, smokers can be anxious about small environmental risks that they cannot control, while continuing to ignore the great risk of smoking. Many patients fear taking medication even though they are assured that it could help their health problem; they see taking pills as unnatural, unfamiliar and with possible, uncertain secondary effects.

An effective message contains more than simple information; it also implies what the recipient should do with the information. For example, after discussing the risks and benefits of exercise with a patient, a clinician should conclude by relating this information to the patient's personal situation and making it clear that the patient should take more exercise.

The Messenger

People respond more to the attitude of the messenger than to his or her status as a professional or authority. People tend to disregard information given by a recognized expert if he shows a lack of caring or empathy (Figure 10.3). By showing commitment to a listener's well-being and by appearing honest, non-experts may influence listeners far more than an expert who disregards the emotional side of communication.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

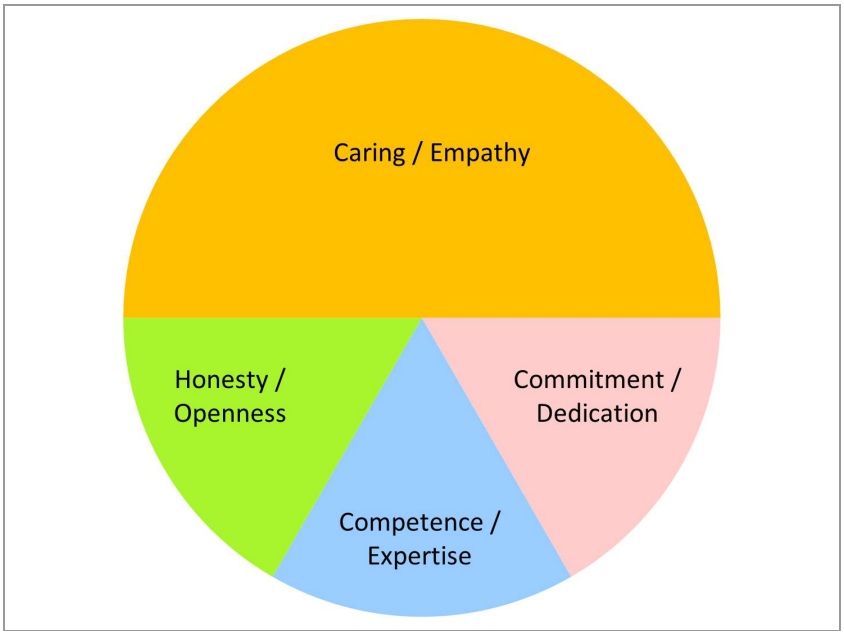


Figure 10.3: Personal qualities of the messenger and their relative effect on how the message is received

The recipient

The recipient is an active participant in communication. The recipient's prior knowledge, beliefs, and attitudes affect his understanding of the message. When communicating with patients, a clinician needs to adapt the message to the patient. Several personal characteristics are likely to affect how a patient interprets the message:

General disposition

Optimistic people tend to feel at low risk. Pessimistic people, and those who are anxious or depressed, tend to overestimate risk and they perceive more threat than optimistic people. They are, therefore, more inclined to use defence mechanisms to reduce feelings of threat. This can go as far as denying the risk entirely.

Affective forecasting

Most people tend to be unrealistically pessimistic about how they will cope with situations they have not experienced. A person who, while still healthy, declares that she would prefer not to be resuscitated if it seems likely that survival might result in serious disability, may have a change of attitude once diagnosed with a serious illness; later she may find that she is better able to cope than she originally thought. In the same way, patients' assessment of the risks and benefits of cancer treatments may vary with the evolution of their emotional state after they learn of the diagnosis.

Perception of threat

Most people feel that they have a lower-than-average chance of getting a severe illness. When shown information about the average risk, people tend to adjust their estimate of their own risk in response and so maintain the belief that they are at less than the average risk. This perception is more pronounced when the health problem is seen as controllable, is likely to occur in the distant future, and occurs in a type of person that the patient considers different from himself. The young smoker believes he has control over his smoking, knows that he is unlikely to feel the effects of smoking for many years, and does not yet consider himself a smoker; he is very unlikely to believe himself at risk of a smoking-related illness.

Confirmatory bias

People are more inclined to retain information that supports their prior beliefs than dissonant information. Cognitive consistency is powerful: even the most objective researcher has a tendency to focus on information that supports his hypothesis. A patient who believes his risk of cancer is low can be inclined not to believe that he has a cancer, or may minimise its significance. Clinicians are inclined to retain their preliminary diagnoses, even in the presence of contrary evidence.

Reduction of vulnerability

The need to feel invulnerable can make people deny or forget information about personal risk. In judging evidence that they are at risk, people may counter the information by questioning the validity or reliability of the source. Similarly, a patient receiving bad news is more likely to request a second opinion than a patient receiving good news. People also tend to find contrary examples to corroborate their denial, so smokers remember their grandfather who smoked twenty cigarettes a day and died in full health at the age of eighty-five.

The channel

The usual channel for clinical communication is the spoken word: clinician and patient talk to each other. However, words can be supplemented with visual aids, such as posters, leaflets, and, occasionally, videos. The recipient must have access to the channel of communication: written materials

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

may not reach people who have difficulty reading.

Encoding and decoding

Information must be coded before it can be passed on. The code might be a language, scientific jargon, or slang; it can also take the form of an illustration or a diagram. For communication to be successful, the messenger and the recipient must share a common understanding of at least part of the code. Even though a clinician speaks the same language as the patient, differences in socio-economic milieu, education and experience can limit their common vocabulary. When a physician discusses an appendectomy with a patient, the physician is talking about a routine procedure that the patient is likely to recover from in a few days. The patient, on the other hand, is hearing about an alarming and unique experience that is likely to be painful, will leave a permanent scar, and will disrupt life for, at the very least, several days.

When communicating with their patients, clinicians should:

- use words and concepts that their patients can understand,
- remember that clinicians are familiar with medical conditions and procedures, whereas their patients are not,
- modify their tone and body language to conform to the common code, as well as be aware that gestures may mean different things in different cultures. In North America, a Greek who shakes his head may be asking a question, not saying "no," while an Italian gesturing for someone to come closer may appear to be waving goodbye.

Information about risk is often coded as numbers or graphs. Most people can grasp the information in these forms, but there are a number of common pitfalls:

Emphasis on gain versus emphasis on loss

People tend to be averse to loss and will do more to avert a perceived loss than to achieve a gain. A well-known experiment describes responses to two sets of options for an imaginary disease:[13] 600 people are affected by a fatal illness, and we must choose between two forms of treatment:

Treatment A will save 200 people.

Treatment B has a one-third possibility of saving all 600 people, and a two-thirds probability of saving no one.

Most people opt for Treatment A because the prospect of saving 200 lives is more attractive than the possibility of losing 600, even though the mathematical value of the two choices is equivalent?200 versus one-third of 600. However, when Treatment A is expressed as a loss

Treatment A will allow 400 people to die.

Treatment B has a one-third possibility of saving everyone and a two-thirds probability that all 600 will die.

Most people opt for Treatment B because the one-third possibility of saving everyone is more attractive than the certainty of losing 400. The two sets of choices are mathematically equivalent, but loss aversion affects the choice. The way the choice is formulated is called the ?frame' and, to ensure informed choice, a clinician should communicate the information using a number of alternative frames.

Definition

Frame: in relation to communication and decision-making refers to the decision maker's understanding of the acts, outcomes, and contingencies associated with a particular choice.[13] Mentioning specific aspects of the risks and benefits can alter the patient's frame and influence their choice.

The default option

The option that is presented as the usual one is more likely to be chosen than the one presented as the alternative.

Numbers versus proportions

Data expressed as proportions tend to be seen as relatively benign, whereas data expressed as a frequency tend to engage people much more. In one experiment, the case of a mentally disturbed patient was presented to physicians. They were told that 20 out of 100 patients similar to this one were likely to commit an act of violence. They were then asked if they would discharge the patient, and 41% said they would not. Another similar group of physicians was presented with the same scenario, but told that the patient had a 20% chance of committing an act of violence: only 21% of these physician refused discharge.

Glass half full or glass half empty

Emily and Ian are psychology students. In the last exam, both answered all the questions: Emily got 74% correct while Ian got 26% incorrect. Which is the better student? It is usually found that positive framing leads to positive feelings and negative framing leads to negative ones, so Emily is usually judged to be better than Ian. In discussing the risk of Alzheimer's disease, dwelling on the 8% of people over sixty-five years old who have it gives a worse impression of the situation than dwelling on the 92% of people who don't.

People don't come in halves

People are more at ease with whole numbers, so fractions and decimal places should be avoided. Few estimates of risk or of therapeutic benefit are so precise that they merit decimals.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Numerator and denominator

People tend to focus on the numerator and ignore the denominator. A disease which afflicts ten in a hundred people tends to be seen as less common than one which afflicts a hundred in a thousand people. When discussing a single risk with a patient it can be helpful to express it in several ways: 10%, or one in ten, or ten in a hundred. However, when asking people to compare risks of different outcomes, one should keep the denominator constant. It is difficult for most people to understand the difference between one in five risk of one outcome and a 25% risk of another outcome. They will find it easier if a 20% risk is compared to a 25% one or if a one in five risk is compared with a one in four risk.

Relative and absolute risk

Relative risks may be particularly difficult to interpret because people do not always know the context; this is another example of the distinction between the patient and the population perspectives. Although hormone replacement therapy doubles the risk of breast cancer, it causes only eight additional cases of breast cancer in 10,000 woman years. For an individual woman, doubling the risk does not increase it greatly because the baseline risk is so small, whereas, on a population level, a doubling of the risk may be significant. It is advisable to use only absolute risks when communicating with individuals, because proportional changes may often obscure a lack of substantive importance.[14]

Links

Patient Decision Aids.

It is often difficult to communicate the notion of risk to a patient in a readily comprehensible way. Annette O'Connor developed a series of patient decision aids as a practical tool to help patients choose between multiple options, each with different risks and benefits. See http://decisionaid.ohri.ca/mission.html

Uncertainty

In communication, uncertainty can occur for several reasons:

1. Words can be ambiguous and their meaning can vary. Clinicians should make certain that their patients understand the precise meaning of the possible outcome and the course of action suggested. If numbers are used instead of words, be aware of the biases inherent in a positive versus a negative framing of the statistics.

2. Terms to describe risk, such as high or low, are open to varying interpretation.

3. The definition of the risk may lack specificity. The risk should always include a time frame. A 10% chance of death from a lung cancer over a lifetime is not the same as a 10% chance in the next five years. A 20% chance of loss of function could be interpreted as a 20% reduction in function, or that 20% of people will have complete loss of function, or that 20% of people will have some loss of function.

When discussing risk with the patient, the information should be communicated in different ways, with both positive and negative framing, to help the patient arrive at a fully informed decision.

Self-test questions

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Self-Testing

1. A patient is complaining of a skin rash, which you think looks like contact dermatitis. What should you question her about?
Contact dermatitis, along with many other skin diseases, is generally not due to intrinsic factors, but is an environmental disease. Questions should cover all spheres of the patient's life, following the mnemonic CH2OPD2:
- *Community acquired? In this case an unlikely source although contact with plants, or use of herbicides or insecticides in the neighbourhood should be considered*
- *Home: consider contact with garden or house plants; use of cleaning products, pesticides, herbicides, construction or decorative materials*
- *Hobbies and leisure: exposure to chemicals, dusts, or micro-organisms, or to particular clothing necessary for the activity ? wet suits for surfing, gloves for gardening ? depending on which part of the body is affected*
- *Occupation: contact dermatitis is common in occupational settings. You may need to ask about previous occupations; don't forget voluntary work; work with known hazards; the protective clothing that is used for protection, such as latex gloves*
- *Personal habits: hygiene products including scents, creams and lotions, source of water used for washing; products used for washing clothes and bedding.*
- *Diet: an unlikely source of contact dermatitis*
- *Drugs: prescription, non-prescription, and alternative medications, particularly medications applied to the skin ? nicotine or hormone patches.*

You should question the patient particularly about changes in the period prior to the onset of symptoms. Where you find a possible source, make sure you get precise information about the period of exposure, how great the exposure is, the variation of symptoms in relation to episodes of exposure, whether other people were exposed and if they have similar symptoms.

2. Name and briefly describe the stages of risk assessment.
***Hazard identification**: explore the environment to identify possible sources of hazard. . In the clinical setting, a good history usually identifies possible sources.*
***Risk characterization**: describe the effects of the hazard. The local public health department will usually be able to advise on the effects of different hazards and about populations at particular risk.*
***Exposure assessment**: describe the duration of exposure and the levels of the hazard. In the clinical setting a good history should include precise details on the level of hazard to which the patient is exposed, the length of time of an exposure, how often the exposure occurs, and since when has this been happening, in other words to gauge the dose to which the patient is exposed. Tests may be available to measure the dose the patient has had. It may also be possible to measure the level of hazard in the patient's environment.*
***Risk estimation**: integrate and analyse the information from the three previous steps to judge how much the hazard contributes to patient's problem.*

3. List the psychological influences on interpretation of quantitative data.
In relation to the person who is at risk
***General disposition:** optimistic people believe themselves to be less at risk than pessimistic people.*
***Confirmatory bias**: people pay more attention to information that confirms what they already believe.*
***Affective forecasting:** people find it difficult to assess how they will react in future possible situations.*

In relation to the hazard
***Perception of threat:** people generally believe that they are at less than the average risk, particularly when they have control over their exposure to the hazard.*
***Reduction of vulnerability:** people wish not to feel vulnerable to a perceived threat and may therefore deny their risk.*

In relation to how the risk is communicated
***Loss aversion:** people will avoid a choice that is presented as a loss.*
***Default option:** people will choose the option that they think is the usual or ?normal' one.*
***Number versus proportion** and **numerator versus denominator:** people pay attention to whole numbers that can represent whole people. They can misunderstand proportions, and pay little attention to denominators.*
***Glass half full or glass half empty:** people will focus on the number that is presented, not on the one that is understood. If a disease incidence is presented as such, they will magnify the risk of becoming ill, rather than the risk of staying well.*
***Relative vs absolute risk**: people interpret a high relative risk as almost equivalent to a high absolute risk, even if the underlying absolute risk in the ?high risk' group is small.*

4. A noisy machine is causing some workers to complain of deafness. What could you suggest to alleviate the problem?
Using the source-path-receiver model,
***Source** ? modify or re-design the machine to reduce the noise at its source, use a different type of machine or process that doesn't produce noise, relocate it away from the workers or enclose it to confine the noise produced.*
***Path** ? use acoustic barriers to absorb or block the noise reaching the workers.*
***Receiver** ? enclose or relocate the workers away from the machine, supply workers with ear protectors and ensure that exposure to the noise in infrequent and for very short periods.*

Reflection Questions

- 1. A 42-year-old female patient living a kilometre away from overhead, high tension power lines is very worried of the lines' effect on her health. How do you approach this case?
- 2. You find her risk from the high tension lines is negligible. How do you reassure her?
- 3. The same patient is not very worried about smoking ten cigarettes a day. How do you explain this?

References

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Further Reading

World Health Organization. The world health report 2002: Reducing risks and promoting healthy life. Geneva, Switzerland: WHO; 2002.

Covello VT. Best practices in public health risk and crisis communication. J Health Communication 2003;8(Suppl 1):5-8.

Klein WM, Stefanek ME. Cancer risk elicitation and communication: lessons from the psychology of risk perception. CA: a Cancer Journal for Clinicians 2007;57(3):147-167.

Tversky A, Kahneman D. Judgement under uncertainty: heuristics and biases. Science 1974;185:1124-1131

1. Marshall L, Weir E, Abelsohn A, Sanborn MD. Identifying and managing adverse environmental health effects: 1. Taking an exposure history. CMAJ 2002 Apr 16;166(8):1049-1055.

2. Porta M. Dictionary of epidemiology. New York: Oxford University Press; 2008.

3. Aidoo M, Terlouw DJ, Kolczak MS, McElroy PD, ter Kuile FO, Kariuki S, et al. Protective effects of the sickle cell gene against malaria morbidity and mortality. Lancet 2002 Apr 13;359(9314):1311-1312.

4. Dorland (editor). Illustrated Medical Dictionary. 37th ed. Philadelphia, PA.: Saunders Elsevier; 2007.

5. Canadian Centre for Occupational Health and Safety. WHMIS - Labelling Requirements. 2001; Available at: http://www.ccohs.ca/oshanswers/legisl/msds_lab.html. Accessed July, 2010.

6. Theoret JJ. Chronology of the Thalidomide Story. Can Med Assoc J 1962;87(18):981-982.

7. Cox LA,Jr. Does concern-driven risk management provide a viable alternative to QRA?. Risk Analysis 2007 Feb;27(1):27-43.

8. Ontario Ministry of the Environment. Drinking water system at the Kashechewan First Nation: Technical Report. 2005; Available at: http://www.ene.gov.on.ca/envision/techdocs/5445e.pdf. Accessed July, 2009.

9. Eggertson L. Despite federal promises, First Nations' water problems persist. CMAJ 2008 Apr 8;178(8):985.

10. Haddon W,Jr. A logical framework for categorizing highway safety phenomena and activity. Journal of Trauma-Injury Infection & Critical Care 1972 Mar;12(3):193-207.

11. Slovic P. Perception of risk. Science 1987 Apr 17;236(4799):280-285.

12. Covello VT, Peters RG, Wojtecki JG, Hyde RC. Risk communication, the West Nile virus epidemic, and bioterrorism: responding to the communication challenges posed by the intentional or unintentional release of a pathogen in an urban setting. Journal of Urban Health 2001 Jun;78(2):382-391.

13. Tversky A, Kahneman D. The framing of decisions and the psychology of choice. Science 1981 Jan 30;211(4481):453-458.

14. Spasoff RA. Epidemiologic methods for health policy. New York: Oxford University Press; 1999.

AFMC Primer on Population Health

Chapter 11 Infectious Disease Control

Clinicians are key players in controlling the spread of infectious disease. They are typically the first to see new cases of a disease that could indicate the beginning of an **epidemic**. They must report these to public health authorities. At times, they are called upon to collaborate with infection control or with public health teams in order to ensure good control practices and to arrest outbreaks of infectious disease. Clinicians are also in a position to recognise changing patterns of non-infectious disease, for example an unusually high number of injuries that may be associated with a change in the built environment. In these cases, clinicians can also play a role in identifying and managing the situation, although the process for doing so may not be as clearly defined as the process of control and management of infectious diseases. Finally, clinicians must be aware that they themselves or the equipment they use can transmit infection to their patients and must take steps to prevent this.

Chapter Objectives

After reading this chapter, you will be able to:

- Know the defining characteristics of an outbreak and how to recognize one when it occurs;
- Demonstrate essential skills involved in controlling an outbreak and its impact on the public, in collaboration with public health authorities as appropriate;
- Define an outbreak in terms of an excess of cases, beyond the number usually expected;
- Describe and understand the main steps in outbreak management and prevention;
- Demonstrate skills in effective outbreak management including infection control when the outbreak is due to an infectious agent;
- Describe the different types of infection control practices and justify which type is most appropriately implemented for different outbreak conditions; and
- Describe appropriate approaches to prevent or reduce the risk of the outbreak recurring:
    1. Understand surveillance systems and the role of physicians and public health in reporting and responding to disease.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Case Study

Dr. Rao received a phone call from Mrs Richards's long-term care home to say that she has had some diarrhoea. He went to see her. She says she had a bit of stomach pain last night but is better today; she just had one bout of fairly liquid stool. Mrs. Richards thinks that other people in the home have been sick also. She heard that one resident had vomited two days ago and another had stomach cramps the other night. Dr. Rao tries to get more precise information; he wants to know if this number of people with stomach complaints is greater than normal. He also realizes he should ask about precisely who was sick, when, for how long, the location of the rooms of the sick people, if any of the staff have been sick, and if the Public Health department has been informed. However, the staff member now on duty was away for the last few days and knows only what was mentioned in the daily report.

Detection and control of outbreaks

Clinicians are often in the best position to recognize an **outbreak** of disease; sometimes they see an unusually high number of people with the same disease and at other times patients report that they know other people with similar symptoms. Even if the clinician does not suspect an outbreak, his reporting contributes to the detection of outbreaks. Filling in discharge summaries and death certificates, and reporting cases of notifiable disease all contribute to health surveillance. Ensuring that the forms are completed accurately and transferred in a timely fashion makes a major contribution to outbreak detection and general health **surveillance**.

Depending on the type of organism involved, the conditions of spread and the target population, outbreaks can be acute and fast-moving, such as gastroenteritis in a nursery school or long-term care home, or they can evolve more slowly, such as the AIDS pandemic. While public health authorities are ultimately responsible for ensuring the detection and control of outbreaks, clinicians are major players in this area because they are usually the first point of contact with the affected population. Likewise, hospital infection control teams rely on the cooperation of clinicians in preventing infections. The basic steps in outbreak control and management are:

1. Establish the existence of an outbreak
2. Define what constitutes a case and identify cases as they occur
3. Formulate hypotheses on the causes, and implement initial control measures
4. Test the hypotheses through analysis of surveillance data or special studies
5. Draw conclusions and re-adjust hypotheses and control measures if needed
6. Plan for long-term prevention and control.

Stage 1: Establishing the existence of an outbreak

Is it an outbreak?

It is important to identify an **outbreak** quickly so that it can be addressed before it develops into a full epidemic. However, there is no hard-and-fast rule for defining an outbreak. A **cluster** of cases could be due to chance or could form the beginning of an outbreak. This is especially true of rare diseases and diseases in small populations, where an absolute increase of a very small number of cases could represent a large relative increase in an illness in the population. For instance, in a community where there are usually two cases of a given disease in one month, four cases in one month (relative increase of 100%, absolute increase of 2 cases) may or may not constitute an outbreak. In such instances, it can be very difficult to decide how much time and money to spend on investigating something that might be just a chance occurrence. Consultation with epidemiologists or statisticians may be required, but the statistical techniques for assessing the influence of chance in disease clusters are still in development.

Furthermore, there may be biases in documenting an outbreak. When a possible outbreak is suspected, more cases of the illness than usual may be reported. Patients who are worried about the illness may be more sensitive than usual to their symptoms and may seek care (and hence their case gets reported) when they would not normally do so. Clinicians may be more assiduous than usual about requesting laboratory tests and reporting cases. In addition, **case-finding** or active surveillance by public health workers will uncover cases which would not ordinarily have been identified.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Definition

Endemic is the constant presence of a disease or infectious agent within a given area or population group.[1]

Cluster is a collection of new cases of an uncommon disease that occur so closely together in space and/or time as to arouse suspicion that this is not a chance occurrence.[1] The distinction between cluster and epidemic is subtle, but cluster suggests a narrower geographic distribution than epidemic, perhaps an early stage in the development of an outbreak or epidemic. It also applies to both communicable and non-communicable diseases.

Outbreak is two or more cases of illness thought to be linked in time and place.[2] An outbreak can also refer to a small, localized cluster of cases, usually of an infectious disease that may be a prelude to a broader epidemic.[1]

Epidemic is the occurrence of illness clearly in excess of normal expectancy.[1] The number of new cases needed to declare an epidemic depends on the disease, the time, and the location. For example, a single case of yellow fever could be considered to be an epidemic in Canada, but it might not be in a tropical region.

Pandemic is an epidemic occurring over much of the world. Examples include influenza (1918?19 and 1957?58), cholera (1961 to present), AIDS (1988 to present), and the tobacco addiction (present).[1]

Great care is needed in deciding whether or not to declare an epidemic: balancing the value of alerting agencies to the need for control measures against possible concerns over causing public panic. National and regional public health agencies are responsible for these decisions. A **pandemic** can only be declared after careful evaluation by the WHO, which may issue travel advisories and, again, such announcements are taken only after serious consideration because of their potential impact on the economy of regions affected. Just as a local outbreak runs a natural course (see Chapter 7), so does a pandemic and the WHO has proposed a sequence of stages shown in Figure 11.1:

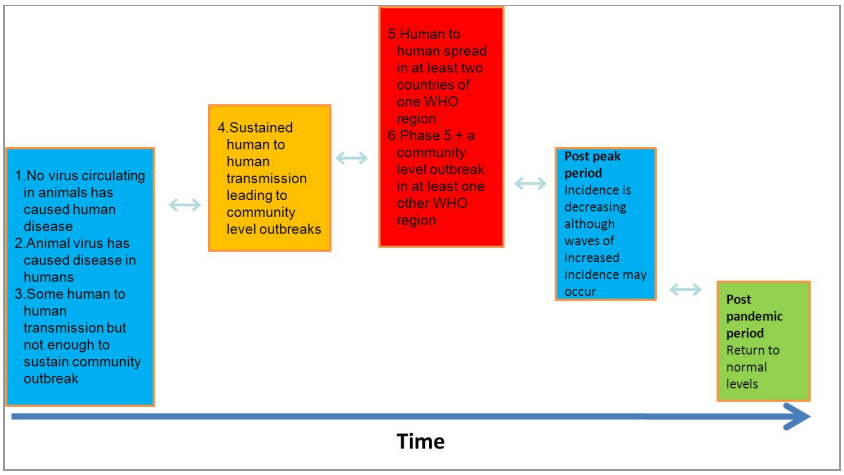


Figure 11.1: World Health Organization levels of pandemic (Source: http://www.who.int/csr/disease/influenza/pandemic_phase_descriptions_and_actions.pdf)

Infectious disease epidemics can result from:[3]

1. *Increased virulence of the infecting organism*. Micro-organisms have a number of mechanisms to alter their **virulence**. For example, *Corynebacterium diphtheria* must be infected with a specific bacteriophage in order to produce the diphtheria toxin that causes diphtheria. Plasmid exchange can confer antibiotic resistance in previously susceptible bacteria. Another example is the virulence of the influenza virus, which varies with its genetic drifts and shifts.

2. *Recent introduction of the organism into new setting*. Europeans introduced measles and smallpox to the Americas, where they swiftly decimated the local populations who had no immunity to them. The two diseases were among the causes of the decline of the Aztec empire. More recently, the severity of the 2002-03 North American epidemic of West Nile virus may have been in part due to the lack of immunity of the host birds in the region, as well as the lack of immunity of the human population (see Figure 11.2, Nerd's corner).

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

West Nile Virus in North America

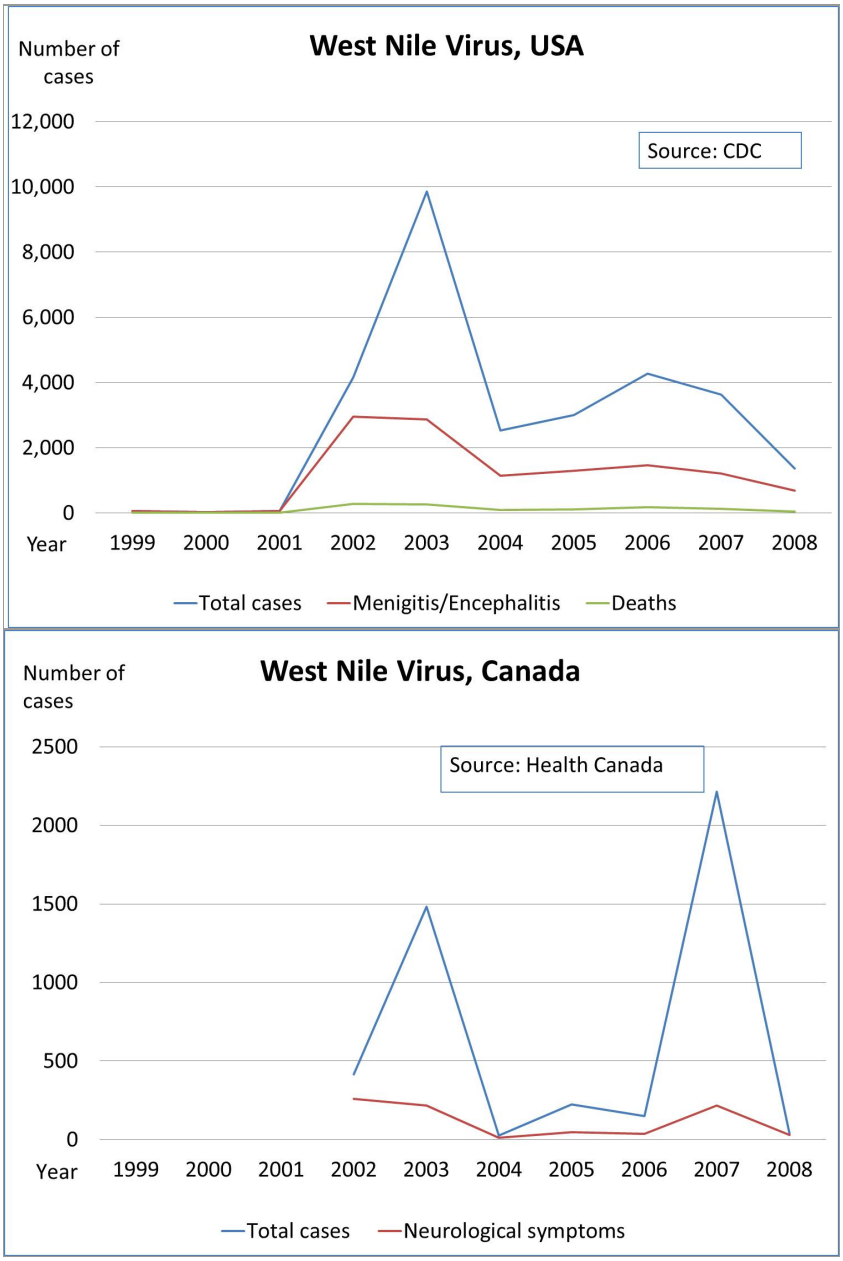


Figure 11.2: Patterns of West Nile virus infections in USA and Canada, 1999-2008.

West Nile virus was first documented in Uganda in 1937. In the 1990s, in Europe, the virus seemed to increase in virulence. In North America, it was first detected in birds in New York. In 2002 and 2003, it caused a large epidemic in North America, much more widespread than in any other continents. Reasons for its epidemic potential in North America may include:[4]

? More virulent strain

? No previous immunity in birds - the primary host

? The carrier mosquito, *Culex pipiens*, mainly feeds on birds, but switches to feeding on humans in fall when host birds migrate

? Age structure of non-immune individuals?older and more susceptible to neurological effects.

3. *Enhanced transmission so that more susceptible people are exposed*. Troop movements and population upheavals during the 1914?1918 war brought many more people than usual into close contact with one another. This enhanced the transmission of the influenza virus that caused the 1918 pandemic.

4. *Change in host susceptibility*. People with HIV infection are particularly susceptible to tuberculosis. The high incidence of tuberculosis world-wide is, in part, due to the HIV epidemic causing a rise in the number of susceptible people.

5. *New portals of entry or increased exposure*. Technical developments in health care that require invasive instrumentation have contributed to the rise in nosocomial infections.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Illustrative Materials

Walkerton, Ontario[5]

On May 18, 2000, a paediatrician at the Owen Sound hospital admitted a nine year old girl with bloody diarrhoea and a seven year old boy with fever and abdominal pain. The boy later developed bloody diarrhoea. On investigating possible links, the attending paediatrician found that the children attended the same school. The same day, people from Walkerton started contacting their Public Utilities Commission because they suspected problems with the water supply.

On May 19th, having learned that more people in the community were suffering from diarrhoeal illnesses, and having conducted a food history with the parents of her two patients, the paediatrician suspected that they were suffering from water-borne *Escherichia coli* infection. She contacted the Bruce-Grey-Owen Sound Health Unit. Later that day, the managing director of a Walkerton retirement home also contacted the health unit to report that three residents had diarrhoeal disease.

The health unit contacted schools in the area and learned that abnormally high numbers of pupils were absent with illness. The unit also learned that the Walkerton emergency room had seen eight patients with diarrhoeal disease. At first, the pattern of disease suggested a food-borne outbreak, but, as many members of the public believed the outbreak to be due to contaminated water, the health unit also began to investigate the water supply.

The Walkerton outbreak was caused by contamination of the municipal water supply with *E. coli* and *Campylobacter jejuni*. By August 31st, 2000, it had killed seven people, caused Haemolytic-Uraemic Syndrome in 27 people and, it was estimated, made over 2,000 people ill. Poor supervision of water treatment combined with heavy rain, which had washed cattle effluent into the water supplying the treatment plant, were the major causes of the outbreak.

Stage 2: Define and identify cases

A crucial early step in investigating a possible outbreak is to define what constitutes a case, as the case definition will be used in the search for more cases, who will then be questioned about their exposures. The case definition describes precisely the symptoms, signs, history, or test results that indicate a probable case of disease. Clinicians may be asked to help in the search to uncover all cases, and they may also have valuable information linking the cases, which can suggest a possible cause. Because of biological variation, the presenting symptoms and signs for cases of any illness vary. The case definition should neither be too broad nor too narrow. In an outbreak of gastro-enteritis, for instance, some people may only have mild abdominal cramps, whereas others have diarrhoea and vomiting, with or without fever, muscle pain, headache, and dehydration, etc. If the case definition includes people with any one of the full range of symptoms (fever, or muscle pain, or headache, etc), it will be broad enough to include a large proportion of the population, many of whose symptoms may be unrelated to the outbreak under investigation. Conversely, if only a narrow case definition is used (fever and muscle pain, and headache), it may exclude many cases, hence underestimating the extent of the outbreak and possibly delaying the implementation of extended control measures. If there are indications of a common exposure, for instance if a large proportion of the initial cases say they ate at the same restaurant, the common exposure can be included as a criterion in the case definition to specify the cases involved and the investigation could focus on identifying the particular foodstuff at the root of the problem.

Links

Sensitivity and specificity

When deciding how inclusive the case definition should be, remember the concepts of sensitivity and specificity in Chapter 6.

Definition

Case definition is a set of criteria that must be fulfilled in order to identify a person representing a case of a particular disease as part of an outbreak. These criteria are not necessarily diagnostic, but may include geographic, clinical, and laboratory criteria, and may be combined into a scoring system. The case definition is used to identify probable cases in epidemiological surveillance. It is not used for diagnosis.[1]

Stage 3: Formulate hypotheses and implement initial control measures

Using the information gathered in the initial steps of the investigation, public health professionals describe the epidemic in time, place, and person. An outbreak progresses, they may draw an **epidemic curve** to track its evolution in time (see epidemic curves in Chapter 7). They describe in detail the circumstances of the outbreak and the demographic characteristics of the people affected. Thus, the patterns described usually indicate the likely source of the outbreak and the population at high risk. This is generally sufficient to suggest some initial control measures.

Management of cases

In general, cases in an outbreak are managed by their usual physicians unless the size of the outbreak justifies setting up special clinics. One role of public health officials is to keep the clinicians and the public informed of control procedures. Routine precautions are recommended to reduce the risk of transmission via contact with asymptomatic patients or contaminated environmental sources. Appropriate precautions to address contact, droplet, or airborne transmission are implemented according to the clinical picture and presumed modes of transmission.

Table 11.1: Typical precautionary measures according to route of transmission (recommendations vary according to the organism involved).

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Type of transmission	Precautionary measures
Routine:	Hand washing, barrier precautions: gloves and gowns if spread is via contact or contamination, masks and eye protection if splash or spray
Contact spread:	Gowns, gloves for direct patient care; dedicated equipment; hand washing
Droplet:	Surgical mask, eye protection or face shields
Airborne:	Negative pressure ventilation (like a laboratory fume hood, air is sucked into the patient's room and then vented outside, rather than flowing from the room into the corridor and other rooms), isolation, N95 respirators, gowns, gloves, protective eye wear for procedures.

Population management

The public health service is responsible for managing outbreaks in the community; in hospitals, the infection control team is responsible. Clinicians are often called upon to collaborate in the management. People at risk from the outbreak?that is, those exposed or likely to be exposed to the probable source?are identified. People at risk need

- *Information about their risk and how to reduce it*. For instance, forestry workers should use adequate clothing and know the symptoms of Lyme Disease so they seek treatment early

- *Personal preventive measures.* For instance, close contacts of meningococcal meningitis should have antibiotic prophylaxis; health care workers may need HIV prophylaxis after a needlestick injury; contacts of people with a disease for which there is a vaccine could be immunized; sexual partners of people with sexually transmitted infections may require treatment

- *Reducing the risk of propagation*. For instance, this can be done by putting asymptomatic contacts into quarantine for the duration of the disease's incubation period, or by asking people to "cough into your sleeve" during the flu season. Adequate treatment of the infection also reduces transmission.

Illustrative Materials

An outbreak of mumps in the Maritimes

In Halifax, Nova Scotia, in 2006 and 2007, two successive outbreaks of mumps affected a total of 34 people.[6] Although there was no obvious link between the outbreaks, they were thought to be part of a group of outbreaks that occurred in North-Eastern USA and Québec, and which may have originated in the UK. The Nova Scotia outbreaks were managed by voluntary isolation of the cases for nine days after symptoms appeared. Households and other contacts of the cases were immunized. All cases had already been immunized against mumps. It was decided not to immunize all those at risk of the illness because

- Mumps immunization levels were already high
- Spread of the outbreak was limited and slow, and there was no evidence of spread to unimmunized groups
- It was assumed that **herd immunity** was effective because of high immunization levels and the limited spread of the disease
- Experience with a similar outbreak in the UK showed a poor cost-benefit ratio
- The population at risk was large.

Stage 4: Test hypotheses through analysis of surveillance data or special studies

Once there is a hypothesis about the cause of the outbreak, it should be tested. If removing the suspected source is followed by a decline in the outbreak, the hypothesis may be correct, although the decline could still be serendipitous. Cross-sectional studies can be carried out in the case of food-borne outbreaks in a limited population. For instance, in the case of an outbreak associated with a social gathering, guests can be asked about symptoms and about foods eaten to identify particular foods that are likely to have been contaminated. Case-control studies may also be necessary, and clinicians may be required to contribute in case-finding or by providing data on their patients.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Illustrative Materials

The Broad Street Pump

John Snow was a British physician considered by many to be the father of epidemiology. He identified the source of the London cholera outbreak of 1854 that killed 500 people over a ten day period. Snow plotted the location of all the cholera deaths on a map of central London and interviewed seventy-seven surviving cases. He noticed that a water pump located on Broad Street was in the centre of the area where most cases lived. Cases in outlying areas also tended to draw their water on Broad Street, ironically because of its reputation for pure water.

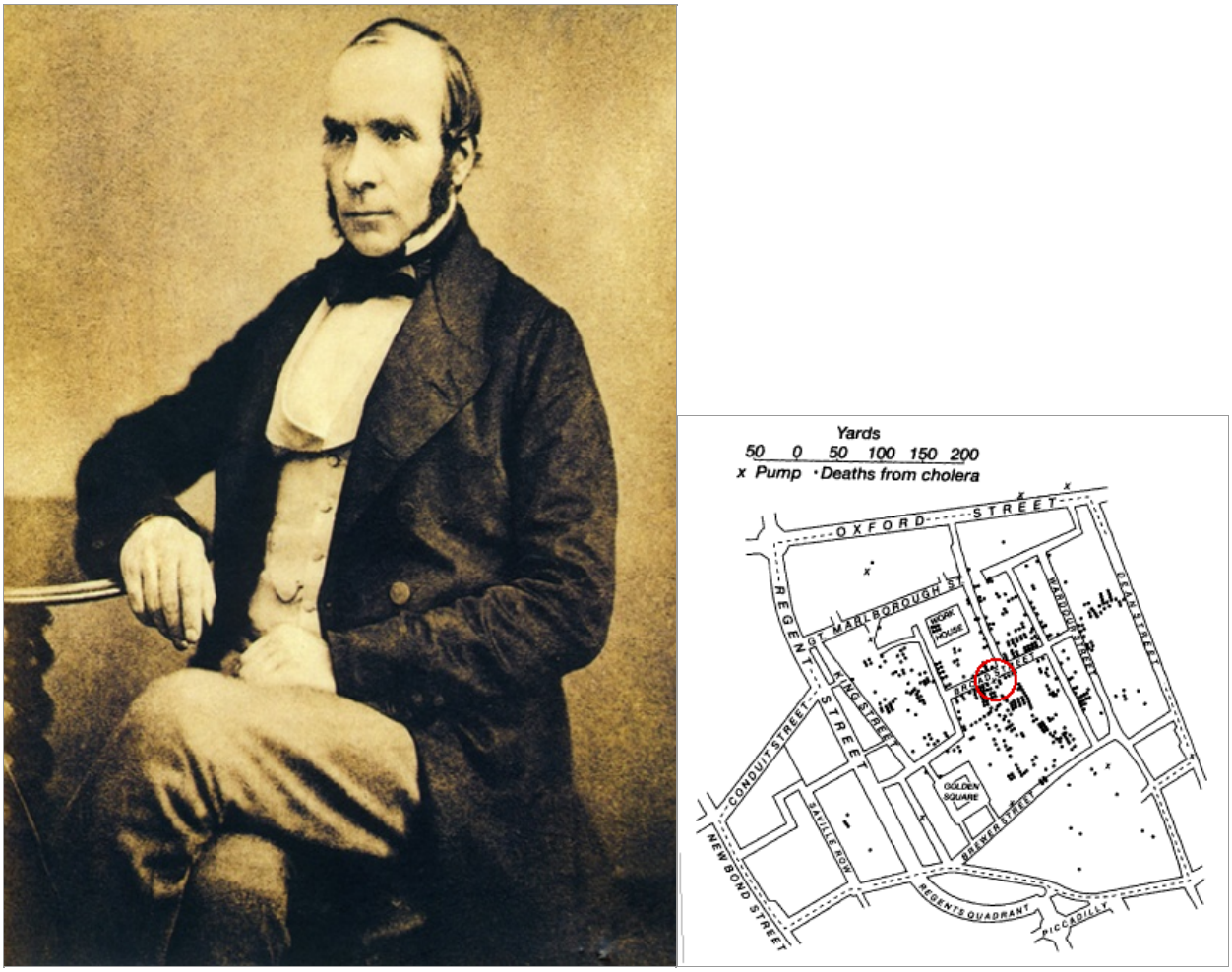


Figure 11.3: John Snow, and his map of cholera cases in central London

Although, at the time, cholera was thought to be caused by bad air, Snow's work led him to suspect an association with water. He convinced the local council to remove the handle of the pump and, according to legend, this stopped the epidemic. In fact, although Snow's work was a breakthrough in surveillance, the epidemic was almost over by the time the handle was removed.

Stage 5: Draw conclusions and re-adjust hypothesis and control measures

Once the cause of the outbreak has been confirmed, the initial control measures may need to be adjusted. If the source cannot be identified, more epidemiological detective work may be required. Based on the results of the outbreak investigation, it may also be possible to recommend action to prevent future similar outbreaks.

Stage 6: Plan for long-term prevention and control

Outbreak investigations should be designed to indicate the possible long-term preventive actions and ways of improving response to future similar outbreaks. When available and safe, active immunization is a very successful way of preventing outbreaks of infectious disease. As long as the vaccine provides lasting immunity, the person remains protected, at least to some degree. Vaccines have created an opportunity for eradication of disease. Smallpox, a virulent and once common illness that afflicted humans for centuries, was eradicated in 1977. Currently, the World Health Organization is working to eradicate polio, which, in 2010, is endemic in only four countries, although still appears in epidemic form in twenty-three.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Links

Immunisation

For more on the eradication of polio, see The Global Polio Eradication Initiative website: http://www.polioeradication.org/

The Canadian Immunization Guide lists the vaccines recommended for Canadians. It can be found at the Public Health Agency's website: http://www.phac-aspc.gc.ca/publicat/cig-gci/

Other measures for long-term control may also be needed. The 2008 outbreak of listeriosis in Canada resulted in discussions about legislative control on food manufacturers. In the case of the water borne illnesses in the Kashechewan First Nations community, recommendations to prevent further problems included upgrading the water treatment plant as well as the water treatment processes, training of personnel, and improved procedures if water quality does not meet the set standards.

Transmission of infectious disease

Knowledge of the natural history of infectious disease is important in controlling its transmission. Contacts of a case should be monitored for the duration of the incubation period of the disease. It might be necessary to isolate cases and quarantine contacts for the period of communicability. Other ways of preventing transmission can be identified using the agent-host-environment model described in Chapter 2:

Agent

1. Destruction or weakening of the agent outside the body (e.g., by cleansers, incineration)

2. Destruction or weakening of the agent inside the body (e.g., use of appropriate antibiotic or antiviral for an appropriate length of time)

Host

1. Physical barriers (e.g., masks, long trousers when hiking in regions that have Lyme Disease)

2. Decrease opportunity for transmission (e.g., by isolating infective cases, quarantine of contacts, coughing into one's sleeve, prompt adequate treatment of cases)

3. Increase specific resistance (e.g., immunization, prophylaxis)

4. Increase general resistance (e.g., nutrition, exercise)

Environment

1. Hygiene and sanitation (e.g., clean water, sewers, adequate housing)

2. **Vector** control (e.g., drainage of mosquito breeding sites, hand hygiene, and respect of infection control measures by health care providers)

3. Health programmes (e.g., free immunization programmes with call/recall system, partner notification for sexually transmissible diseases).

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

A geographer's perspective.

Jacques May, a medical geographer, was the first to propose the ecological model of health. He viewed disease as a reduction in survival caused by maladjustment of a person to the environment.[7] The capacity of individuals to adjust to their environment is genetic, although it is generally mediated by the traits and behaviours associated with their culture. May cites the example of a high haemoglobin level, which improves the survival of people living in high mountain areas, and is therefore normal for them, but which would indicate disease in people living at sea level.

These ideas led to the development of May's model of the elements of transmissible disease, which include:[8]

- Agent: causes the disease (e.g., *Plasmodium* causing malaria);
- Vector: transmits the agent to the host (e.g., *Anopheles* mosquito in malaria);
- Geogens: physical environmental factors (e.g., humidity, temperature, vegetation which allow the survival of *Plasmodium* and *Anopheles*);
- Host: susceptible human.

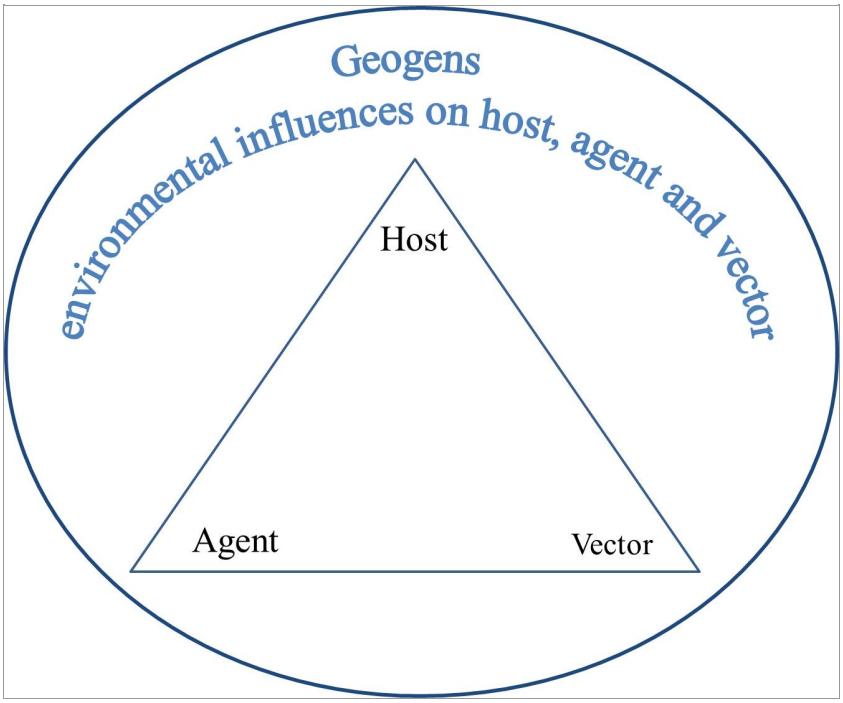


Figure 11.4 May's model of transmissible disease showing the epidemiological triangle situated in the environment.

Transmissible diseases have a number of characteristics that are important to consider when intervening to interrupt their transmission. These include

The **incubation period**: the length of time between infection and appearance of first symptoms or signs. People who have been in contact with infectious disease should be observed for symptoms during this period. The time between the likely exposure to infection and the development of symptoms, indicating the probable incubation period, can help in the differential diagnosis. For instance, a case of gastroenteritis that occurs rapidly after eating a suspect food is likely to be caused by a toxin, such as that produced by *Staphylococcus*, whereas a delay of 6 hours or more indicates an infectious agent, such as *Salmonella*.

Window period: the time between infection and when the infection can be detected. Some diseases, for instance HIV infection, have a long window period, but can be transmitted during it. In the early stages of HIV infection, negative tests can falsely reassure the patient, who may continue risky behaviour, putting others at risk.

Period of communicability: the length of time during which a person can transmit a disease. To prevent transmission, the appropriate precautions should be taken through the whole of this period. Some organisms, such as the hepatitis B virus or *Salmonella typhi* can cause a chronic carrier state in which the patient, having apparently recovered from the disease, continues to transmit it to susceptible people.

Further Reading

Typhoid Mary

Typhoid Mary, a cook New York at the turn of the last century, was a salmonella carrier who was probably responsible for about 55 cases of typhoid. To read an account of her, go to: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1911442/pdf/bullnyacadmed00595-0063.pdf

Attack rate: the proportion of exposed persons who become infected. This is a marker of **infectivity** which refers to the ability of an organism to invade, survive, and multiply in the host. For instance, HIV is not very infective, so people who have had needlestick injuries can be somewhat reassured, although its consequences are so severe that immediate prophylactic action is usually advised. At the other end of the scale, measles is

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

highly infective even before the characteristic rash appears, so that routine immunization is the best means of reducing transmission.

- **Pathogenicity**: the proportion of infected persons who develop clinical disease. For instance, young children can develop antibodies to hepatitis A without showing signs of the disease. This happens far less frequently in adults. So hepatitis A is less pathogenic in children than it is in adults. Another example is *Corynebacterium diphtheriae*, which produces diphtheria only when it is carrying the bacteriophage that produces the toxin responsible for causing diphtheria.

- **Virulence**: the proportion of persons with clinical disease who become severely ill or die (severity). The virus that causes the common cold is not virulent; Ebola virus is.

- **Case fatality rate**: of all the persons who contract a disease, the proportion that die of it. The case fatality rate is related to **virulence**. The more virulent the disease, the higher the case-fatality rate is likely to be and the more urgent it becomes to control its spread. For this reason, physicians who diagnose virulent diseases, such as meningococcal illness, should notify the public health authorities as soon as possible by telephone, whereas the notification of less virulent diseases is not so urgent.

- **Herd immunity**: the resistance of a group or community to invasion and spread of an infectious agent. If enough people in the community are immune to an agent, the chain of transmission is very likely to be broken before the agent reaches non-immune people. So, the immunity of the "herd" protects non-immune members. This happens only with agents that are transmitted directly from person to person. If there is a vaccine against the agent, this will contribute to herd immunity and it may be possible to eradicate it.

Nerd's Corner

Why "quarantine"?

The word "quarantine" is derived from the Italian *quaranta giorni* meaning 40 days. In 14th-Century Venice, in an effort to stop the spread of plague, ships arriving in port and travellers arriving by land had to stay outside the city for 40 days. Unfortunately the quarantine did not affect the rats and their fleas that carried the plague and could transmit it from the quarantined travellers to the citizens.

Modes and control of transmission

There are six common modes of transmission of infection. (See Table 11.2) If the mode of transmission is known, precautions can be put in place to prevent outbreaks. Precautions will vary according to the microorganism involved and the context of the case. For instance a case of influenza in a normal household setting does not require strict precautions, where as one in a long term care home might. In hospitals, the infection control team can be a source of advice on which precautions to use. Outside hospitals, the local public health authority can be consulted.

Table 11.2: Modes of transmission of infections and corresponding ways of controlling spread. Note that the precautions to be taken vary with the microorganism and the context of the illness.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Contact	Direct	Direct physical contact (body surface to body surface) between infected individual and susceptible host. Examples: Influenza virus; Infectious mononucleosis; chlamydia. Precautions: Hand hygiene; masks; condoms.
	Indirect	Infectious agent deposited onto an object or surface (**fomite**) and survives long enough to transfer to another person who subsequently touches the object. Examples: RSV; Norwalk; rhinovirus; perhaps influenza. Precautions: Sterilizing instruments; disinfecting surfaces and toys in school.
	Droplet	Via coughing or sneezing, or (in health care) during suctioning. Droplets are relatively large (>5 µm) and can be projected up to about one metre. Examples: Meningococcus; influenza (though there is some debate); respiratory viruses. Precautions: Masks; cover mouth; stand clear.
Non-contact	Airborne	Transmission via aerosols (airborne particles <5µm) that contain organisms in droplet nuclei or in dusts. Can be spread via ventilation systems. Examples: TB; measles; chickenpox; smallpox (and maybe influenza: controversial, as more likely via droplets). Precautions: Masks; negative pressure rooms in hospitals.
	Vehicle	A single contaminated source spreads the infection (or poison). This can be a common source or a point source. Examples a) Point source: Food-borne outbreak from infected batch of food; cases typically cluster around the site (such as a restaurant) b) Common source: The Listeriosis outbreak in Canada in 2008 was linked to a meat production facility in Ontario. It caused 20 cases across five provinces. Cases may be widely dispersed due to transport and distribution of the vehicle. Precautions: Normal safety and disinfection standards. Deliberate contamination of Tylenol in 1982 led to the use of tamper-proof containers for medicines.
	Vector-borne	Transmission by insect or animal vectors. Example: Mosquitoes ? malaria vector, ticks ? Lyme disease vector. Precautions: Protective barriers (window screens, bed nets); insect sprays; culling animals.

Links

Guidelines on hospital infection control

For more information on containing hospital infections, the Public Health Agency of Canada provides guidelines at: http://www.phac-aspc.gc.ca/nois-sinp/guide/ps-sp/index-eng.php

Case Study

Once back in his office, Dr Rao called his local public health unit. He was told that a nurse had already visited Mrs Richards' long-term care home and that the investigation was underway, but there was, as yet, no clue to its cause. Along with the home's management team, the public health nurse was looking at ways of isolating affected patients and reassigning staff members so that staff caring for the ill patients would not care for the unaffected patients. The public health officer thanked Dr Rao for calling.

Problems associated with medical care

Nosocomial infections

Nosocomial infections are those that result from treatment in a hospital or in a health care unit. They spread from staff to patients, from equipment to patients or from patient to patient. In hospitals, they are a common cause of delay in discharge. There are two main reasons for the current problem. One is that the increased use of ambulatory instead of inpatient care means that the patients who are admitted to hospital are sicker and more vulnerable to infection than in the past. The other is that developments in medical technology result in more invasive procedures being carried out using delicate equipment that must be sterilized carefully, but not in the autoclave. In other words, more susceptible people are exposed and new portals of entry are available: two of the risk factors for epidemics. In the case of *Clostridium difficile* infections in hospitals, additional factors are involved, such as the increased use of antibiotics, the reduction of ward cleaning staff, increasing burden of work on nursing staff and, sometimes, the antiquated care facilities including shared toilets and lack of isolation rooms.[9]

To counteract the problem, clinicians should pay strict attention to isolation and infection control procedures, assiduously use sterile techniques and wash their hands before and after patient encounters. They also should ensure that all equipment is properly sterilized and should not reuse

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

equipment intended for single use. Environmental measures, such as proper cleaning and ventilation of clinics and hospitals, good spacing of beds in care units and seats in waiting rooms, separation of sterile and unsterile areas also reduce transmission. Other environmental measures, such as the convenient placement of hand-washing basins and hand sterilizers, as well as of equipment for precautionary measures increase staff compliance with hygiene guidelines.

Antimicrobial resistance

Long before antimicrobials existed, microbes had antimicrobial potential. This became recognised as antimicrobials came into use in the 1920s and 1930s. Until about the 1970s, pharmaceutical science was able to keep ahead by developing new antimicrobials that challenged resistance. The development of antimicrobials effective against resistant strains is now slowing so that antimicrobial resistance has become a worldwide problem.

Several factors encourage the growth of antimicrobial resistance:

- *Inappropriate use*: antimicrobials for self-limiting infections, the latest antimicrobial instead of an effective older one, broad spectrum antimicrobials instead of appropriate narrow spectrum ones.
- *Inadequate use*: too low a dose, too short a course, missed doses, poor choice of route of administration, anything that can expose infecting microbes to an antimicrobial without eradicating them allows the survival of microbes that have some resistance to the antimicrobial. This is a particular problem with infections such as tuberculosis which require a patient to continue therapy for up to a year or more.
- *Use of antimicrobials in veterinary practice as a preventive measure or for promoting growth*: intensive modern farming means that animals are raised in conditions that promote the spread of infection. Antimicrobials are often added to animal feed to prevent the loss of livestock (and profits) to infection and to promote growth (and profits). This continuous, low level dose encourages the emergence of antimicrobial resistant strains which can then infect humans directly or pass resistance on to other microbes which infect humans.

Further Reading

Global antimicrobial resistance

For more about the global perspective on antimicrobial resistance and strategies to counter it, visit http://www.who.int/mediacentre/factsheets/fs194/en/

To reduce the spread of resistance, physicians have to curb their prescribing habits and educate patients on the proper use of antibiotics and the reasons for this. Physicians should make sure that their diagnosis is correct. A patient who presents with a cough and a runny nose might not have an infection: she might have an allergy. Even if she has an infection, as long as the symptoms are mild, it is likely to be viral and self-limiting. In either case, antibiotics are not indicated. In hospitals, where patients are sicker and antimicrobials are used intensively for prolonged periods, attention to infection control practices reduces the spread of resistance.

Nerd's Corner

Survival of the fittest microbe[10]

As with all life, microbes are subject to evolutionary pressures. Survival of the strain depends on its efficiency in spreading from one host to another. Microbes whose transmission depends on close proximity between the host and a susceptible person do not benefit from being highly virulent, for they require the host to be fit enough to move around and transport them to new hosts. A good example is the common cold, which does not incapacitate enough to limit the host's mobility; however, it does irritate the upper respiratory tract, causing the coughing and sneezing that efficiently transmit the agent via droplets. Another example is the female pinworm, who deftly deposits her eggs around a child's anus along with an allergen that causes itching. The child scratches, the eggs get on the fingers, and are then transferred to toys or other people. Microbes that cause sexually transmitted infections are spread by intermittent close contact so benefit from a different strategy. Gonorrhoea, chlamydia, and syphilis can infect silently while their hosts continue an active sex-life, thus ensuring continued spread.

Pathogens that depend on an intermediate vector (such as a flea or mosquito) for transmission can spread even from an immobilised host as long as the vector remains relatively healthy. These pathogens benefit from reproducing rapidly in their host because the greater the number of microbes the host is carrying, the more likely it is that the microbe will be transferred to a passing vector. *Plasmodium* can be present in large numbers in the blood stream during the periodic fevers that characterize malaria and typically confine the patient to bed. The sicker the patient, the less likely he is to worry about that mosquito feeding on his blood and picking up the plasmodium parasite to transmit to the next victim. Similarly, agents able to survive for long periods of time outside of human hosts can afford to be virulent. *Bacillus anthracis* and *Clostridium tetani* are examples of microbes that survive for years in the soil and can kill people rapidly. Again, the ability to reproduce quickly helps the bacteria spread; the patient dies with a high bacterial load and the bacteria return to the soil to patiently await another host.

In general, a disease that spreads via attendants (such as nurses or physicians who carry the pathogen from patient to patient but are not themselves infected) tends towards high virulence. And yet, if the spread is to continue, the microbe can't incapacitate the attendant, so the transmitted dose has to be small. A small dose of a less virulent pathogen might not establish infection. Following this logic, creating barriers to the transmission of certain community acquired microbes may eventually reduce their virulence while providing immediate control of outbreaks. This is because in order to overcome the barriers, the microbes have to infect their hosts for longer periods of time, keeping the hosts in good enough health to transmit the infection. In theory, therefore, it may be possible to encourage organisms to evolve into less virulent forms. It is possible that one of the reasons for the huge impact of the 1918?1919 influenza pandemic was that the virus appeared at a time that favoured the transmission of a virus that would be poorly transmissible under normal social conditions.

Notifiable (or reportable) disease

Physicians are required by law to report cases of certain infectious diseases. Unfortunately, many do not. Reporting disease is necessary in order to direct attention to possible outbreaks or single cases that will require rapid action to control. The notifiable disease database can also be used to evaluate resulting control measures, as well as to trace long-term fluctuations in incidence, indicating changing patterns of behaviour that may

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

require intervention on a number of levels to control the disease.

In Canada, maintaining the notifiable disease information system is a provincial and territorial responsibility. The World Health Organization specifies a number of diseases that must be reported worldwide. The Public Health Agency of Canada specifies some that are notifiable nation-wide and each province or territory can add other diseases to be reported in that province or territory. Not all infectious diseases are notifiable. The Public Health Agency of Canada selects diseases as notifiable according to certain characteristics, such as;[11]

- Its interest to national or international regulations or prevention programmes
- Its national incidence
- Its severity
- Its communicability
- Its potential to cause outbreaks
- The socio-economic costs of its cases
- Its preventability
- The risk it poses in the public perception
- The need for a public health response
- Evidence that its pattern is changing.

Some provincial and territorial public health authorities require physicians to notify them when they suspect an outbreak of any infectious disease. Some lists include non-infectious diseases that can be caused by environmental hazards, such as poisoning with heavy metals or with carbon monoxide.

Physicians are required to notify the public health authority when history and clinical examination causes them to suspect a notifiable disease; they are not required to await the results of testing. Indeed, diseases that could pose an immediate, severe threat to the public's health should be reported by telephone as soon a case is suspected. Laboratories notify the public health authority of cases of notifiable disease when test results are positive. Because some diseases are difficult to diagnose reliably without laboratory testing, while others do not require testing, the list of laboratory-notifiable disease differs slightly from that of physician-notifiable disease.

Self-test questions

Discussion Points

1. Why have antibiotic-resistant strains of bacteria become a problem in recent years?
2. What precautions should a clinician take against the spread of infectious disease in a clinic setting?
3. Why is the common cold not a notifiable disease?
4. What diseases are notifiable in your province Territory?

Self-Testing

1. List the steps of outbreak control.
   - *establish the existence of an outbreak;*
   - *define what constitutes a case and identify cases;*
   - *formulate hypotheses on the causes and implement initial control measures;*
   - *test hypotheses through analysis of surveillance data or special studies;*
   - *draw conclusions, re-adjust hypotheses and control measures if needed;*
   - *plan for long-term prevention and surveillance.*
2. What are the advantages of immunization in the control of infectious disease?
   Vaccination provides long-lasting protection that reduces a person's chances of getting the specific disease. It is a form of passive prevention as it does not rely on continued cooperation for its success. It can break the chain of transmission, thus preventing or arresting outbreaks and, if there is no animal reservoir it can even eradicate the disease.
3. A hospital patient has contracted Clostridium difficile. What precautions should be taken?
   Clostridium difficile is spread by direct and indirect contact. Indirect spread can occur via carers and via surfaces. Adequate hand-washing or hand-cleansing is essential as is thorough cleaning of the room after an infected patient has been in it.

References

1. Porta M. Dictionary of epidemiology. New York: Oxford University Press; 2008.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

2. Holland WW, Detels R, Knox G (editors). The Oxford textbook of public health. Oxford: Oxford University Press; 1991.

3. Kelsey JL, Whittemore AS, Evans AS, Thompson WD. Methods in observational epidemiology. New York: Oxford University Press; 1996.

4. Kilpatrick AM, Kramer LD, Jones MJ, Marra PP, Daszak P. West Nile virus epidemics in North America are driven by shifts in mosquito feeding behavior. Plos Biology 2006 Apr;4(4):e82.

5. O'Connor DR. Report of the Walkerton Inquiry: the events of May 2000 and related issues. 2002; Available at: http://www.attorneygeneral.jus.gov.on.ca/english/about/pubs/walkerton/part1/. Accessed July, 2009.

6. Watson-Creed G, Saunders A, Scott J, Lowe L, Pettipas J, Hatchette TF. Two successive outbreaks of mumps in Nova Scotia among vaccinated adolescents and young adults. CMAJ 2006 Aug 29;175(5):483-488.

7. May JM. The ecology of human disease. Ann N Y Acad Sci 1960 Dec;84(Culture, Society, and Health):789-794.

8. Curtis S. Health and inequality: Geographical perspectives. London, Thousand Oaks, New Delhi: Sage Publications Ltd; 2004.

9. Loo VG, Libman MD, Miller MA, Bourgault AM, Frenette CH, Kelly M, et al. Clostridium difficile: a formidable foe. CMAJ Canadian Medical Association Journal 2004 Jul 6;171(1):47-48.

10. Ewald PW. Guarding against the most dangerous emerging pathogens. Emerging Infectious Diseases 1996 Oct-Dec;2(4):245-257.

11. Doherty JA. Final report and recommendations from the National Notifiable Diseases Working Group. Canadian Communicable Disease Report 2006 1st October;32(19).

"Every system is perfectly designed to achieve precisely the results it achieves."

(Don Berwick, Institute for Healthcare Improvement, Boston)

Chapter Objectives

After reading this chapter, you will be able to:

- Describe the Canadian health services and how they affect the health of a population and the individuals it comprises;
- Know and understand the pertinent history, structure, and operations of the Canadian health care system;
- Describe at a basic level:
  - Methods of regulation of the health professions and health care institutions;
  - Supply, distribution, and projections of health human resources;
  - Health resource allocation;
  - Organization of the Public Health system; and
- Describe the main functions of public health related to population health assessment, health surveillance, disease and injury prevention, health promotion, and health protection.

Case Study

After the outbreak of diarrhoea in the long stay home, Julie and Paul start to worry about Paul's mother, Catherine Richards. They know she would rather be at home. They think they could organize their home to take her in, but Dr. Rao wonders if this is feasible. Mrs Richards needs continuous supervision, at least two people to help her in the bath, and she is occasionally incontinent. She also needs physiotherapy to prevent contractures and bronchopneumonia. Still, Julie and Paul think they could learn to give Mrs. Richards the care she needs. Dr. Rao knows that more chronic illness care is provided by family members than by health care services, but he has seen many carers suffer from the stress of being constantly "on call." Sometimes family caregivers find it difficult to get the support they need from health care services. Even getting short term respite care is often difficult.

Dr. Rao contacts the geriatric team in the Weenigo hospital for advice on medical management. Their liaison worker suggests that the Richards contact the local social worker to discuss the home services available. One of Paul and Julie's main problems will be to know what exactly is covered by Medicare and if they are eligible for financial help. Their next problem will be finding out where to get services.

Elements of a health care system

Health care systems are complex organisations comprising regulatory, funding, and service provision bodies that provide **access** to health care in accordance with societal goals and values. The metaphor of a house (see Figure 12.1) can be useful in describing health care systems. The roof corresponds to the societal goals and values that shelter service provision, which is founded on legislation and regulations that control the relationships among providers (the rooms of the house), funding agencies (the power source) and citizens. Regulations also control who can provide care (back door) and who can access it (main door). Note that the model can be applied to the country as a whole or to smaller regions. It can also be applied to specific programmes, such as cancer or HIV/AIDS care.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

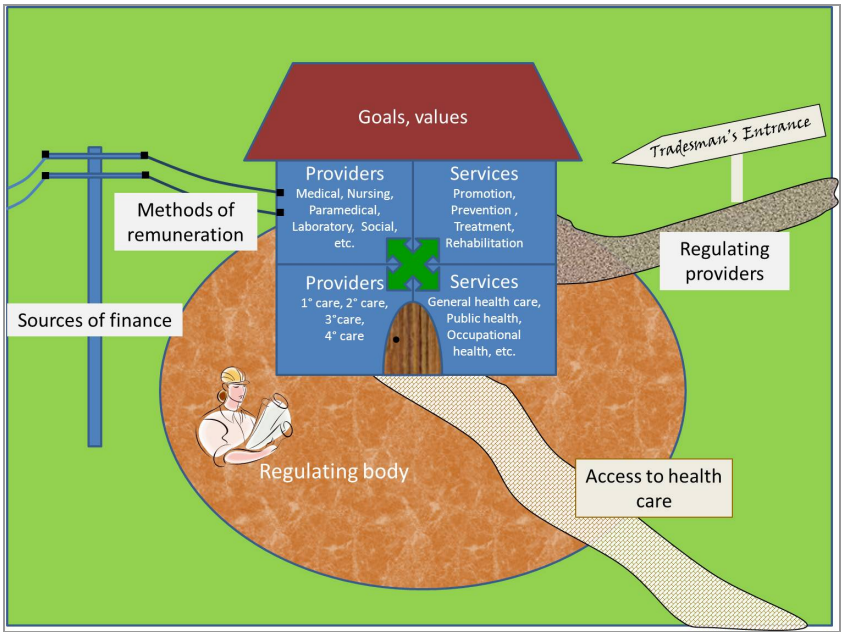


Figure 12.1: Description of a health care system. (Developed from an OECD model[1])

Societal goals and values

In common with most of the world's developed countries, Canada treats health care as a resource for all, regardless of ability to pay. It is not seen as a commodity that can be bought and sold in the open market. Over the past one-and-a-half centuries, several factors have led Canadians to expect that the government become increasingly involved in the organization of health services. These factors include

- Urbanisation, which broke the informal social networks that provided health and social care for those in need
- The Great Depression, during which governments became involved in relief efforts, and after which people never again wanted to have to beg for health or welfare services
- The two World Wars, which left people disabled and widowed for their country, leading people to ask what their country should do in return.

Meanwhile, increasing knowledge in medicine led to more effective **health care technology**. Governments, noting that healthy people are more economically productive than unhealthy people, began to see the advantage of providing health care services. The Nerd's corner box titled "Timeline: Canadian Health Services" shows the gradual evolution of legislation in response to societal pressures.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

Timeline - Canadian Health Services

Table 12.1: Milestones in the development of the Canadian health care system.

1867	*The Constitution Act* Although health care was not specifically mentioned in this Act, it gave provincial legislature power for the "Establishment, Maintenance and Management of Hospitals, Asylums, Charities, and Eleemosynary [alms giving] Institutions." This power was mainly regulatory. Powers of "Quarantine, and Establishment and Maintenance of Marine Hospitals" were retained by the Federation.
1914	Sarnia, Saskatchewan . This rural municipality successfully experimented by offering physicians a retainer to practice in the area. The plan guaranteed physicians an income while allowing them to charge for their services.
1916	Saskatchewan. The province passed the *Rural Municipality Act* that permitted rural municipalities to levy property taxes to pay for the retention of physicians.
1917	Saskatchewan. Provincial legislation allowed municipalities the right to collect taxes to finance hospital care.
1920	The creation of a national federal Department of Health in response to the influenza pandemic of 1918?1919 and to address the welfare of returning soldiers. This new department was given responsibilities for implementing campaigns against venereal diseases, tuberculosis, and for promoting child welfare. It took over responsibility for quarantine and ensuring food and drug standards from the Department of Agriculture.
1934	Newfoundland. The Cottage Hospital and Medical Care Plan provided care in remote communities.
1935	The provinces successfully challenged the federal government's plan to provide certain social and health benefits financed by collecting taxes. The British Privy Council ruled that health care lay outside the federal government's responsibility.
1947	Saskatchewan introduced public insurance for hospital services. This followed the federal government's attempt to do so, which failed because federal and provincial governments couldn't agree on financial arrangements.
1957	*Federal Hospital Insurance and Diagnostic Services Act* (HIDS), passed in 1957 and implemented in 1958, offered the provinces an average of 50% of the funding of hospital services if certain conditions were met, such as the provinces offering universal coverage to their residents, ensuring adequate standards covered by the act, and keeping adequate records and accounts.[2,3]
1961	All 10 provinces participate in HIDS.
1966	The *Medical Care Act* provided for universal coverage for physicians' services.
1971	All provinces had programmes that complied with the *Medical Care Act*.
1984	The *Canada Health Act*.

The Canada Health Act

In Canada, as set out in the Constitution Act of 1867, the provision of health care falls under provincial jurisdiction. In theory, this could mean that each province and could administer health care in a completely different way (and in some ways they do differ). The Canada Health Act, introduced in 1984, augmented the Constitution Act. The Canada Health Act aims "to protect, promote and restore the physical and mental well-being of residents of Canada and to facilitate reasonable access to health services without financial or other barriers." The Act applies to all services deemed medically necessary for the purpose of maintaining health, preventing disease, or diagnosing or treating injury, illness or disability, and includes accommodations and meals, physician and nursing services, drugs, and all medical and surgical equipment and supplies.

The Canada Health Act specified that the federal government would provide funding to the provinces and territories on condition that they complied with the principles of the Act. In this way the federal government could ensure some measure of continuity across provincial health systems. The federal government originally provided half of the funding for health services, but this portion has fallen steadily over the years, eroding the federal government's ability to influence provincial governments' policies. The principles of the Act are:

- **Public Administration**: The provincial or territorial health insurance plan must be administered and operated on a non-profit basis by a public authority accountable to the provincial or territorial government.
- **Comprehensiveness**: The plan must insure all medically necessary services provided by hospitals, dentists working within a hospital setting, and medical practitioners.
- **Universality**: The plan must entitle all insured persons to health insurance coverage on uniform terms and conditions.
- **Accessibility**: The plan must provide all insured persons reasonable access to medically necessary hospital and physician services without financial or other barriers.
- **Portability**: The plan must cover emergency services for all insured persons when they are visiting another province or territory within Canada. When they move to another province or territory, all insured persons should be able to transfer their insurance to that province or territory.

The Act specifically prohibits **extra-billing** and **user fees** for insured services, which are

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

- Hospital services that are medically necessary for the purpose of maintaining health, preventing disease, or diagnosing or treating injury, illness or disability, and include accommodations and meals, physician and nursing services, drugs, and all medical and surgical equipment and supplies. However, the Act does not define which services or drugs are "medically necessary"
- Any medically required services rendered by medical practitioners
- Any medical or required surgical-dental procedures that can only be properly carried out in a hospital.

The Canada Health Act is still the basis of our health care system, although it only covers hospital services including professional services within hospitals and physician services outside hospitals. There have been changes in the way funds are allocated to provinces and territories, which allow the federal government to set the direction of health care, but the principles still hold. As a substantial source of finance, the federal government can influence provincial and territorial health care systems.

The regulating body

Health care in Canada is a shared responsibility. In reality, Canada has not one, but 14 health care systems?a different system in each province and territory. The federal government sets standards and principles and assists in financing provincial and territorial health care services. The provinces and territories are responsible for the administration and delivery of services for most of the population. The federal government retains responsibility for providing services to First Nations, Inuit and Métis communities, members of the RCMP, members and veterans of the Canadian Forces, prisoners in federal penitentiaries, and refugee claimants. The federal government also has a role in coordinating the promotion and protection of the public's health; it contributes to disease surveillance and prevention, it supports health promotion through the Public Health Agency of Canada, and it regulates drugs, medical devices, food, and consumer safety through Health Canada.

Provincial governments work within the parameters of the Canada Health Act to provide health care services according to the needs of the population. They plan, fund, and evaluate hospital care, physician care, allied health care, prescription drug care in hospitals, and public health, as well as negotiate fees with health professionals. Most provinces discharge their provincial obligations through regional boards. This decentralizes decision-making and enhances responsiveness to community needs. These boards oversee publicly provided services, including hospitals, nursing homes, home care, and public health services.

In addition to providing hospital care under the Canada Health Act, most provinces and territories have special plans for low-income residents and seniors, such as out-of-hospital drug benefits, ambulance costs, and hearing, vision and dental care. Some provinces and territories fund community health service clinics, which provide a range of professional services in the community. Some fund extramural programmes, which provide various types of care in patients' homes, particularly palliative care, post-operative care, home oxygen, long term care assessment, rehabilitation, etc. Figure 12.2 shows trends in the proportion of total spending (public and private) in the different service sectors.

Regulating the providers

Governments ensure that care providers meet certain standards in a number of ways, including

- Setting the standards for publicly funded institutions that provide care. Some require regular accreditation assessments. Within the institutions, standards generally require professionals to audit their practice regularly.
- Regulating health professionals. The practice of a regulated professional is: 1) covered by provincial or federal legislation and 2) governed by a professional corporation or regulatory authority, for instance a College of Physicians or an Order of Nurses. Given that many of these regulatory bodies are provincial, variation exists between provinces and territories (for examples of regulated professions, see Nerd's Corner box). There are practitioners who define themselves as formal providers of health care, but who are not members of any professional corporation and therefore they may not have had to prove their fitness to practice and may not respect a code of ethics.

Here Be Dragons

Role of professional corporations.

It is important to realize that the provincial regulatory authorities that govern professional practice differ from the groups that represent professional interests. For instance, for physicians in Ontario, the College of Physicians and Surgeons of Ontario is the provincially-recognized regulatory body, whose duty is to serve and protect the public by regulating Ontario physicians and surgeons, whereas the Ontario Medical Association represents the interests of Ontario physicians and exercises no regulatory function.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

Some regulated health professions

Some professionals are regulated only in certain provinces. For instance, British Columbia is one of the few provinces that regulate traditional Chinese medicine and acupuncture; it also regulates massage therapy, which is not regulated in Ontario. Regulated midwifery is spreading across Canada although, at the time of writing, some of the maritime provinces have yet to enact or implement legislation relating to it.

Table 12.2: Examples of regulated health professionals.

Professionals	Area of expertise or practice
Audiologists and speech-language pathologists	Hearing and understanding, speech, language, and swallowing disorders
Chiropodist/podiatrist	Assessment of the foot; treatment and prevention of its diseases
Chiropractors	Diagnosis, treatment, and prevention of mechanical disorders of the musculoskeletal system
Dental hygienists	Preventive oral hygiene
Dental technologists	Design, construction, repair, or alteration of dental prosthetic, restorative, and orthodontic devices
Dentists	Evaluation, diagnosis, prevention and treatment of disease of teeth, the jaw, the mouth, the maxillofacial area, and the adjacent and associated structures
Denturists	Oral procedures and related activities pertaining to the design, construction, repair, or alteration of removable dentures for the partially or fully edentulous patient
Dieticians	Evaluation of the nutritional state of people in order to choose and implement a nutrition strategy that takes account of their need to improve or re-establish health.
Kinesiologists	Assessment of human movement and performance, and its restoration and management to maintain, rehabilitate, or enhance movement and performance
Massage therapists	Assessment of the soft tissue and joints of the body, and the treatment and prevention of physical dysfunction and pain of soft tissue and joints by manipulation to develop, maintain, rehabilitate, or augment physical function, or relieve pain
Medical laboratory technologists	Medical tests on blood, body fluids, and tissues
Medical radiation technologists	Application of radiation therapy, radiography, nuclear medicine, magnetic resonance imaging
Midwives	Assessment and monitoring of women during pregnancy, labour, and the post-partum period, as well as of their newborn babies; the provision of care during normal pregnancy, labour and post-partum period, and the conducting of spontaneous normal vaginal deliveries
Nurses	Care of individuals of all ages, families, groups, and communities, sick or well, and in all settings
Occupational therapists	Helping people to learn or re-learn to manage the everyday activities that are important to them, including caring for themselves or others, caring for their home, and participating in paid and unpaid work and leisure activities
Opticians	Supply, prepare, and dispense optical appliances, interpret prescriptions prepared by ophthalmologists and optometrists, and fit, adjust, and adapt optical appliances
Optometrists	Assess the eye and visual system, sensory and ocular motor disorders and dysfunctions of the eye and the visual system, and diagnose refractive disorders
Pharmacists	Evaluate and dispense prescription medications; advise on their correct use and mode of action
Physicians and surgeons	Assessment and diagnosis of health problems, prevention and treatment of disease in order to maintain or restore health
Physiotherapists	Through understanding of how the body moves and what prevents it from moving, manage and prevent many physical problems caused by illness, disease, sport- and work-related injury, aging, and long periods of inactivity
Psychologists	Assessment, treatment, and prevention of behavioural and mental conditions
Respiratory therapists	Monitor, evaluate, and treat individuals with respiratory and cardio-respiratory disorders
Social workers	Help individuals, families, groups, and communities to enhance their individual and collective well-being; help people develop their skills and their ability to use their own resources and those of the community to resolve problems. Social work is concerned with individual and personal problems, as well as with broader social issues such as poverty, unemployment, and domestic violence.

Sources of finance

There are two basic funding models for publicly financed health care systems. The first, the social insurance model, uses compulsory contributions

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

to a social insurance fund. Governments can direct how the premiums are levied and what amount should be levied, so that premiums can be linked to a person's income. In some countries, citizens can choose from a number of insurance providers; in others, the choice is limited to a single national not-for-profit insurer.

In the second model, as used in Canada, general taxes fund health care. However, in Canada, only hospital and physician services are universally funded. Other services are funded through a variety of sources, such as social insurance (often used for drug insurance), social security, and out-of-pocket and private insurance. Meanwhile, provincial workers' compensation and health and safety at work programmes are funded by a form of social insurance in which employers pay premiums that are graded according to the inherent risk of the industry and the past performance of the employer. Figure 12.2 shows the sources of the health care finance in Canada and how that money is spent.

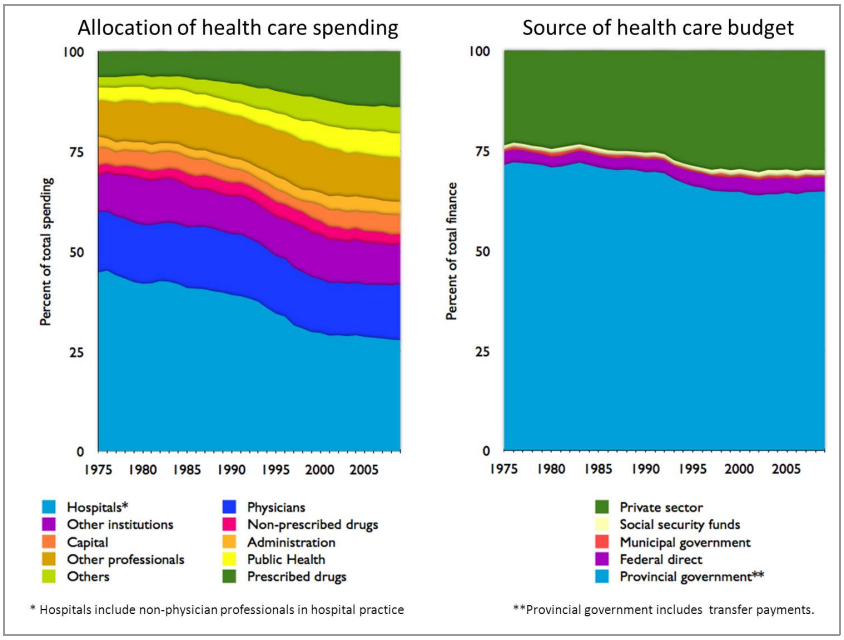


Figure 12.2: Trends in sources and allocation of health care budgets in Canada, 1975 to 2009. Source of information: CIHI.[4] Note that both graphs represent the proportions of budgets, not the absolute budget, which is increasing

Public versus private financing of health care

Currently, a major policy discussion in health care financing concerns whether or not to allow people to pay out of pocket or to buy private insurance for services that are provided by public funds, for example to avoid waiting a long time for an operation. Proponents of private financing say that it would take the pressure off the public system. However, human resources link the two systems. The number of qualified professionals is limited, therefore, when professionals move into the private system, the public system is depleted. Moreover, private insurers generally avoid insuring people with, or at risk of, serious and chronic conditions requiring complex and expensive care, and very few people could afford such care without private insurance. As a result, public funds would continue to be used for the most expensive care, so reduction in cost to the public system would be small, if any. Finally, private insurers tend to have large administration costs. For example, in Canada where health care coverage is mainly public, administration costs are 17% of health spending, whereas the administration costs of health care in the U.S., where health care coverage is mainly private, are 30% of spending.[5]

Illustrative Materials

The Chaoulli affair

Canada is unique among **OECD** countries in prohibiting private insurance for services that are covered by the provincial systems. However, in Québec, Jacques Chaoulli, the physician of a patient waiting for a hip replacement, challenged this. The case went before the supreme court of Canada in 2004. In 2005, by a majority decision, the court ruled that the sections of the *Health Insurance Act* and of the *Hospital Insurance Act* that prohibit private insurance violate the *Quebec Charter of Rights and Freedoms*. No ruling was made on whether or not the Acts in question violate the *Canadian Charter of Rights and Freedoms*.

To many, this ruling represents a threat to the Canadian single insurer system, possibly opening the door to private insurance for services covered by public health insurance, which could eventually drain human resources away from publicly insured care, thereby reducing the level of care available for those who cannot afford private insurance. Furthermore, as single insurer systems put the single insurer in a strong bargaining position to negotiate prices downwards, the resulting multi-payer system would probably allow health care costs to rise.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Further Reading

Private and public finance

For more information on the topic of private versus public payment for health care, see "Myth: A parallel private system would reduce waiting times in the public system," March 2005, at the Canadian Health Services Research Foundation website: http://www.chsrf.ca/PublicationsAndResources/Mythbusters/ArticleView/05-03-01/5bda3483-f97b-4616-bfe7-d55d0d66b9a0.aspx

Further Reading

The 70/30 split

For more about the Canadian health care system and its financing, read the Canadian Institute for Health Information's *Exploring the 70/30 split: how Canada's health care system is financed,* available at: http://www.cihi.ca/cihiweb/dispPage.jsp?cw_page=AR_1282_E

Controlling Costs

One of the major problems facing health care administrators is how to control costs which, in developed countries, have been growing constantly in the last century. Figure 12.3 shows the rising cost per capita in Canada. Population aging has contributed somewhat to this increase, but the cost increase is mainly due to increased access to health care, technological advances, and limited incentives to control costs. Note that Figure 12.2 showed the *proportion* of the budget allocated to various expenditures; even though the total spent on hospitals is rising, the amount spent on prescription drugs is increasing more, so that the proportion of the budget spent on hospitals is declining.

Health care utilization can change as a result of supply and demand. Both have increased faster in rich countries than in poor for a number of reasons. In terms of supply, governments in wealthy countries have higher incomes with which to pay for services, and technical innovations create a wider range of services. In terms of demand, people are more able to pay for health services. Moreover, health messages create a more informed population and encourage them to consult health professionals. In particular, elderly Canadians are making more use of family physician services.[6,7] The net effect is that health care costs are rising faster than the national wealth as measured by the gross domestic product (see Figure 12.3).

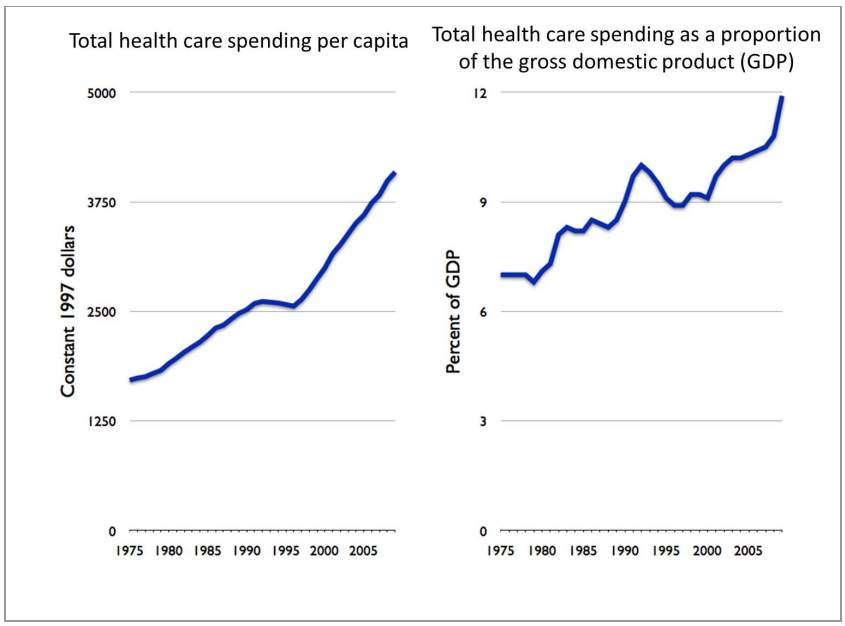


Figure 12.3: Total health care expenditure per capita and as a proportion of the gross domestic product (GDP), Canada, 1975?2009
Source: CIHI[4]

Developments in technology and pharmaceuticals produce new treatments, which are generally more expensive than the ones they replace. Being generally more effective that the old treatments, the new treatments may be used more frequently, so that, even if a new treatment is cheaper, it may end up costing the health service more. Costs can also increase when a new technology, developed for a specific application, is used more broadly. For instance, the CT scan was originally developed for imaging only the brain, but has since become indispensable for imaging many other parts of the body. In the same way, drugs that are approved for limited indications are sometimes marketed for a much broader range of indications.

Finally, health care is a large sector of the economy. Health and social services in Canada employ about 12% of the working population.[8] This means that about 12% of the working population have an interest in maintaining current levels of spending on health care in order to maintain their income. In addition, companies that research, develop and produce technologies and pharmaceuticals provide employment, generate income for shareholders, and often produce exportable goods. Because of this, these companies have a degree of political power as well as a vested interest in health care. In this context, political will tends to maintain or increase health care spending instead of decreasing it.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

Canadian health care spending[9]

In 2007, $160 billion was spent on health care: nearly $4,900 per person. Of this, just over $3,000 was public spending and just over $1,300, or 30%, was private. The public to private spending ratio has been fairly steady at around 70:30 since the mid 1990s.

Table 12.3: Breakdown of health care spending in Canada, 2007

2007 Canadian health care spending	
Public and private sector	% of total spending
Institutional services (hospitals and publicly funded residential care)	39
Professional services (Physicians, dentists and others not on the payroll of institutions)	24
Drugs (prescribed and over the counter)	17
Public health	5
Other	15

? Of these categories, institutional spending has been falling in comparison to drug costs, which have risen dramatically

? About 44% of government health care spending in 2005 was on people aged 65 and older; this proportion has changed little since 1998. People aged 65 or over formed 14% of the Canadian population in 2010

? About 5% of Canadians work in health related occupations. In 2005, there were just over 62,300 physicians in Canada. In 2006, there were nearly 253,000 registered nurses, 67,300 licensed practical nurses, and 5,051 registered psychiatric nurses. The physicians were split almost equally between general or family physicians and specialists. However, although 20% of Canadians lived in rural areas, only 15.7% of general practitioners and family physicians do so

? In 2007, 85% of Canadians reported having a regular medical doctor. Of those without a regular medical doctor, 6% said they had looked but couldn't find one

? Per capita spending on health in Canada is similar to that in France, Germany, the Netherlands, and Australia. The United States spent nearly twice as much per capita as Canada. New Zealand and Japan spend just two thirds as much per capita as does Canada.

Recent trends include:

- Hospitalization rates for conditions which can usually be treated outside hospital are falling
- Canadians are becoming less likely to die in hospital
- Re-admission rates after heart attack are falling
- In 2005, 27% of nurses delivering direct care thought the quality of patient care had deteriorated over the previous year; 16% thought it had improved.

Economists argue that there are only two ways to reduce health care costs: either reduce the use of services (i.e., reduce demand) or improve their efficiency (which would result in increasing supply for the same cost), either by improving administrative support or by improving quality. They also note that, unlike the case in most other markets where demand leads supply, in the health care market increasing supply tends to increase demand, so that improved efficiency is unlikely to be sufficient to control costs. Table 12.4 outlines alternative approaches to increasing efficiency in health services, showing possible side-effects of each approach.

Table 12.4: Approaches to reducing health care costs[10,11]

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Reduce use of services

Approaches	Examples	Issues raised
Copayments or user fees	Financial incentives for patients to reduce their use	As poverty is a major determinant of health, co-payments or user fees ensure that those most in need have least access to care. If patients delay consultation in the hope of avoiding user fees, this could increase subsequent costs. Moreover, user fees tend not to reduce use of services or health care costs overall.
Limit resources	Day surgery and ambulatory treatment, both popular with patients, only became common with the closure of hospital beds	Forces efficiencies and innovative methods. Politically difficult. Can increase pressures on other parts of the system or create other unintended consequences. Closure of acute hospital beds increases pressure on informal carers, home care, and long stay care; more intensive use of remaining beds stresses staff and increases cost per bed, thus savings are rarely as great as predicted. Wait-lists for non-urgent procedures may expand, perhaps requiring more expensive care if the condition deteriorates.
Use gate-keepers	Access to secondary care only through primary care.	Generally reduces waste by ensuring that the correct secondary provider is consulted and by allowing the primary care provider coordination and case management role.

Improve efficiency

Approaches	Examples	Issues raised
Use professionals' skills appropriately	For instance, use nurse practitioners for prevention and routine follow-up, physicians for complex diagnostic and treatment problems.	Some claim that non-physician care is seen as second rate. However, most people are satisfied with it when it is appropriate. When the tasks of each person are clear and accepted, professional job-satisfaction is likely to be improved.
Improve practice	Educate and support patients in self-management of chronic conditions.	Can reduce hospital admissions and emergency room visits. Current thinking holds that improvement in patient autonomy in itself improves health.
	Improve management of chronic conditions.	Timely and appropriate interventions to control the condition and reduce its impact on function, thus reducing need for services
	Use of evidence-based guidelines for optimal management.	Observations on wide variations in practice with little apparent variation in need suggest that over-treatment could be reduced by use of guidelines. However, guidelines also vary and are hard to impose. Patients' wishes must also be respected.
	Reduce medical error.	Medical error is an important preventable cause of morbidity, which costs the system a great deal of time, money, and resources. Systems approaches to medical error reduction and quality assurance are effective in improving quality of care, morbidity associated with error, and reducing costs.

Improve support systems

Approaches	Examples	Issues raised
Information systems	Electronic medical records, portable databases, digital imagery available via a secure network, provision of information on drug costs, etc.	Contribute to reducing duplication of tests, overmedication and drug interactions, and coordination of care, application of guidelines and generation of data for evaluation of practice. Systems have to be designed with the users and the objectives in mind. In certain situations a cheap paper system may be superior to an expensive electronic one.
Administrative systems	Call-recall systems for preventive and follow up care.	Improve the uptake of care but there is some concern that recall systems may generate unnecessary interventions. Running costs should not outweigh advantages of the system.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Financial systems	Make sure that remuneration systems rewards high quality, efficient care.	All remuneration systems have advantages and disadvantages (See table 12.6)

Illustrative Materials

Manitoba tackles overuse of health care

Manitoba responded to overuse of its health care system by forcing people who visit too many doctors or pharmacies to limit their use to one doctor and one drugstore. The province reviewed the records of patients who made more than 67 office visits in a year or saw more than 12 physicians annually. The heaviest user made 247 office visits to 71 different physicians in one year. Of 99 people identified as heavy users of the system, 34 had medical conditions that justified the use. In only 28 cases were restrictions on service use imposed. The project was expected to save the province $116,000 in a province that spends nearly $2 billion a year on health care (around 0.005%).[12]

Health care service provision

There are two basic ways of delivering publicly funded care. One way is that the government itself may organize delivery of the services. This happens in the UK, Cuba, and some Scandinavian countries where health care workers, including physicians, are public employees. Hospital and clinic buildings are owned by the state and services are managed publicly.

The other way, for instance in Canada, is that service providers can be public or private, although finance for medically necessary hospital and physicians services comes from a public insurance system. Private service providers include for-profit, not-for-profit, charitable, and religious organizations?in fact, all providers except publicly administered ones. The private sector can offer hospital, long term care and community services that are funded by provinces. Most physicians, working in hospitals or not, contract with the provincial insurance plan to provide services as private service providers. For patients with private insurance or who can afford to pay out-of-pocket, and who want services not covered by the provincial insurance plan, there is a wide range of professionals working outside of hospitals providing services such as physiotherapy, occupational therapy, optician, podiatry, and psychology.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Here Be Dragons

Public or private?

*Table 12.5: The Canadian public-private mix[13,14]

		Delivery		
		Public	Private not-for-profit	Private for-profit
Finance	Public	▫ Public Health ▫ Provincial psychiatric institutions ▫ Home Care in some provinces	▫ Most hospitals ▫ Addiction treatment	▫ Most physicians ▫ Ancillary services in hospitals (laundry services, meal preparation and maintenance) ▫ Laboratories and diagnostic services in most provinces ▫ Some hospitals
	Private	▫ Enhanced non-medical (e.g., private room) and medical (e.g., fibreglass cast) goods and services in a publicly owned hospital	▫ Some home care and nursing homes in some provinces	▫ Cosmetic surgery ▫ Long-term care ▫ Extended health care benefits such as prescription drugs, dental care and eye care in some provinces ▫ Some MRI and CT scan clinics ▫ Some surgery clinics

*NB: This table, based on work carried out in 2002, still offers a generally correct reflection of the situation. However, gradual change in federal, provincial and territorial policies is causing some movement between the cells. The different provincial and territorial systems also change at different rates.

In debates on health care it is important to acknowledge the difference between funding services and providing services. For instance, Dr Rao now works in a group practice in which the physicians bill the insurance plan for the services they provide and the practice members administer these funds as they see fit, paying for the clinic buildings, employing the staff, paying their own salaries etc. However, the insurance plan is publicly funded, so that people who use Dr Rao's services do not have to pay for them. Similarly, institutions that provide publicly insured services can be private, but patients do not have to pay for services.

In Canada, the public or private health care debate is generally about who pays for the care: the provincial insurance plan or whether the patient should be allowed pay directly or through private insurance.

The other hot topic is whether or not hospitals should be private for-profit. Currently, almost all Canadian hospitals are private not-for-profit, operated by regional health authorities.[15] As such they are perceived as public institutions, but technically they are not. There have been studies that show an increased mortality in for-profit hospital and fears that the ?bottom-line' will harm patient care.[16] There are also fears that for-profit hospitals will open the door to the free market and erode the principles of the Canadian health care.[15]

Illustrative Materials

Quebec's population health approach to providing care

In 2003, Quebec adopted a bill which transformed its community services. Its Local Community Service Centres (CLSC) were merged with its residential and long-term care centres (CHSLD) to become Health and Social Service Centres (CSSS). As well as providing certain services to individual clients, the CSSS are charged with ensuring that the needs of the population of their territories are met. A CSSS must

- Monitor the health status of its population and coordinate action to improve it
- Manage and coordinate the general and specialized services available to its population, taking the appropriate measures to provide case management, assistance, and support to users of the health and social service network
- Ensure that services are effective, efficient, relevant, and meet users' expectations and the population's needs
- Inform, consult, and receive input from its population, as well as assess satisfaction with regard to the services and their outcomes.

All CSSS provide a range of front line services, which may include prevention, primary care treatment, rehabilitation, crisis support, and public residential care. To ensure provision of other health and social services and secondary and tertiary services that do not exist in the CSSS territory, the CSSS must negotiate agreements with the providers of these services. Health policy making is more fully described in Chapter 14.

Remuneration of providers

There are a number of basic ways of paying health care providers for their services. Back in the 1960s, Avedis Donabedian proposed a useful way of thinking about the quality of care that can be extended to provide a conceptual basis for designing payment systems.[17,18] Costs of care (and hence payments) can be based on the *structure* of an institution (number of hospital beds, staff numbers, etc.), or payments can be based on the services provided: the *process* of care, in Donabedian's term. Alternatively, quality can be judged, and payments provided, according to the

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

outcomes of care (success rates). In Canada, most physicians are paid by a fee per item of service, i.e. a process of care payment, while block funding is used for most institutions, based on a combination of structure and process (see Table 12.6).

Table 12.6: Approaches to paying health care providers

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Remuneration of physicians and other professionals

Method of payment	Discussion points
Fee for service	The professional, acting as an independent, private contractor, is paid a set amount for each service provided: a process of care system. Although methods of remuneration of physicians in Canada are changing, most are still remunerated by this method, as are many other professionals in private community practice.
	This encourages professionals to provide services that are less time-consuming and that attract higher fees.
	Fee-for-service payment is often blamed for encouraging the development of walk-in medical centres at the expense of services that manage chronic illness and complex cases. However, walk-in services respond to patient demand for easily accessible, rapidly responsive services.
	Fee-for-service payment tends to encourage provider-driven overuse of services, increasing the risk of iatrogenic illness and medical error.
	Fee-for-service remunerates a single professional for a single service, so provides no incentive for the development of teamwork.
Salary	In Canada, most non-physician health services in institutions are provided by salaried professionals. A salary implies that the professional is an employee of an organisation and is responsible to managers for services provided. In Canada, salaried physicians, although providing services within institutions, generally receive their salary from the provincial insurer, not directly from the institution in which they work.
	In the case of physicians, salaries are said to remove the incentive to offer more than the minimum of services. There is also a fear that being responsible to a manager hinders professional autonomy in deciding what care is appropriate.
Retainers	Can be useful to attract physicians into areas of low patient volume (such as remote areas). The retainer ensures a minimum salary and this can be coupled with fee per service to encourage service provision.
	In Canada, retainers are frequently used in specialties such as anaesthetics and psychiatry. In Quebec, many specialists have opted for retainers coupled with modified fees per service.
Capitation	Payment according to the number of people on a patient list. The fee structure can include a premium for complex cases and may be adjusted for the sociodemographic profile of the patient population. The practice is paid whether or not the patient consults. The practice team may include a number of disciplines and the payment remains the same no matter which member of the practice team provides services.
	Capitation probably discourages the provision of unnecessary services and encourages the provision of effective preventive services.
	It can encourage the development of multidisciplinary team work.
	There is concern that the requirement to register with a physician can reduce patient and physician freedom of choice. Most capitation plans allow for patients to change physician and for physicians to refuse patients.
	Family practice in the UK is generally paid by capitation. In Canada, capitation is used in some primary care practices in some provinces.
Target payments	Payment for reaching a target level of services delivered. Useful in preventive services. For instance, a physician could receive a payment according to the proportion of patients on his list who have had cervical cancer screening. This moves towards a payment system based on the outcomes of care.
	Requires a denominator to develop accurate targets?generally a list of patients who have chosen to be cared for by the physician under a capitation scheme.
	There are concerns that patients could be coerced into undergoing unnecessary or unwanted interventions in order to achieve the target.
Mixed	Mixed remuneration schemes are very common, for instance, supplementing a capitation programme with targeted payments. They seek to get the best from each of several types of payment.

Payment methods for institutions

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Method of payment	Discussion points
Block funding	This is commonly used in Canadian hospitals. The institution is paid a certain amount of money every year to provide services. The amount paid is generally calculated according to the previous year's amount and the types of services offered, adjusted for demographic change, changing health care costs, and inflation. It is a structure based payment system. Block funding, especially when coupled with sanctions for going over budget, favours cost control. However, it can also merely shift the costs to another budget. For instance, reducing the length of hospital stay can reduce hospital costs (as long as the number of hospital stays don't increase), but the cost of home services, and the costs to families, may increase as sicker people need home care.
Funding by episode of care	This is similar to fee-per-item of service funding, and represents a process of care payment system. In general, the fees are scheduled according to the patient's diagnosis classified in a way that reflects the average cost of care required for that diagnosis. This method encourages efficiency for single episodes; it does not provide any incentive to control the number of episodes.

The amount of money most Canadian hospitals receive is calculated according to their inputs: the number of beds they have, the number of staff they employ, or simply according to the previous year's budget with allowances for changes in clientele, services offered, and inflation. Hospital managers responsible for the budget are expected not to overspend, so they must limit the number of services provided. In tension with this, physicians, mainly on fee-for-service, have an interest in increasing the number of services they provide. Physicians increase hospital costs when they use hospital facilities to provide services (for instance, when a surgeon operates). But the physician has no responsibility for hospital costs and, as long as he can use hospital facilities, his income is guaranteed. The Kirby report suggests that hospitals should instead be paid according to the services they provide.[19] This would reduce the conflict between the interests of service managers and those of physicians. However, without other controls on spending, this plan could cause costs to rise. There is no perfect method of remuneration for service. Research on the effects of remuneration methods on quality of care and health outcomes is lacking, and definitive conclusions about which method of payment might be best are not available.

Further Reading

Remuneration of providers

To read a perspective on the effects of physician remuneration systems, see "Myth: Doctors do it for money," September 2003, at the Canadian Health Services Research Foundation website:

http://www.chsrf.ca/PublicationsAndResources/Mythbusters/ArticleView/03-09-01/84c4445b-89da-4a46-8082-60d7ae969ee4.aspx

To read about changes in the attitudes of physicians regarding fee-for-service payments, see "Myth: Most physicians prefer fee-for-service payments," January 2010, at the Canadian Health Services Research Foundation website: http://www.chsrf.ca/PublicationsAndResources/Mythbusters/ArticleView/10-01-01/13b5e8bb-e7c2-4544-8da5-b1aa5d9e38db.aspx

For more about options for funding hospitals go to: http://www.chsrf.ca/PublicationsAndResources/researchreports/articleview/11-01-19/f20713d8-905d-43a5-bc79-1c457f95e495.aspx

Illustrative Materials

A radical experiment in remuneration for services

One of the challenges for limiting health care expenditure is the separation between those who are responsible for constraining costs and those who make the decisions that incur the costs. Physicians, with their patients, decide on a management plan, while hospital or laboratory managers are likely to be responsible for controlling the costs that result from the management plan. Some experiments have tried to draw responsibilities for budgets and for patient management closer together. In the UK, General Practice fund-holding gave family physicians the budget with which to pay for the health services their patients needed, including the family physician services, diagnostic and hospital services. Within their practice physicians were free to allocate budget for services they thought necessary, including office staff and the services of other professionals, such as practice nurses and social workers. Their budgets were based on the number of patients on their practice list and were adjusted for age of the patients and the proportion of patients on their list with certain chronic diseases. Any savings that were made could go into practice development. There were indications that fund-holders managed to generate economies while reducing patient waiting times. However, the project was highly politically charged and was stopped without proper assessment when a new government came to power.[20] Some physicians believed it was unethical for physicians, in their role as patient advocates, to control the budgets for treating them. Committed fund-holders believed that fund-holding made it easier for them to negotiate better services.[20] Some elements from the project have been introduced to Canada, most notably in Québec as the Family Medicine Groups[21] and in Ontario as the Family Health Teams.[22]

Ensuring equity

High quality primary care is the cornerstone of equitable, efficient, and effective health care.[23,24] In addition to physician services, the notion of primary care includes nurse practitioner services, well-baby and other preventive care, home care, elderly care, and crisis intervention. In some countries, primary care physicians act as gatekeepers to higher levels of care, directing patients to the most appropriate services and following up

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

after consultation, thus reducing inappropriate use of other, more expensive levels of services. Well-managed primary care services ensure preventive care and timely follow-up for all their patients and can act as a resource in advocating for community health. A number of different models of primary care services have been tested in Canada. Their common objective is to integrate services, bringing physicians out of their traditional, unidisciplinary practice and into collaboration with other types of professionals as pivotal members of multidisciplinary teams that provide the services patients need.

The Canada Health Act aims to reduce inequity in access to services by publicly insuring the population for many of its health care needs. The publicly funded Medicare system distributes the costs of care across the population. In principle, nobody has to pay to access medically necessary services in hospitals and physicians' offices, but fees may be demanded for other services. Moreover, affordability is only one aspect of accessibility, and there is also cause for concern in the other aspects of access.

Access to health care services

The concept of access to health care services includes:[25,26]

- **Availability**: the relation between the volume of services provided and the demand for them
- **Accessibility**: the geographical relationship between the location of services and the people who need them
- **Accommodation**: the relationship between the manner in which the services are provided and the constraints of people who need them
- **Affordability**: the relationship between the cost of services and the ability of users and potential users to pay
- **Acceptability**: the extent to which people who need services are comfortable using them.

Availability

While there are many forms of health care available in Canada, in general, publicly financed health care systems provide allopathic medical care only. Allopathic medicine is also highly regulated to ensure the safety of users; standards for other forms of care are, in general less regulated.

Having a regular care provider, waiting times for appointments, and ?unmet need' are the usual measures of health care availability. The availability of health care is often measured by waiting times for care and by unmet need for care. Uptake of preventive care may also indicate its availability.

Illustrative Materials

Availability of health care in Canada[27]

In 2007,

85% of Canadians said they had a regular family physician,

14% found it difficult to get care during office hours because there were no appointments or the office waiting time was too long,

42% of Canadians who had attended an emergency room, said their condition could have been treated by a primary care provider had one been available,

10% of Canadians said that they didn't get health care when they needed it.

64% of Canadians aged 65 and over said they had had a flu vaccine in the last year.

In 2005,

74% of Canadian women aged 18-69 said they had had a Pap smear in the last 3 years,

71% of Canadian women aged 50-69 said they had had a mammogram in the last 2 years.

Accessibility

Rural areas tend to lack physicians and health professionals. In spite of financial incentives to practise in rural areas, physicians cite social, family, and professional reasons for preferring urban practices. Some Canadian medical schools are providing medical training in rural areas in order that the experience will induce physicians to remain in them. For practical and economic reasons, rural areas will likely continue to have less access than urban areas to specialized care. Other ways of providing care, such as joint or shared care and telemedicine need to be developed.

Accommodation

Many people are constrained by work or family responsibilities and so unable to attend clinics during the usual clinic hours. Similarly, people who do not have private transport may be virtually excluded from clinics not served by public transport. Services must take into account the particular problems of vulnerable populations who often are most in need of care and least able to access it. For example, people with limited mobility may need wheel chair access ramps, or people with limited vision need adequate lighting and safeguards around stairwells, many people with substance abuse problems, because of their chaotic lifestyles need mobile outreach services.

Affordability

In principle, all Canadians have access free of charge to medically necessary physician and hospital services. However, extra-billing and user fees are a continuing problem, even though prohibited under the Canada Health Act.[28] Other costs can also be a factor in reducing access. For instance, transport-related costs can be considerable particularly in cases, such as cancer, where treatment may require repeated visits to a specialist centre.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Services of non-physician professionals in the community are not included under the Medicare plan. Under the influence of federal direction as well as a number of reports on health care,[15,19,29] provinces and territories are now beginning to extend the range of publicly financed services to cover certain types of home care and at least some pharmacy care. In general, provinces and territories provide some community services for the most vulnerable sectors of their populations, but other people must pay directly or have private insurance for these services. Nonetheless, as care shifts to the ambulatory setting and the physician's office, the costs of treatment may be passed on to the patient. Financial barriers to care, therefore, still exist, and may increase.

As poverty is a major determinant of health, financial barriers increase inequity in health; they ensure that those most in need of the service have least access to it.

Further Reading

User fees

For further reading on user fees and extra billing, see "Myth: User fees would stop waste and ensure better use of the health care system," September 2001, at the Canadian Health Services Research Foundation website: http://www.chsrf.ca/PublicationsAndResources/Mythbusters/ArticleView/01-09-01/519f623f-b805-4550-93a5-648e60b8aad3.aspx

Acceptability

Acceptability of services depends on a range of culturally determined factors which affect the expectations and attitudes of the user. Minority groups, such as immigrants, English speakers in Quebec, French speakers in other Canadian provinces, or Indigenous Peoples throughout Canada may feel ill at ease with services geared towards the expectations of the majority.[30,31] In particular, poorly educated people on low income can be inhibited in negotiating their way through health services managed and offered, as they are, by highly educated people with a relatively high income.

First Nations, Inuit and Métis (Indigenous) health services

Traditional Indigenous teachings highlight the importance of maintaining and restoring a balance among the physical, mental, emotional, and spiritual aspects of health through social and environmental sensitivity. These teachings were discounted by arriving Europeans who brought with them a way of life that threatened the lives and health of Indigenous Peoples. Infectious diseases arriving with the different waves of immigrants had a devastating impact on Indigenous Peoples because they had no immunity to them. Through colonization, the European way of life became the norm and Canada's Indigenous Peoples found themselves excluded and denigrated. As a result, their health deteriorated compared to that of the dominant society. The health gap between many First Nations, Inuit and Métis communities and the rest of Canada broadened and to this day remains huge (see Chapter 1).

Health services alone cannot significantly reduce the health gap between Indigenous Peoples and other Canadians. Reducing the gap would require attention to income, education, social and physical environment, and housing and sanitary infrastructure, as well as the restoration of traditional lands, governance and culture. However, in spite of the much greater service need among Indigenous Peoples, their health services lack coordination. Although the federal government retains responsibility for providing care for registered Indian and Inuit groups, the services are increasingly delivered by provinces, territories, and band councils on reserves and in Indigenous communities. These services may not be well oriented towards the communities' needs. They tend to be staffed by non-Indigenous people and, until recently, the First Nations communities had little say in the planning of their health services. There are no specific services for First Nations people living off-reserve. Mainstream institutions and professionals who serve Indigenous people living off-reserve rarely have the resources or training to provide culturally sensitive care.

In 2000, the National Aboriginal Health Organisation (NAHO) was established. Funded by Health Canada, it is an "Aboriginal-designed and -controlled body committed to influencing and advancing the health and well-being of Aboriginal Peoples by carrying out knowledge-based strategies."[32] In the same year, the Institute of Aboriginal People's Health was established as one of the Canadian Institutes of Health Research (CIHR) to support research and build research capacity in Indigenous peoples' health.[33] Nonetheless, the political nature of health service provision and the wide variety of issues to be addressed continue to create barriers to health for First Nations, Inuit and Métis Peoples and will likely do so for some time.

Further Reading

National Aboriginal Health Organization

For more about health issues of Aboriginal People in Canada, see the National Aboriginal Health Organization's website at http://www.naho.ca/

The Aging Population

There is widespread debate over the likely impact of the aging of the Canadian population on the future demand for health services. The basic fact is that Canadians are living longer and this, combined with the falling birth rate, means that the proportion of the population aged 65 or over is rising (See Figure 12.4).

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

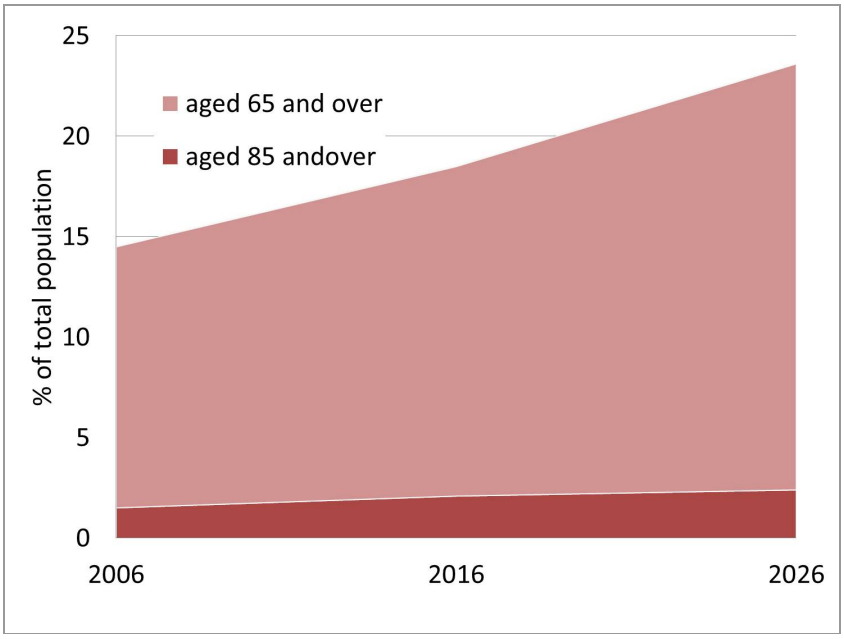


Figure 12.4 The aging population: projections of proportion of elderly people, Canada, 2006 to 2029. Note that the proportion of the population aged 85 and over will increase from 1.5% in 2006 to 2.4% in 2026

Further Reading

Animated population pyramid

For a graphic description of the Canada's changing population structure that shows the effects of the postwar baby-boom, explore Statistic Canada's animated population pyramids at

http://www.statcan.gc.ca/kits-trousses/animat/edu06a_0000-eng.htm

The impact this will have on health service use is widely debated, and many factors are involved. It is commonly feared that, as life expectancy increases, disability and chronic disease will rise in the population, so health services should prepare to care for a greater proportion of the population than ever before. However, set against this are the possibilities that people are becoming more aware of their health and population health promotion will maintain health for longer, reducing the duration of ill-health a person experiences before dying. (See squaring the morbidity curve, Figure 8.2).

Economists have found that adults use health services most intensively in the four to six months before they die, no matter at what age death occurs. In fact, younger people who die may cost the health services more than older people because the efforts to save them are more heroic. Controlling for the proximity to death shows that the major factor driving use of acute services (and hence costs) is the number of deaths rather than the age of the population.[34] This means that young populations with a high death rate, such as some in remote and rural areas of Canada, may need relatively more services than populations with a lower death rate even if the latter are older.

Focusing on the aging of the population may be a distraction from the more pressing issue of how to allocate resources to care for people at the end of life. The growing use of medications, intensive interventions and high-technology equipment influence the cost of care. A greater acceptance and use of palliative options at the end of life may well reduce the cost of care for a dying person. This could result in a drop in the overall cost of care, although the increasing use of primary care services by the elderly and their greater need for long term care may in part offset this. So the future influence of population aging is uncertain.

Further Reading

Myth of the aging population

To read about the effect of the aging population in more detail, see "Myth: The aging population will overwhelm the health care system" January 2002, at the Canadian Health Services Research Foundation website: http://28784.vws.magma.ca/mythbusters/html/myth5_e.php

Information and dissemination of information in health care

Information technology promises ways of managing and disseminating information for improving the coordination of services, communication between care providers, and dissemination of research findings. Canada is currently investing in the ?Health Infoway', which should improve efficiency in service provision and communication. In spite of the many organizations that disseminate evidence-based guidelines and models of excellence for practice, success in getting the evidence into practice is still elusive.

Some developments in informatics that are of interest to clinicians include:

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

- *Electronic medical and health records.* These have been deployed in many institutions and medical offices throughout Canada. They provide efficient storage and retrieval of information about patients, they can be built into local networks to allow quick transfer of information between hospitals and referring physicians, and they can be used as basis for call and recall systems that facilitate preventive programmes and disease follow-up. A major concern is to ensure confidentiality when transferring information between institutions and offices; another is to develop compatibility between systems in building networks. Finally, electronic medical and health records, if the system is well designed, can provide useful information for evaluation of practice.
- *Telemedicine.* First-line carers and patients can consult with specialists via videoconferencing. Variations on this concept include telemonitoring in which anthropometric data from patients can be communicated to a specialist service. There has also been experimentation with telesurgery using robots or with video monitoring through which a specialist can guide a general surgeon. Videoconferencing and its derivatives can be useful in bringing specialist services to remote areas.
- *Teletraining.* There are a number of useful services that provide continuing web-based training for clinicians. Podcasts, videocasts and interactive training programmes are all available.

Some specific types of services

Occupational health services

Occupational health services in Canada are divided into two distinct sections. The first aims to protect workers and prevent work-related injury and ill health. It is generally known as health and safety at work, although the names of the legislation and authorities governing it vary between provinces and territories. The second section aims to rehabilitate people who have been injured at work or who suffer from occupational illness, and to compensate workers whose health has been damaged by work. This is generally administered by a workers' compensation board?again, names differ. Some large corporations in Canada maintain their own occupational health services, but must still respect occupational health legislation.

Health and safety at work

The provinces and territories are responsible for providing occupational health services. They do so within the framework of the *Canadian Labour Code*. Employees of most industries are covered by provincial legislation, and the types of industry covered vary slightly by province or territory. Certain industries that cross provincial and national boundaries are covered under federal legislation. Some employees, such as those working in domestic employment (for instance, domestic workers in private households), are generally not covered. The authorities responsible for providing services also vary by province or territory.

Further Reading

Health and safety at work

For more about health and safety at work in Canada, visit the Canadian Centre for Occupational Health and Safety website at: http://www.ccohs.ca/ and the Human Resources and Skills Development website at : http://www.hrsdc.gc.ca/eng/labour/health_safety/index.shtml

In spite of provincial differences in organization, all Canadian workers have certain rights and duties under the Labour Code:

Rights of workers include

- Right to **know**. That is the right to know about work-related hazards. This includes the training and supervision necessary to protect the worker's health.

- Right **to participate** in health and safety. Employers with 300 or more employees are required to establish a health and safety committee, which includes representation from workers and from management. Its objective is to develop a health and safety policy. Other employers (those with fewer than 300 employees) may choose to do so. Workers have the right to participate in the committee and policy deliberations. They may also participate through the internal complaint resolution process covered under the Canadian Labour Code.

- Right to **refuse**. Employees can refuse to work if the work is dangerous to themselves or others and if the danger is not a normal condition of employment.

The duties of workers include; using the safety equipment provided, respecting health and safety procedures, instructions and policies that relate to the hazard or to the use of safety equipment, and reporting potential workplace hazards as well as injuries and hazardous events.

The duties of employers are; to provide a safe workplace, to respect safety standards, and to ensure that workers receive the information and training they need to protect their health.

Workers' compensation

Organized workers' compensation in Canada began with the 1913 Meredith Report, and is administered by Workers' Compensation Boards. One of the major goals of most workers' compensation boards is to get workers with occupational illness back to work. To do so, compensation may cover rehabilitation costs, including, for instance, outpatient physiotherapy services. As such, workers' compensation can provide a broader range of services than does the provincial health care plan. The ill worker remains under the care of his or her usual treating physician who, on request of the patient, must provide a letter to support the patient's claim for compensation and will continue follow-up of the patient's condition. Apart from medical care, return to work may require changes in the patient's work-station, entailing work with, for instance, occupational therapists or occupational

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

hygienists and the patient's employer.

Five basic principles of workers' compensation.

The basic principles of workers' compensation were set out in the 1913 Meredith Report and survive today. They are:

1. *No-fault compensation.* Workplace injuries are compensated regardless of fault. Neither worker nor the employer can sue, nor are they expected to admit responsibility, except in a case of gross negligence.

2. *Collective liability.* The total cost of the compensation system is shared by all employers in proportion to the costs of claims for each occupational sector. For instance, the forestry sector pays higher premiums than the office work sector. All employers contribute to a common fund and financial liability becomes their collective responsibility. However, employers' contributions may be adjusted to reflect the cost of successful claims of their workers.

3. *Security of payment.* Injured workers are assured of prompt compensation and future benefits.

4. *Exclusive jurisdiction.* All compensation claims are directed solely to the compensation board. The board is the decision-maker and final authority for all claims. The board has the authority to judge each case on its individual merits.

5. *Independent board.* The governing board is both autonomous and non-political. The board is financially independent of government or any special interest group. The administration of the system is focused on the needs of its employer and worker clients, providing service with efficiency and impartiality.

Further Reading

Workers' compensation

Information about workers' compensation, including the Meredith Report and workers' compensation boards, is available at the Association of Workers' Compensation Boards of Canada website at: http://www.awcbc.org/en/index.asp

Public health services in Canada

Development of the public health system

Public health activities remained fairly uncoordinated until relatively recently. As cities grew, water and sanitation services developed with engineering advances, and were coordinated by municipalities. Quarantine and isolation became important in the early and mid-19th century, when waves of immigrants brought cholera and typhus with them to Canada. In the early 20th century some towns began to chlorinate water, rural towns began to pasteurize milk, and Ontario began to immunize against smallpox and diphtheria.

Further Reading

Grosse Île

Grosse Île in Québec was the main quarantine station in Canada in the mid-nineteenth century. It is now a national park. Information about it and its history can be found at the Parks Canada website: http://www.pc.gc.ca/eng/lhn-nhs/qc/grosseile/index.aspx

In 1974, the Lalonde Report emphasized the need to look beyond the care of the sick in order to improve the health of the population, so that the health care system should include action on environment, lifestyle and health care organization, as well as biology. It was the first Canadian report on the health system to mention the importance of health promotion and prevention in maintaining population health. In 1986, with the first international conference of health promotion being held in Ottawa, the then Minister for Health and Welfare, Jake Epp, presented "Achieving Health for All: A Framework for Health Promotion." This report reflected the *Ottawa Charter for Health Promotion* (which had been ratified at the same conference); the Framework set out the direction for health promotion in Canada. It said that Canada should attempt to reduce inequities, to increase the prevention effort, and enhance people's capacity to cope. It suggested that this could be achieved by fostering public participation, strengthening community health services, and coordinating public health policy.

Until the turn of the millennium, public health continued to take a back seat to personal health services and received about 3% of overall health funding. Although some provinces, such as British Columbia and Quebec, had developed coherent structured public health systems, including provincial organizations providing public health expertise (BC-CDC in British Columbia and the INSPQ in Quebec), in most provinces public health provision remained poorly coordinated.

In 2000, an E. coli outbreak killed seven people in Walkerton, Ontario, and affected thousands of others. Then, in 2001, around 6,000 people in

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

North Battleford, Saskatchewan, contracted cryptosporidiosis because of problems with the water supply. In 2002 and 2003, SARS, a previously unknown disease, reached near pandemic levels causing over 8,000 cases in 16 countries. Forty-four Canadians, mostly in Toronto, died. Meanwhile, experts in public health were warning of an impending influenza pandemic. During the same period, the 9/11 attacks in New York and several terrorist attacks in Europe and Asia occurred, and extreme weather conditions were causing death and injury around the world. These man-made and natural disasters increased awareness of the need for public health services disaster planning around the world.

The events in Canada demonstrated the weakness of the public health infrastructure. In response, the 2006 *Public Health Agency of Canada Act* came into force establishing a Chief Public Health Officer for Canada who would

- Advocate for effective disease prevention and health promotion programmes and activities
- Provide science-based health policy analysis and advice to the federal minister for health
- Provide leadership in promoting special health initiatives
- Improve the quality of public health practice.

The Act also created the Public Health Agency of Canada, whose mission is to promote and protect the health of Canadians through leadership, partnership, innovation, and action in public health. The agency would

- Concentrate and focus federal resources
- Enhance collaboration between different levels of government
- Allow faster, flexible response to emergencies
- Improve and focus communication
- Allow for longer-range plans than the usual annual planning cycle of governments
- Achieve greater success in attracting and retaining public health professionals.

In keeping with its origins, the agency's immediate priorities were emergency preparedness, immunization, and chronic disease prevention.

Definition

Public Health.

The Chief Public Health Officer of Canada uses Last's[35] definition of Public Health: "The organized efforts of society to keep people healthy and prevent injury, illness and premature death." It is a combination of programmes, services, and policies that protect and promote the health of Canadians.[36]

Public Health Responsibilities in Canada

According to the Public Health Agency of Canada, **public health** is concerned with six essential activities:[36]

1. *Health protection:* This includes ensuring that water, air and food are safe, maintaining the regulatory framework for the control of infectious disease and protection from environmental threats, as well as advising on food and drug safety regulations.
1. *Health surveillance:* The ongoing, systematic use of routinely collected health data for the purpose of tracking and forecasting health events or health determinants. It includes
    - The collection and storage of relevant data
    - The integration, analysis, and interpretation of these data
    - The production of tracking and forecasting products with the interpreted data
    - Publication and dissemination of those products
    - Provision of expertise to those developing or contributing to surveillance systems, including risk surveillance.

The information produced by surveillance is used in planning services and prevention programmes.

1. *Disease and injury prevention:* The investigation, contact tracing, and development of preventive and control measures to reduce the risk of infectious disease emergence and outbreaks as well as the promotion of safe, healthy lifestyles to reduce preventable illness and injuries.
1. *Population health assessment:* Understanding the health of communities or specific populations, as well as the factors that underlie good health or pose potential risks, to produce better policies and services.
1. *Health promotion:* Preventing disease, encouraging safe behaviours, and improving health through public policy, community-based interventions, active public participation, and advocacy or action on environmental and socio-economic determinants of health.
1. *Emergency preparedness and response:* Planning for natural disasters (e.g., floods, earthquakes, fires, dangerous infectious disease) and man-made disasters (e.g., those involving explosives, chemicals, radioactive substances, or biological threats and social disruption).

As with general health care services, responsibility for public health is shared between the federal, provincial, and territorial governments. Provinces

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

and territories are responsible for providing services to their population; however, the organisation and the services offered vary across the country.

Further Reading

Provincial public health services

To find out how public health services are delivered in each of the provinces and territories, visit the National Collaborating Centre for Healthy Public Policy website: http://www.ncchpp.ca/en/structuralprofile.aspx

Coordination of world health

The World Health Organization is the directing and coordinating authority for health within the United Nations system. It was created in 1948 and is responsible for providing leadership on global health matters, shaping the health research agenda, setting norms and standards, articulating evidence-based policy options, providing technical support to countries, and monitoring and assessing health trends. It currently focuses on six approaches to improving health:

1. Promote development

2. Foster health security

3. Strengthen health systems

4. Harness research, information and evidence

5. Enhance partnerships

6. Improve system performance.

Recent successes were the agreement on a code of practice for international recruitment of health personnel, and a framework for action on interprofessional education. It is, however, confronting several other major global problems, including food safety and fraudulent medical products. The WHO must also monitor the health-related Millennium Development Goals, and develop strategies to reduce the harmful effects of alcohol, and to address the rise of chronic non-communicable diseases, including the issue of marketing food and beverages to children, which contributes to obesity and poor dental health.[37]

Responsibility for Promoting Health

A brief look at the list of health determinants confirms that no one authority can be responsible for the health of a population. Health promoters believe that it should be the responsibility of all members of a population or community, and a core responsibility for all government.[38,39] Although those working within the health services have a particular responsibility to advocate for health, expertise in public and population health can be found in a variety of different academic disciplines, professions, and organizations. Experts in public health come from a variety of different backgrounds and may use different names to describe their expertise.

Community mobilization is the strategy of choice for sustainable health promotion. In this approach, community members are involved in defining the problems and in proposing solutions. Unlike classic top-down health education, where health professionals study and prioritize problems and then develop solutions, community mobilization involves community members in the process of defining and transforming problems. It is a long-term process that empowers communities, allowing them to take over health-promoting action.

Nerd's Corner

Who does public health?

Public health teams comprise a wide range of professionals with interests in diverse aspects of health. Many have a basic training in the clinical sciences, such as physicians, nurses, psychologists, social workers, dieticians, kinesiologists etc. They work at all levels in the health system from the provincial and federal government to the local and regional health authorities and some work in community health service centres. Those who work in public health departments manage and deliver public health programmes, which generally include transmissible and environmental disease protection, well baby, vaccination and sexual health programmes. Statisticians and epidemiologists carry out health surveillance.

The Royal College of Physicians and Surgeons designates **public health and preventive medicine** as the branch of medicine concerned with the health of populations. Through interdisciplinary and intersectoral partnerships, the public health and preventive medicine specialist measures the health needs of populations and develops strategies for improving health and well-being through health promotion, disease prevention, and health protection.[40]

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Public health law

Much of public health law was drawn up in the 19th century when the discipline of public health was gaining recognition. Quebec, having invested in public health for some time, is the first province to have developed a sophisticated system that was consolidated by its adoption of its *Public Health Law* in 2001. The most notable recent event in Canada was the creation of the Public Health Agency of Canada. Public health law is still evolving; it has the following characteristics:

- Recognition of the special responsibility of government in public health
- The focus on population health
- The regulation of the relationships between the state and populations, and between the state and individuals who may pose a risk to the public's health
- Government provision of public health services
- The government power to coerce individuals and businesses in order to protect the public's health.

Public health law respects the following principles:

- Duty of government to protect the health and well being of the population
- Power to set standards of health and safety and ensure compliance
- Restraint in the exercise of power; should act only on the basis of clear criteria where necessary, procedural due process.

Public health law assigns the government the legal power and duty to ensure conditions for people to be healthy (e.g., to identify, prevent, and ameliorate risks to health in the population). It places constraints on autonomy, privacy, liberty, propriety, or other legally protected interests of individuals for the protection or promotion of community health, but it also places limitations on the power of the state to constrain these interests.[41]

Self-test questions

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Self-Testing

1. What are the advantages and disadvantages of publicly funded and privately funded health care?

	Advantages	Disadvantages
Publically funded	Possibility of a single payer system that reduces costs. Everyone can be covered so that people do not face financial ruin because of illness. Health care resources can go to those that need them. Redistributes wealth in the population, so affecting a major determinant of the health of populations and individuals.	Can reduce individuals' liberty. Health care is a major political issue.
Privately funded	Generally ensures that care is available when and where the patient demands it, particularly for acute, self-limiting conditions.	Resources go only to those that can afford them, therefore those least likely to need them. Can drain human resources from co-existing public systems. Coverage for chronic disease is likely to be inadequate or excluded due to the cost of care for chronic disease. People at high risk of chronic disease or who already have may not be able to afford premiums or get coverage at all.

1. What are the likely effects of the aging population on the health care system?
The aging population is unlikely to increase the cost of acute care. The causes of increasing acute care costs are rooted in the development of technologies, particularly pharmaceuticals, and changing health behaviour. The cost of chronic care is likely to increase. A strong primary health care system that includes different types of professionals and services will be necessary to meet the needs of the aging population.
2. Assume you are treating a patient who has had a disabling cerebrovascular accident. Which allied health professionals could contribute to the management of the disability?
A wide variety of professionals and non-professionals could contribute to the management of the disability depending on which functions have been affected. Mobility problems can be aided by physiotherapy. Occupational therapy can help the patient and family find ways to adapt to the dysfunction. Speech therapy can be required when Broca's speech area has been affected or when the patient has difficulty swallowing. Previously existing hearing and sight problems can require further attention to maintain optimal function. Social and psychological care can help the patient and family adapt to loss of function and social workers can help the patient access the benefits to which he or she is entitled. Respiratory therapists can help prevent respiratory problems due to immobility; pharmacists can oversee prescription drugs and warn of interactions; dieticians may be required to advise in adjusting the diet for secondary prevention and to ensure an adequate nutritional state.

Non-professional services, such as home care and respite care can be of great benefit to the patient and the carer. Patient associations or associations of elderly people can improve social support and provide a social network to prevent isolation of patients and families. Associations, such as the heart and stroke association of Canada can provide patient information to improve health and to access benefits.

Discussion Points

- 1. What non-physician services are available outside hospital in your area?
- 2. How are health services for Aboriginal Peoples organized in your province or territory?

References

Further Reading

Report of the National Forum on Health. Canada health action: building on the legacy. Ottawa: Health Canada, 1997. http://www.hc-sc.gc.ca/hcs-sss/pubs/renewal-renouv/index-eng.php

The Kirby Report. Report of the standing senate committee on social affairs, science and technology. Study on the state of the health care system in Canada, 1999-2002. Ottawa: Government of Canada, 2002. http://www.parl.gc.ca/common/committee_SenRep.asp?Language=E&Parl=37&Ses=2&comm_id=47

The Romanow Report. Report of the Commission on the future of health care in Canada. Building on values. Ottawa: Government of Canada, 2002. http://publications.gc.ca/site/eng/237274/publication.html

Reid TR, The healing of America: a global quest for better, cheaper, and fairer health care. Harmondsworth, UK: Penguin, 2009.

1. Rothgang H, Cacace M, Grimmeisen S, Wendt C. The changing role of the state in healthcare systems. European Review 2005;13(supp 1):187-212.

2. Collings FD. National hospital insurance in Canada. Public Health Rep 1959 Dec;74:1049-1058.

3. Lindenfield R. Hospital insurance in Canada: An example in federal-provincial relations. Soc Serv Rev 1959 June;33(2):148-160.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

4. Canadian Institute for Health Information. National Health Expenditure Trends 1975 to 2009. 2009; Available at: http://secure.cihi.ca/cihiweb/products/National_health_expenditure_trends_1975_to_2009_en.pdfOttawa, Ont.

5. Woolhandler S, Campbell T, Himmelstein DU. Costs of health care administration in the United States and Canada. N Engl J Med 2003 Aug 21;349(8):768-775.

6. Menec VH, Lix L, MacWilliam L. Trends in the health status of older Manitobans, 1985 to 1999. Canadian Journal on Aging 2005;24(Suppl 1):5-14.

7. Watson DE, Heppner P, Reid R, Bogdanovic B, Roos N. Use of physician services by older adults: 1991/1992 to 2000/2001. Canadian Journal on Aging 2005;24(Suppl 1):29-36.

8. Statistics Canada. Labour Force Statistics. Available at: http://www.statcan.gc.ca/subjects-sujets/labour-travail/lfs-epa/lfs-epa-eng.htm, December, 2009.

9. Canadian Institute for Health Information. Health Care in Canada. Ottawa, Ont: CIHI; 2008.

10. Davis K. Taking a walk on the supply side: 10 steps to control health care costs. Available at: http://www.commonwealthfund.org/Content/From-the-President/2005/Taking-a-Walk-on-the-Supply-Side--10-Steps-to-Control-Health-Care-Costs.aspx. Accessed August, 2009.

11. Ginsburg PB. Controlling health care costs. N Engl J Med 2004 Oct 14;351(16):1591-1593.

12. Stewart J. Manitoba targets patient overuse as it tackles abuse of health care system. CMAJ 1995 May 1;152(9):1483-1484.

13. Deber RB. Delivering Health Care Services: Public, Not-For-Profit, or Private? Discussion Paper No. 17. 2002 August;2010(August):163http://www.teamgrant.ca/M-THAC%20Greatest%20Hits/Bonus%20Tracks/Delivering%20Health%20Care%20Services.pdf.

14. Groupe de travail (Roland Arin,President). La complémentarité du secteur privé dans la poursuite des objectifs fondamentaux du système public de santé au Québec. Québec: Ministère de la santé et des services sociaux; 1999.

15. Romanow RJ. Building on values: The future of health care in Canada. Privy Council of Canada publication number CP32-85/2002E-IN November; 2002.

16. Taylor DH,Jr. What price for-profit hospitals?. CMAJ Canadian Medical Association Journal 2002 May 28;166(11):1418-1419.

17. Donabedian A. The quality of care. How can it be assessed? JAMA 1988 Sep 23-30;260(12):1743-1748.

18. Donabedian A. Evaluating the quality of medical care. 1966. Milbank Q 2005;83(4):691-729.

19. Standing Senate Committee on Social Affairs, Science and Technology. Study on the state of the health care system in Canada, 1999-2002 (The Kirby Report). 2002 Octoberhttp://www.parl.gc.ca/common/committee_SenRep.asp?Language=E&Parl=37&Ses=2&comm_id=47#Substantive%20Reports.

20. Kay A. The abolition of the GP fundholding scheme: a lesson in evidence-based policy making. British Journal of General Practice 2002 Feb;52(475):141-144.

21. Évaluation santé et services sociaux. Évaluation de l'implantation et des effets des premiers groupes de médecine de famille au Québec. 2008 décembrehttp://msssa4.msss.gouv.qc.ca/fr/document/publication.nsf/fb143c75e0c27b69852566aa0064b01c/eadedf781c2bf119852575a200649099?OpenDocument&Highlight=0,GMF.

22. Ontario Ministry of Health and Long Term Care. Family Health Teams. Available at: http://www.health.gov.on.ca/transformation/fht/fht_mn.html. Accessed January, 2010.

23. International Conference on Primary Health Care,Alma-Ata, USSR. Declaration of Alma-Ata. 1978 September.

24. Chan M. Primary health care - Now more than ever. 2008; Available at: http://www.who.int/dg/speeches/2008/20081014/en/index.html. Accessed September, 2010.

25. Penchansky R, Thomas JW. The concept of access: definition and relationship to consumer satisfaction. Med Care 1981 Feb;19(2):127-140.

26. Clark DW. Dimensions of the concept of access to health care. Bull N Y Acad Med 1983 Jan-Feb;59(1):5-8.

27. Leatherman S, Sutherland K. Quality of healthcare in Canada: A chartbook. Ottawa, Ont.: CHSRF; 2010.

28. Ministers of Health. Canada Health Act Annual Reports. 2002-2003 to 2007-2008http://www.hc-sc.gc.ca/hcs-sss/pubs/cha-lcs/index-eng.php.

29. Report of the National Forum on Health. Canada health action: Building on the legacy. 1997; Available at: http://www.hc-sc.gc.ca/hcs-sss/pubs/renewal-renouv/index-eng.php. Accessed June, 2010.

30. Asanin J, Wilson K. "I spent nine years looking for a doctor": exploring access to health care among immigrants in Mississauga, Ontario, Canada. Soc Sci Med 2008 Mar;66(6):1271-1283.

31. Levin R, Herbert M. The experience of urban aboriginals with health care services in Canada: implications for social work practice. Soc Work Health Care 2004;39(1-2):165-179.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

32. National Aboriginal Health Organization. Available at: http://www.naho.ca/english/index.php. Accessed October, 2009.

33. Institute of Aboriginal Peoples' Health. Available at: http://www.cihr-irsc.gc.ca/e/8668.html. Accessed October, 2009.

34. Werblow A, Felder S, Zweifel P. Population ageing and health care expenditure: a school of 'red herrings'? Health Econ 2007 Oct;16(10):1109-1126.

35. Last JM. A dictionary of public health. New York: Oxford University Press; 2007.

36. Butler-Jones D. The Chief Public Health Officer's Report on the Sate of Public Health in Canada, 2008. Minister of Health ; 2008.

37. The CPHA and Global Health Governance. CPHA Update 2010;101(3):191.

38. World Health Organization. Ottawa charter for health promotion. 1986; Available at: http://www.who.int/hpr/NPH/docs/ottawa_charter_hp.pdf. Accessed July 28Geneva, 2011.

39. World Health Organization. Bangkok Charter for Health Promotion in a Globalized World. 2005; Available at: http://www.who.int/healthpromotion/conferences/6gchp/bangkok_charter/en/. Accessed August, 2009.

40. Royal College of Physicians and Surgeons of Canada. Objectives of Training in Public Health and Preventive Medicine. 2003(April).

41. Gostin LO. Public health law reform. Am J Public Health 2001 Sep;91(9):1365-1368.

Chapter Objectives

After reading this chapter, you will be able to:

- Describe the approaches to assessing quality of care and methods of quality improvement;
- Describe and understand the multiple dimensions of quality in health care, i.e., what can and should be improved.

Case Study

Dr. Rao visits Mrs. Richards in her long-term care home. He notices her blood pressure is high. He is about to write a prescription for an antihypertensive medication. First, however, he reviews her current medications and sees that a locum physician prescribed a non-steroidal anti-inflammatory drug (NSAID) for osteoarthritis the previous summer. He suspects that the hypertension is an adverse drug reaction to the NSAID, and so he advises her to discontinue the NSAID and take Tylenol rather than starting an antihypertensive medication. He schedules another visit to check the effects of his management.

Patient safety

Every health care professional wants to do a good job and help patients. However, things do not always go according to plan. Sometimes patients' laboratory results get lost and tests have to be repeated. Sometimes patients are given medications to which they are known to have contraindications. Practices are not always based on the best available evidence. Occasionally, errors and poor quality practice are due to professionals' lack of knowledge and skills; more often, they are due to much broader problems within the complex system in which they work. Prevention of these errors requires a culture of open communication and quality improvement rather than a culture of blame.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Definition

From the outset, three concepts should be distinguished:

1. **Adverse outcome**: an unsatisfactory outcome that arises as part of the natural course of the disease, for instance, a gastrointestinal bleed due to a stress ulcer in a patient with pneumonia;

2. **Adverse event**: unintended injury or complication that was caused by health care management, rather than by the patient's underlying disease, and that leads to death, disability or prolonged hospital stays.[1] Adverse events include complications that arise from treatment, inherent (and often unforeseeable) risks of interventions, and errors. For instance, a gastro-intestinal bleed may still occur in a patient on warfarin treatment that is adequately monitored.

3. **Medical error**: an adverse event that was preventable. For instance, a gastro-intestinal bleed in a patient on warfarin treatment that was not monitored adequately. Medical errors include: errors of execution, such as **slips** (operating on the wrong side), **lapses** (forgetting to check for allergies), or **fumbles** (needlestick injuries). These are errors in automatic functioning that are often due to fatigue or stress. Other errors are due to **mistakes** or **failures of reasoning** that result in a poor choice of plan (treating for heartburn when the patient has classic symptoms and signs of myocardial infarction) can be due to a failure in attitude resulting in inadequate investigation or gaps in knowledge.

Because many adverse events arise from inherent risks and not from errors, it is prudent to use the term "adverse event" unless error has definitively been established.

Patient safety has become a universal concern, and is more likely to be improved by overall quality improvement than by merely reacting to adverse events. In Canada, the Romanow Commission reported that the responsibility for ensuring quality and safety is distributed among a number of different players and organizations in the health care system, including professional and regulatory bodies.[2] In such a system of divided responsibilities, lines of accountability can be unclear, so activities may not be carried out or may not be followed up. Furthermore sharing of information and learning is inhibited. Following the publication of the Canadian Patient Safety Strategy, a Patient Safety Institute was set up in 2003. Also, the Health Council of Canada was created in 2004 to provide a perspective on health care issues, including quality and safety, and their impact on the health of Canadians. Most provinces have created agencies responsible for examining the quality of care and for informing the public and the government of their findings. Many hospitals have some kind of safety audit, often in the form on morbidity and mortality conferences, and physicians can be required to carry out regular assessments of their practice. However, coordinated and integrated structures to support patient safety are still lacking.[3]

Further Reading

Quality and safety in health care

The Canadian Patient Safety Strategy sets out approaches to ensuring patient safety

http://rcpsc.medical.org/publications/building_a_safer_system_e.pdf

The Canadian Patient Safety Institute provides for exchange of ideas and tools and resources relating to patient safety

http://www.patientsafetyinstitute.ca/English/Pages/default.aspx

The Health Council of Canada produces general information about health care quality and outcomes

http://healthcouncilcanada.ca

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

One method of classifying adverse events is to relate them to the stage of medical intervention at which they occur.[4]

Diagnostic

- Error or delay in diagnosis
- Failure to employ indicated tests
- Use of inappropriate tests
- Use of outmoded tests or therapy
- Failure to act on results of monitoring or testing

Treatment

- Error in the performance of an operation, procedure, or test
- Error in administering the treatment
- Error in the dose or method of using a drug
- Avoidable delay in treatment or in responding to an abnormal test
- Inappropriate (unnecessary) care

Preventive

- Failure to provide prophylactic treatment
- Inadequate monitoring or follow-up of treatment

Other

- Failure of communication
- Equipment failure
- Other system failure

Nerd's Corner

Appropriate use of services

A simple way to define quality is in relation to how health services are used: they may be underused, overused, or misused.[5]

- *Underuse:* Failure to apply effective care that would improve outcome. Examples include low income women not getting routine Pap smears, women aged 50-69 not getting routine mammography, and patients not adhering to therapy for high blood pressure.
- *Overuse:* Providing treatment even if cost exceeds benefit. Examples include prescription of antibiotics for minor upper respiratory tract infections and ordering an X-ray of an ankle when the Ottawa Ankle Rule indicates that a fracture is improbable.
- *Misuse:* Order tests or provide treatments that increase cost or the risk of complication and that are not significantly better than other options. Examples include using the latest, most expensive antibiotic when a cheaper or generic one would be equally effective, or giving a drug to a patient who is known to be allergic to it.

This simple classification of quality is useful when patient safety is the only focus, but it is one-dimensional. In fact, quality assessment encompasses all the direct and indirect factors that determine patient safety as well as those relating to patient and provider satisfaction (see Table 13.1).

Adverse drug events

Adverse drug events are a particular area of concern, especially in older adults and people with complex disease. In these cases, a ?prescribing cascade' may occur; this happens when an adverse drug event is not recognized as such but is diagnosed as a new medical problem that is then treated with another drug.[6] Adverse drug events are probably between the fourth and sixth leading causes of death of hospitalized patients.[7] In long-term care facilities, one study estimated that two thirds of residents experienced an adverse drug event over a four year period, one in seven of these events resulted in a hospitalization.[8] The Beers Criteria and similar lists indicate drugs to avoid in older adults to avert adverse drug events.[8,9]

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

Table 13.2 The ABCDEF of adverse drug reactions[10]

Types	Descriptions	Examples
Augmented	Dose-related - common and predictable reaction - related to the pharmacological mechanism of action of a drug	Digoxin toxicity
Bizarre	Non-dose related - uncommon and unpredictable - not related to the pharmacological mechanism of action of a drug	Penicillin hypersensitivity
Chronic	Dose-related and time-related - uncommon - related to the cumulative dose	Chronic oral steroid use
Delayed	Time-related - uncommon - usually dose-related - occurs some time after the use of a drug	Radiotherapy-induced cancer
End of use	Withdrawal of drug - uncommon - occurs soon after withdrawal of drug	Opiate withdrawal syndrome
Failure	Unexpected failure of therapy - common - dose-related - often caused by drug interaction	Reduced efficacy of oral contraceptive when used with antibiotics

Quality in health care

Definition

Quality of care is the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with professional knowledge.[5]

Defining quality

Quality improvement is much more than the prevention of adverse events. It includes a number of dimensions. Different authors include different dimensions in their descriptions of quality. The list in Table 13.1 is a fairly comprehensive one.

Table 13.1: The dimensions of quality[11,12]

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Dimensions of quality	Health service values
Safety	People should not be harmed by an accident or mistakes when they receive care.
Effectiveness	People should receive care that works and is based on the best available evidence.
Patient-centred	Health care providers should offer services in a way that is sensitive to an individual's needs and preferences.
Timeliness	People should get the care they need when they need it.
Appropriateness of resources	Caregivers and professionals should have the appropriate skills for their practice and have access to the information, equipment, supplies, and facilities they need to look after people's health needs.
Efficiency	Supplies, equipment, time, ideas, and information should not be wasted.
Equitableness	People should get the quality of care that is appropriate to their needs regardless of whom they are and where they live.
Accessibility	People should be able to get the right care at the right time in the right setting from the right provider.
Continuity	All parts of the system should be organized, connected, and work with one another to provide high quality care.
Population health	Health care professionals should recognize their responsibility to prevent sickness and improve the health of the wider community and population.
Rewarding	Caring for people who need health care should be a rewarding occupation. Some people consider this a health service value; others do not. Nevertheless, people who consider their jobs rewarding generally produce work of higher quality than those who are unhappy in their work. Salary level is usually only a minor part of what makes a job rewarding.

Unfortunately, dimensions of quality may conflict with each other. For example, the most effective therapy may not be the most patient-centred or the most rewarding for the provider. Hence, it may be necessary to establish a hierarchy among these criteria.

Approaches to managing quality

There are two ways of approaching quality improvement in health care.[13] The inspection model (also known as ?name, blame, and shame'), although ineffective and inefficient, is still in common use. This model operates according to the theory of bad apples (so called because one bad apple can spoil the whole box), which holds that once the ?bad apples' are removed from the system, their spoiling effect is also removed. There are two major drawbacks to this theory. The first is that most people who make mistakes are not bad apples, but act erroneously because of pressures in the system surrounding them. Blaming people encourages them to hide bad practices or close calls, creating a ?culture of silence'. The second drawback is that, unlike bad apples that sit individually and unmoving in a box, the health care system boils with interacting people and complex processes, offering numerous opportunities for error. Monitoring every critical point for error is difficult, consumes resources and has not yet been shown to be effective in preventing mistakes. Furthermore, the standard of quality it achieves can be no higher than the standards imposed by the inspection: usually the bare minimum.

The second and more effective way of ensuring quality is to focus on the system that delivers the health care. This ensures there are multiple checks and safety nets in place to prevent errors, while encouraging communication about errors and near misses, and also inviting ideas for improving quality.

Links

Quality improvement

For more on the concept of quality improvement, see "Myth: We can eliminate errors in health care by getting rid of the ?bad apples'" September 2004, at the Canadian Health Services Research Foundation website: http://www.chsrf.ca/PublicationsAndResources/Mythbusters/ArticleView/04-09-01/3e36e6e7-2120-49cd-98cb-d02395009483.aspx

and "Myth: We can improve quality one doctor at a time" March 2007, also at the Canadian Health Services Research Foundation website: http://www.chsrf.ca/PublicationsAndResources/Mythbusters/ArticleView/07-03-01/d759d156-3fc7-4c61-85d6-c50e9c964bc2.aspx

Providing high-quality health care is the goal of all clinical professionals. Many professional colleges and regulatory bodies (for instance, provincial colleges of physicians and surgeons) have incorporated quality improvement practices into their mandate and operational approaches. Most have quality assurance committees and routinely conduct audits of professional practices. Whereas in the past many of these bodies merely responded to complaints, they are now studying and implementing ways of improving quality. In turn, their relationship with clinicians is shifting from a focus on discipline to one of continuing professional education and training.

Investigating quality

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

In most medical situations, vulnerable people are protected from possible hazards by several layers of barriers. The most immediate barrier is the skill and knowledge of the people caring for them. Routines in the way procedures are carried out also protect patientsFor instance good sterile technique reduces wound infection and good management and organization helps staff to apply safety procedures and make safe decisions.

No protective barrier protects 100% of the time. Most have transient or lasting defects. Most of the time this does not result in an adverse event because the other barriers present are sufficient to protect the patient. Injury results only when defects in several barriers are aligned, exposing a vulnerable patient to a hazard. As Swiss cheese is known for its holes, the idea of barriers with defects or holes brought James Reason to propose the Swiss cheese model of error (Figure 13.1).[14]

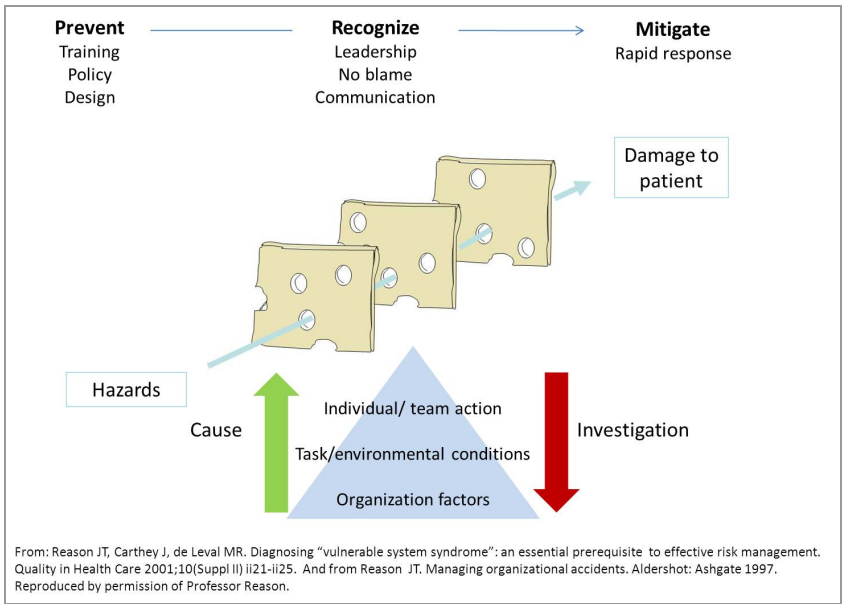


Figure 13.1: The Swiss cheese model of error (adapted from Reason[14,15])

According to the bad apple theory, investigation of problems in quality focuses on the error itself or its immediate cause. This is rarely sufficient to prevent recurrences. The Swiss cheese model focuses on the system to examine all barriers that should have prevented a lapse in quality: individual and team actions, the environment or context in which the error occurred, and the organizational factors that influence the actions and their context. Thinking along these lines, the Toyota Company developed the ?Five Whys' as a simple way to find the root causes of an error. This is adapted to a health context in Figure 13.2.

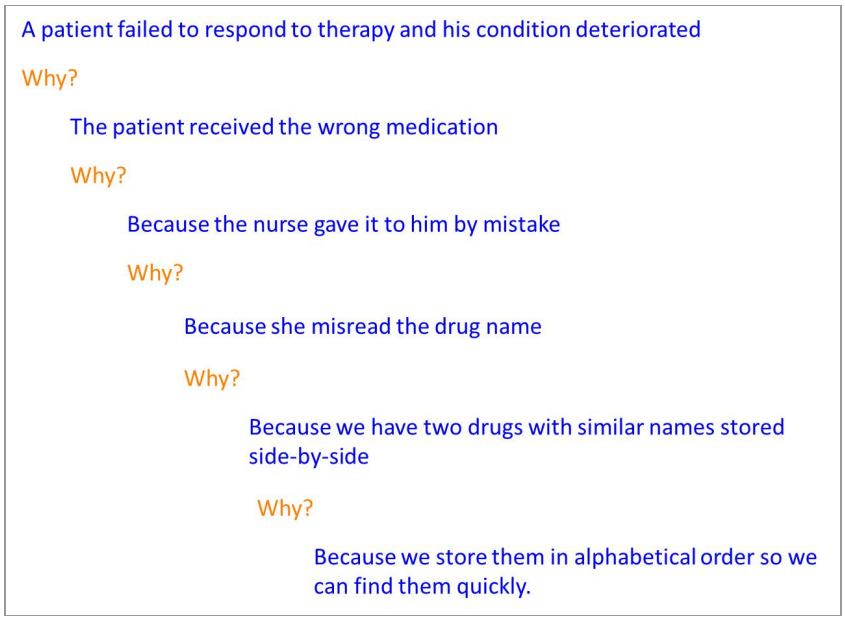


Figure 13.2: Example of the Five Whys.

In reality, there are usually several answers to each ?Why?' question. For instance, the nurse might have been inexperienced or rushed or the light might have been low. Following all the answers leads to a full description of the situation, which gives a number of options for preventing further errors.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Case Study

At a practice meeting, Dr Rao mentions Mrs Richards' hypertension. Here is an excerpt from the Minutes of the meeting:

Dr Middleton: over-medication is a huge problem in healthcare generally.

Nurse Jennings explained that she includes a drug history in her preventive checks, but patients are often quite vague about what they are taking.

Mr Bertoli (practice manager): if we had the money we could buy a computer application that checks recent prescriptions for drug interactions.

Dr Middleton: yes, but when patients get a prescription from a specialist or a walk-in clinic, our computer application wouldn't be able to take account of it.

On his way home, Dr Rao drops into Mr Roy, the local pharmacist, who says that the pharmacies in the area have a system for tracking prescriptions, although it's not perfect. Dr Rao asks if the local physicians could be linked into the system so they could check what drugs their patients are on before prescribing others. Mr Roy says he'll look into it.

After some more discussions in the practice, Nurse Jennings and Mrs Austen, the receptionist, agree that Mrs Austen will instruct patients with chronic illnesses to bring all their drugs with them to their appointment, including the over the counter medicines. Nurse Jennings will list the drugs and how the patients take them and put this list in the chart for the doctors to see.

Dr Rao welcomes this initiative; he sees that they will be able to evaluate it using the number of patients' charts with drug lists. They might even be able to follow how many prescriptions are changed as a result. However, he worries that it will take too much of Nurse Jennings' time so he resolves that he will check with her in a couple of weeks to see how she is doing. He will also check with Mrs Austen to see if the arrangement seems to be increasing the patients' waiting time. Dr Rao wonders if he can persuade the managers of the long-term care home to keep a drug list for each resident.

Improving quality of care

Several principles guide actions to improve quality:[16]

- The primary responsibility for quality rests with the production worker; in health care, this is the clinician.
- Quality assurance systems aid, support and strengthen production systems; they are not simply a form of inspection.
- Workers want to improve quality but need extra energy and skills to do so.
- Vague exhortations, threats and blame are not effective quality improvement interventions.
- Statistical methods should be used to measure performance and compare it with standards.
- Quality assurance is designed to ensure improvement and not to exert control.
- Quality assurance systems should cover all aspects of an organization, including the process of management itself.
- Quality assurance without consumer involvement is inadequate and incomplete.
- Quality assurance focuses on systems and system failure, not on individuals and individual failure.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

An approach to quality improvement

According to Juran, quality improvement consists of three main activities: quality planning, quality control, and quality improvement.[17]

1. Quality planning

- Determine who the patients or populations are and identify their needs:

What are the health problems in the practice population and what can be done about them?

- Orient services being offered to meet these needs:

Do staff members or professionals need training? Does the clinic environment promote health?

- Make sure that service processes and systems are capable of producing the services:

What can the practice do?

2. Quality control (measure against goals)

- Evaluate quality:

How does the practice measure up in the different dimensions of quality?

3. Quality improvement

- Establish an infrastructure:

Information systems for monitoring quality and a team responsible for quality

- Identify projects:

What can the practice do? What would an alliance with other health services contribute? Could local authorities be persuaded to make healthy policy choices?

- Establish teams with enough resources and motivation to lead projects:

The teams will look at the causes of the problem, decide on a remedy, and decide what information is needed to know if the remedy has improved matters.

Instead of waiting for errors to happen and patients to be harmed before reviewing clinical practice, clinicians and health care institutions should be proactive in quality improvement. Quality improvement can be thought of as a continuous cycle that starts by examining needs, goes on to changing practice, checking to see if the planned changes have happened, checking to see if goals have been reached, and starting all over again. This should result in a continuous improvement in quality. The cycle is known as the Plan, Do, Study, Act (PDSA) cycle or the Shewart or Deming cycle, after two of the experts who developed it:

- *Plan:* Plan what you want to change and make measurements (using the dimensions of quality as a guide) to confirm the problem. Diagnose what can be improved and decide on the best way to make the change.

- *Do:* Implement the change.

- *Study:* Using the indicators in the "Plan" step of the cycle, measure the impact of the change. Look for positive, negative and unexpected impacts, because changes in complex systems can have negative side effects and unintended consequences.

- *Act:* Act on the results of the study. The net result should be more modifications and changes. Depending on the results, you then re-enter the "Plan" step again. (See Figure 13.3).

It is often better to do many mini-PDSA cycles rather than one big one. Clinicians do not usually have the resources to take on a major PDSA cycle. A sequence of successful small cycles is likely to achieve greater improvement in quality than a failed large one.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

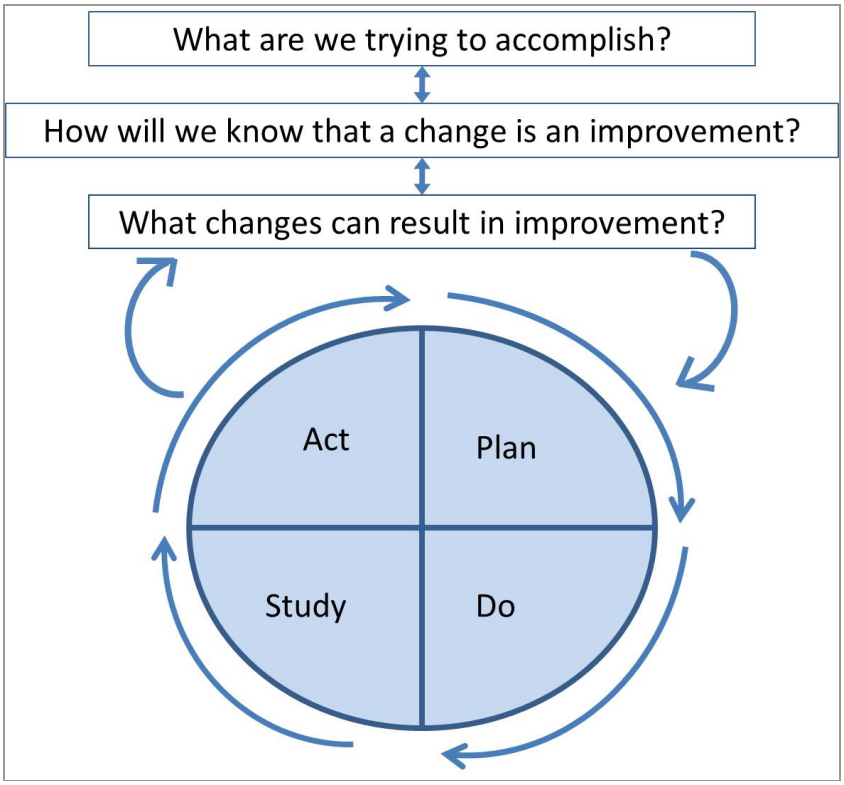


Figure 13.3: The Plan, Do, Study, Act cycle. (Adapted from the Institute for Health Care Improvement[18])

Nerd's Corner

Linking quality improvement to health improvement

The iterative PDSA cycle is similar to the measurement iterative loop proposed by Tugwell et al.[19] This is a framework in which information on the health needs of the population or practice population is presented in a logical sequence, from quantifying the burden of illness to identifying its likely causes to identifying interventions likely to prevent or reduce the illness and evaluating their efficiency. Once an intervention is identified as feasible and effective, it can be applied and the result evaluated. The cells of the framework can be adjusted for the change achieved, and the process can be started again.

Practical steps in quality improvement

The practical steps in conducting a quality improvement are as follows:

1. Decide what to try to improve, including purpose and scope.
2. Gather a team that includes those who will be affected by the likely changes.
3. Establish specific improvement objectives that are feasible and, if possible, measurable.
4. Establish measures or indicators of objectives, preferably ones that are easily available.
5. Choose what to change to make the improvements.
6. Evaluate the change.
7. Write a report and look at what else you might improve.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

Some leaders in quality improvement

Walter Shewart (1891?1967) is known as the father of statistical quality control. Having a strong scientific background, he adapted the scientific steps of observation, hypothesis generation, and hypothesis testing to quality management, where the steps became Plan, Do, and See what happens.

W. Edwards Deming (1900?1993) saw that quality depends on systems more than on individuals. In Japan, his work led to the rise of many very powerful Japanese companies, such as Sony, Panasonic, and Toyota. He generally aimed to reduce complexity, duplication, and waste. He saw variation as an opportunity for improvement, not as a problem.

Joseph M. Juran (1904?2008) trained as an electrical engineer and worked around the world as a management consultant. His particular contribution was to focus on the role of management in ensuring quality, reinforcing the idea that the system as a whole and its processes are at the base of most quality problems.

Avedis Donabedian (1919? 2000) practiced family medicine and was recruited into the School of Public Health of the University of Michigan. In 1966 he wrote the paper that introduced the concepts of structure, process and outcome, which is still the dominant paradigm for the evaluation of quality of health care and health care programme evaluation.

James T. Reason (b, 1938) started with an interest in controlling risks in the aviation industry. Reason has worked in a range of hazardous industries, including the health system. In particular, he focuses on the study of the safety and reliability of complex, well-defended hazardous systems.

Self-test questions

Self-Testing

1. What are the steps in the quality improvement cycle?
***Plan**: evaluate the current situation, prioritise what should be done, decide what can be done, decide what to do, and decide how to measure success. This step should be carried out in collaboration with the people that the action will affect.*
***Do**: carry out what was planned. Plans may need to be adjusted.*
***Study**: assess what was done, its success, its positive and negative impacts, and obstacles and facilitating factors.*
***Act**: act on the results of the "study" step and proceed to the next step, which can be a reassessment of the current action or a "plan" step to reassess how the current action has affected the overall situation and whether or not to change priorities and goals.*
2. What are the dimensions of quality?
Health care should be safe, effective, patient-sensitive, timely, efficient, equitable, and accessible. It should provide continuity of care. Health care professionals should have the skills and resources appropriate to their role in care. Health care should contribute to the health of the whole population. Finally, as the basis of excellent care is motivated personnel, those providing care should find their work rewarding.

Discussion Points

- 1. In your current attachment or practice, what do you think would improve the quality of care offered?
- 2. Develop a plan to address it.

References

Further Reading

The Institute for Healthcare Improvement provides brief, free online courses for students and many other resources for quality improvement:http://www.ihi.org/ihi

Reason J. Managing the risks of organisation accidents. Aldershot: Ashgate, 1997.

1. Baker GR, Norton PG, Flintoft V, Blais R, Brown A, Cox J, et al. The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada. CMAJ 2004 May 25;170(11):1678-1686.

2. Romanow RJ. Building on values: The future of health care in Canada. Privy Council of Canada publication number CP32-85/2002E-IN November; 2002.

3. Leatherman S, Sutherland K. Quality of healthcare in Canada: A chartbook. Ottawa, Ont.: CHSRF; 2010.

4. Leape LL, Lawthers AG, Brennan TA, Johnson WG. Preventing medical injury. QRB 1993;19(5):144-149.

5. Chassin MR, Galvin RW. The urgent need to improve health care quality. Institute of Medicine National Roundtable on Health Care Quality. JAMA 1998 Sep 16;280(11):1000-1005.

6. Rochon PA, Gurwitz JH. Optimising drug treatment for elderly people: the prescribing cascade. BMJ 1997 Oct 25;315(7115):1096-1099.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

7. Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. JAMA 1998 Apr 15;279(15):1200-1205.

8. Fick DM, Cooper JW, Wade WE, Waller JL, Maclean JR, Beers MH. Updating the Beers criteria for potentially inappropriate medication use in older adults: results of a US consensus panel of experts. Arch Intern Med 2003 Dec 8-22;163(22):2716-2724.

9. McLeod PJ, Huang AR, Tamblyn RM, Gayton DC. Defining inappropriate practices in prescribing for elderly people: a national consensus panel. CMAJ 1997 Feb 1;156(3):385-391.

10. Edwards IR, Aronson JK. Adverse drug reactions: definitions, diagnosis, and management. Lancet 2000 Oct 7;356(9237):1255-1259.

11. Institutes of Medicine. Crossing the quality chasm: A new health system for the 21st century. Washington: National Academy Press; 2001.

12. Ontario Health Quality Council. QMonitor: 2009 Report on Ontario's health system. 2009; Available at: http://www.ohqc.ca/pdfs/ohqc_2009_report_-_english.pdf. Accessed July, 2010.

13. Berwick DM. Continuous improvement as an ideal in health care. N Engl J Med 1989 Jan 5;320(1):53-56.

14. Reason JT, Carthey J, de Leval MR. Diagnosing "vulnerable system syndrome": an essential prerequisite to effective risk management. Quality in Health Care 2001 Dec;10(Suppl 2):21-25.

15. Reason JT. Managing the risks of organizational accidents. Aldershot: Ashgate; 1997.

16. Gray JA, Austoker J. Quality assurance in screening programmes. Br Med Bull 1998;54(4):983-992.

17. Juran JM. Juran on quality by design. New York: Free Press; 1992.

18. Institute for Healthcare Improvement. Improvement Methods. Available at: http://www.ihi.org/IHI/Topics/Improvement/ImprovementMethods/. Accessed February, 2010.

19. Tugwell P, Bennett KJ, Sackett DL, Haynes RB. The measurement iterative loop: a framework for the critical appraisal of need, benefits and costs of health interventions. J Chronic Dis 1985;38(4):339-351.

Chapter Objectives

After reading this chapter, you will be able to:

- public policy can influence population-wide patterns of behaviour and affect the health of a population.
- Identify ethical issues arising from restricting individual freedoms and rights for the benefit of the population as a whole (e.g., issues in designating non-smoking areas or restricting movements of person with active tuberculosis);
- Be familiar with economic evaluations such as cost-benefit / cost effectiveness analyses as well as issues involved with resource allocation.
- Describe at a basic level:
  - health resource allocation;

Public Policies

As described in Chapter 2, the factors that influence health are not solely individual or family characteristics but are linked to the economy, the political system, and to culture as a whole. As early as the 19th century, people recognized the link between politics and health. More than a century ago, Rudolph Virchow wrote that "medicine is a social science, and politics nothing but medicine on a grand scale."[1] One hundred years later, Michael Rachlis noted that "patterns of health and illness are reflections of our values, culture and institutions".[2]

Nerd's Corner

Rudolph Virchow

Dr. Rudolph Virchow (1821?1902), a Prussian, was Chair of Pathological Anatomy at the University of Berlin. He designed a sewerage system for the city of Berlin. During the Franco-German war, Dr. Virchow led the first hospital train to the front lines where he attended to injured soldiers. He is known as one of the fathers of modern pathology and one of the founders of social medicine. A number of entities are named after him, including Virchow's angle, Virchow's cell, Virchow's disease, Virchow's node, and Virchow's triad.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Definition

Policyrefers to a course or principle of action adopted or proposed by a government party, individual, or interest group.[3]

Politics refers to the art and science of government or to activities concerned with the acquisition and exercise of authority or power.[3] Politics has also been defined as the study of who gets what, when, and how.[4] The term encompasses more than the party political systems of national, provincial, and municipal government; it also refers to the struggle for power within any kind of organization, such as office politics, hospital politics, etc.

Politics refers to the debates, battles and negotiations that lead to **policies** being chosen and implemented.

Health policy issues can be classified into four major groups according to the level at which they affect the system:[5,6]

- *Systemic issues* concern the main features that shape the health system overall, such as the nature of public institutions involved in health care, the public/private mix, and the relationship between health and other sectors. An example would be, "who should pay for health care: the government or the person who uses it?"

- *Programmatic issues* concern the priorities for health care, the actual nature of health care programmes, and the way in which resources should be allocated. An example would be, "what types of services should the government pay for?"

- *Organizational issues* concern the way in which resources can be used productively and the ways to ensure a high quality service. An example would be, "how should we organize primary care centres to best utilize interprofessional care teams?"

- *Instrumental issues* concern the management of the various instruments of good organization, such as the human resource development and information systems. An example would be, "how can we best move to a unified system of electronic health records?"

At the same time, health policy issues can be classified by their likelihood of generating conflict among interest groups:

- *Highly political issues* are those likely to elicit a strong response from the public, physicians, nurses, or other stakeholders. Reductions of service are generally in this category. In Canada, discussions surrounding the central tenets of the Canada Health Act nearly always fall in this category, especially those surrounding the private and public payment for health services.

- *Low political issues* are those unlikely to elicit a great deal of response. Choosing to purchase rather than to lease cars for personnel might fall into this category.

Health Policy and Healthy Public Policy

Health policy dictates who should do what to whom, when they should do it, how much is available for doing, where they should do it, and who should pay for it. Health policy includes such things as the regulation and licensure of health care providers, the number of hospital beds in a community, arrangements for the insurance and payment of health services, the mix of public and private services, the variety, type and quality of services available in the community, and other issues such as access and cost of health services (see Chapter 12).

Health policies change according to the current major health problems and advances in technology and knowledge. Policy change is also governed by vested interests that can either obstruct or facilitate evidence-based policy proposals. Health policies are formed in a number of domains within the health care system. Some of these domains are shown in Table 14.1.

Table 14.1: Different domains of health policy decision-making

Policy domain	Type of question	Themes
Public policy	Do we fund transplants?	Issues and priorities
Administrative policy	Where do we locate transplant services?	Service delivery problems
Clinical policy	Who should receive transplants?	Intervention possibilities.

Healthy public policy concerns policies written in other sectors (such as education or transportation), which may have an important impact on health because they modify known social determinants of health. The intent is to ensure that this health impact is positive, and that policy makers include health considerations in their planning. For example, transportation policies should bear in mind the health benefits of **active transport**; school curriculum design should consider health issues in planning how much time to allocate to different activities; city zoning regulations should consider wind direction and air quality in deciding where to locate industrial complexes. Similarly, policies that affect social inequity, social exclusion and access to education can affect health.[7] In other words, improvements in **health policy** only cannot achieve health; a **healthy public policy** is also required. Healthy public policies are driven by broad social goals that evolve over time, as outlined in Table 14.2.

Table 14.2: Healthy public policy shifts in Canada[8]

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Policy goals	Basis	Actions
Public health measures (19th century)	Observation of the role of contamination in producing disease	Water management systems, food inspection, quarantine regulations
Universal health care coverage (mid 20th century)	Observation of the effects of the 1930s depression in Canada coupled with the belief that disease can be cured	Medical services focussing on individual and the body
Health promotion (late 20th century)	Observation of the role of the social environment in producing health	Funding of health promotion professionals and initiatives focussing on the social environment
Reduction of inequalities (not yet achieved)	Mid 20th century studies showing health gradient across socio-economic groups	Quebec's antipoverty law. Problem of "health imperialism" (health seen as the major public goal). Depends on other sectors (e.g. education) understanding their impact on health in order for change to occur.

Further Reading

For examples of important milestones in Canadian healthy public policy, check out the Canadian Public Health Association's ?12 Great Public Health Achievements' of the last 100 years: http://cpha100.ca/

Health impact assessment in relation to public policy is a developing field in public health that aims at producing evidence that can contribute to policy-making on questions that are as simple as the fluoridation of drinking water to questions as complex as transportation or urban planning policies.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Illustrative Materials

How an economic crisis may affect medicine

Economic crisis and medicine

Towards the end of 2008, over-inflated national economies began to collapse leading to a world-wide severe recession which lasted several years. A student-run forum in Queen's University School of Medicine examined the likely impact of the recession on medicine. The main themes that emerged from the discussion illustrate the broad potential impact that economic policies, the underlying cause of the recession, can have on medicine and health care.

Patients

Loss of work and job-opportunities leading to greater susceptibility to health problems and to poverty, a major health determinant.

Physicians and professionals

Greater obstacles to studying medicine, particularly for students from lower income backgrounds.

Threat to residency places, job opportunities and physician remuneration. Decreased research funding, as well as re-allocation of funds, for instance between prevention and curative interventions.

Health services

An economic recession produces an increased load on medical and social services. Financial crises in health often stimulate reform, leading to reformulation of health policy with new priorities being identified and new controls and regulations, such as increasing accountability for health care spending. Reformulation of health policy may be seen as a threat or as an opportunity for improvement. It can also be seized upon as an opportunity for promoting privately funded health care as a way to reduce pressure on the public system (see Chapter 12 for reasons why private funding does not reduce pressure on the public system)

Ethics

Threat to maintaining the value that health care is not a commodity. In our global market for health care services, economic pressures may drain trained care-givers from developing countries where they are most needed. In contrast, economic immigrants into rich countries may support the economies of their country of origin with the money they send home. The social responsibility of physicians may become more irksome and therefore less respected, leading to ethical questioning of privilege and obligation in the medical profession.

From Noyahr LA, Leung K, Uy P. The economic crisis and medicine forum: a guided inquiry model for exploring topics in medicine and society. Presented at Canadian Conference on Medical Education. St John's, Newfoundland, May 2010.

According to the 1988 WHO Adelaide recommendations,[9] **healthy public policy** is "characterized by explicit concern for health and equity in all areas of policy and by accountability for health impact. The main aim of health public policy is to create a supportive environment to enable people to lead healthy lives. Such a policy makes health choices possible or easier for citizens. It makes social and physical environments health-enhancing. In the pursuit of healthy public policy, government sectors concerned with agriculture, trade, education, industry, and communications need to take into account health as an essential factor when formulating policy. These sectors should be accountable for the health consequences of their policy decisions. They should pay as much attention to health as to economic considerations."

Here Be Dragons

Always look at the big picture

Health policy-making can be difficult. It should include consideration of a remarkably broad range of possible costs and benefits, and focusing on one aspect of the problem may lead to unexpected harmful consequences, as seen in the story of the Cairo pig farmers.

Pigs, swine flu, and garbage[10,11]

Until the 2009?2010 H1N1 (swine flu) pandemic, urban pigs flourished in Cairo. They fed on organic household waste, provided a steady income for their owners, and a cheap source of food for the 10% of the population whose religious beliefs allowed them to eat pork. In early 2009 the Egyptian government culled a million of these pigs in order to resolve the health risk posed by "disorderly pig-rearing" in cities. However, international news media reported that the cull was to prevent the spread of swine flu. The cull had a number of probably unforeseen consequences. First, the Egyptian government was derided for using an inappropriate measure to prevent the spread of swine flu. Second, the urban pig farmers, already a vulnerable, excluded group, lost their source of income. Finally, an excellent waste recycling service was removed, resulting in pressure on the city waste removal services which could not cope with the increase in volume. Organic waste piled up in the streets, causing a nuisance and possibly a greater health hazard than urban pig farming.

Influences on health policy-making

In the ideal health care scenario, patients would demand only what they need and the supply of care would meet all health care needs. There would be complete overlap of the need, demand and supply circles in Figure 14.1. However, demand is essentially a health behaviour. As such it is subject to all the influences described in models of health behaviour, including advice from professionals and others with vested interests in providing health

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

care. The identification of need, defined as the capacity to benefit from health care (see Need for care), is subject to the many differing perspectives on what need is and how it should be defined. For instance, after examining available evidence, in 2001 the Canadian Preventive Care Task Force concluded that mammography screening for breast cancer in women aged forty to forty-nine was not a need, although, using the same evidence, in 2002 the U.S. Task Force concluded that it was. Furthermore, with little new evidence, in 2009 the U.S. Task Force changed to recommending against routine screening in women under fifty. It is evident that both task forces were subject to influences that caused variation in their perspectives on the evidence, and thus on their perceptions of what the need was.

Further Reading

The politics of breast cancer screening

The saga of breast cancer screening for women ages 40 to 49 in the United States is perhaps an extreme example of how policy making occurs. For more information, see V. L. Ernster's "Mammography screening for women aged 40 through 49: A guidelines saga and a clarion call for informed decision making" in the *American Journal of Public Health*. 1997 Jul, 87(7):11036. (Available at: http://ajph.aphapublications.org/contents-by-date.0.shtml)

For a view of more recent controversy from those responsible for producing guidelines see: Woolf SH. The 2009 Breast Cancer Screening Recommendations of the US Preventive Services Task Force.JAMA. 2010;303(2):162-163.

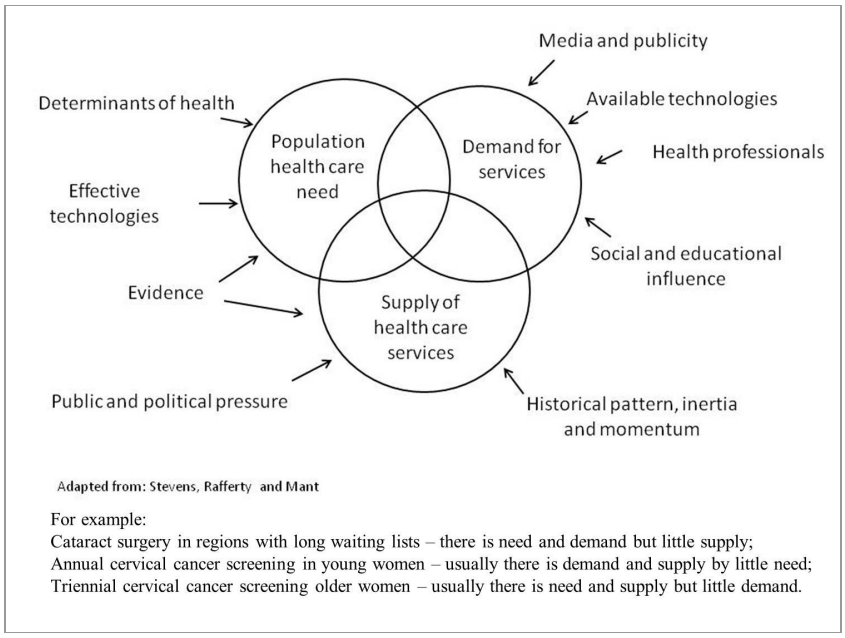


Figure 14.1: Factors influencing policy in health care[12]

There is money to be earned by supplying health care. In Canada, it accounts for about 20% of all government spending.[13] Each year, billions of dollars go to purchase all kinds of goods and services, including professional and non-professional services, high- and low-tech equipment, and pharmaceuticals.[14] Furthermore, other sectors are also interested in attracting government spending; education, welfare, infrastructure, among others, all have convincing reasons for money to come to them. As a result, many different groups are interested in the policies that govern spending on health and promote their agendas, sometimes quite aggressively. Although most agree that the patient's interests come first, generally the patient's interests are assessed from the perspective of an assessor who often has particular interests. Issues such as the public/private mix in health care, primary health care reform, and inter-professional scope of practice are sometimes deeply rooted in power struggles between professional groups, health care institutions, and other interested groups?even to the extent that the needs of patients and populations are obscured.[15]

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

A perspective on policy making[16]

Policy-making is a formal struggle over ideas, values and interests, played out by the rhetorical use of language and the enactment of social situations, much more than merely turning evidence into practice. Scientific evidence answers the question "What works?" Policy-making is about "What do we do?" Ostensibly, scientific research is about the objective establishment of facts; in fact, it is value-laden. The values of the researcher form the assumptions underlying the research question, as well as the interpretation of the findings. One only has to look at the letters page of a quality scientific journal to guess that there is no such thing as hard evidence?there are only competing constructions of evidence, which can support widely differing positions.

Policy-making is essentially about using judgment. In practice, it depends on what is said, by whom, and whether others find the arguments persuasive. Arguments are composed of *logos* (the facts and the reasoning), *pathos* (the emotional content), and *ethos* (the credibility of the speaker and the way the argument is presented). A persuasive argument accurately penetrates the audience with all three elements.

The policy cycle

Policy-making is a complex, involved, and continuous process. The history of tobacco illustrates how policies on tobacco production and consumption have changed since its arrival in Europe. It took nearly fifty years from the time Doll and Hill published a study showing the harmful effects of smoking tobacco[17] to the passing of effective legislation against it in Canada.

Many policy analysts use the **policy cycle** as a framework to understand the process of how policies come about.[18] The **policy cycle** describes how an issue moves from its initial inception through to implementation, evaluation and a new agenda. Table 14.3 summarizes the cycle; note the parallel between the policy cycle and the PDSA cycle of quality improvement.

Table 14.3: Problem solving and the policy cycle. From Howlett and Ramesh 1995[18]

Five stages of the policy cycle and their relationship to applied problem solving

Stages in policy cycle	Phases of applied problem solving	Description and comments
Agenda setting	Problem recognition	How an issue comes to the attention of policy makers. The process is not always rational, and it can often be difficult to see why some issues rise to the top of political agendas while other, seemingly more important issues, remain unaddressed[19]
Policy formulation	Proposal of solution	Decision-makers (governments, health regions, hospitals, care teams etc.) formulate policy options. Government policy-making usually occurs behind the scenes and is carried out by professional policy analysts[19]
Decision-making	Choice of solution	How decision-makers decide what to do?or not do?about an issue
Policy implementation	Putting solution into effect	Putting the decisions into effect. Not as simple as it sounds, as it usually entails changing habits and ingrained ways of doing things
Policy evaluation.	Monitoring results	(all too often neglected) Examining implementation and outcomes to check if the policy has been properly implemented and if the desired outcomes were achieved[18,20]

It is tempting to think of policy-making as an ordered process that moves forward, logically, through each step of the cycle. However, the real world is not so rational. Sometimes policies are formed without consideration of possibly better alternatives or without formulation of the problem. The reality is that the process can begin at any of the steps and does not always move in logical sequence through all the steps. Those who want to change policy must always be on the look-out for opportunities and influences that will help advance their position. Achieving policy change generally requires more tenaciousness, patience and persuasive argument than scientific evidence.

In some cases, although evidence is used to inform the policy-making process, budgets and the feasibility of implementation dictate choices. Even if budgets and implementation are not obstacles to acting according to evidence, the questions asked of the evidence and how the replies are interpreted can result in widely different viewpoints.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Links

Politics of prevention

For the story of petroleum companies opposing the banning of lead in gasoline

visit the University of Ottawa website at: http://www.medicine.uottawa.ca/sim/data/Lead_in_Gasoline.htm

For the history of Canadian tobacco policy, please visit the National Collaborating Centre for Healthy Public Policy website at: http://www.ncchpp.ca/515/Timeline-Tobacco+story.htm

Nerd's Corner

Industry: a political force

Industry succeeds in selling products long after they are shown to be a probable hazard to health. Here are some of the methods they use:[21]

1. Public Relations

- *Express a concern for the health of the users of the product.* In 1954, the US tobacco industries produced a statement saying that their top executives accept an interest in people's health as their basic responsibility. Obviously, people with such an interest would never sell a harmful product.
- *Stress personal responsibility and the freedom of the individual.* Industry claims to provide choice for individuals: "Just because we have electricity doesn't mean that you have to electrocute yourself." Of course, industry claims, advertising simply informs.
- *Funding civic activities and demonstrating social responsibility.* In 2000, the Philip Morris tobacco company spent $115 million on social causes in the U.S. and a further $150 million to publicize its beneficence. This $265 million was 1.7% of the company's domestic tobacco revenues.

2. Influence government and key organizations

- *Election campaign contributions.* Some companies contribute funds to all the major parties in order to buy favour with the candidate who wins the election.
- *Lobbying.* Many companies retain the services of lobbyists who plead the company's cause to those in power.
- *Revolving door between private organizations and public bodies and government.* Industry may hire scientists as advisors in an attempt to colour the interpretation of results of scientific study. Jobs can be promised to members of government.
- *Funding for "grass roots" groups.* Some companies fund pressure groups that masquerade as consumer groups.

3. Create doubt and influence the results of scientific enquiry

- *Dispute generally accepted scientific results.* Harshly criticize studies that find against their interests and discredit the source of the study.
- *Fund scientific projects to produce good will and influence reporting of results.* Studies funded by industry are more likely to report results in favour of the industry than studies not funded by industry.

4. Product marketing

- *Target youth.* Influencing the habits of young people can ensure that products are consumed for a lifetime.
- *Product placement.* The common method is for the hero of a movie or television show to be seen using the product on screen. Other forms would include the presence of fast food, junk food, and soft drinks in schools or hospitals, which are assumed to endorse the products sold on their premises.
- *Offer "safer" versions of the product.* "Light" cigarettes are generally only light when smoked by the machines that measure nicotine and tar content. When humans smoke them, the smoker's fingers block the ventilation holes. Low fat cereal alternatives often have higher sugar content than the original products.
- *Create addiction.* The tobacco industry has modified nicotine levels to make their products more addictive.

However, industrial interests can also benefit health, the following is an example:

Suncor ? a major industry employer in the Fort McMurray area of Alberta, has been extracting petroleum products from the surrounding oil sands since the 1960s. Approximately half of its 12,000 employees are in the Fort McMurray. Within this, largely isolated, environment, health and safety are important to the industry and to the population. To address worker safety, mandatory drug/alcohol screening takes place along with third party transport to work (bus pickup and delivery). Family health is addressed by a well-supported employee and family assistance program.

To improve health and wellness among workers, families and the community, Suncor has sponsored a Community leisure facility for use by all in the city of 100,000. This facility is well used by people of all ages, and encompasses programmes directed towards families. http://www.macdonaldisland.ca/

It is understood that the underlying motives of these actions may be to entice workers and families to stay within the municipality. Besides this intended outcome, the use of the leisure facility also benefits individual and community health. Linkages between communities and industry are possible, and can be beneficial for the worker, family and municipality.

The budget cycle

Whereas the policy cycle is sometimes almost pure theory, the **budget cycle** is concrete. The government's year is highly structured and budget-making follows a strict schedule. Although the schedule varies from province to province, from municipal to federal governments, and from year to year, there is a fairly predictable pattern. A typical budget cycle is illustrated in Figure 14.2. Policies that require a reallocation of budget must be adopted by early to mid-autumn if they are to be implemented the following year. Early in the New Year, the budget is, for all intents and purposes,

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

finalized, although it is not made public until March.

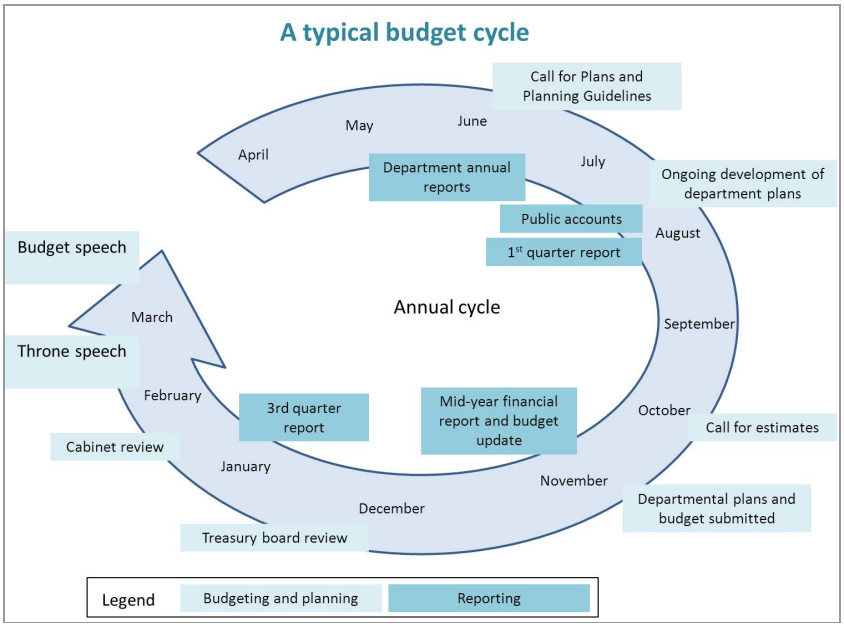


Figure 14.2. A Canadian budget cycle[22]

Welfare economics and cost benefit analysis

Welfare economics holds that, whereas social and economic decisions ought to be controlled by the market rather than by regulation, there are certain situations in which the market fails to distribute goods in an equitable manner. Nevertheless, if markets fail to distribute goods equitably, as in the case of free-market health care, distribution should be regulated.[18] Economic evaluation of policy options can aid decision-making on how to distribute goods. It is a way of comparing different interventions to decide which provide the greatest potential impact on patients and on society. It should assess all possible ways of resolving the problem under consideration so that resources can be allocated to the actions most likely to enhance the quality of people's lives and to give the greatest benefits.[23] (See chapter 7)

Definition

Economic analysis is the "the comparative analysis of alternative courses of action in terms of both their costs and their consequences."[24]

Illustrative Materials

Two examples of economic evaluations

A cost-utility analysis

In many Canadian provinces, at least two doses of hepatitis B vaccine are given to nine-to-eleven year-olds (the age depends on the province). The hepatitis A vaccine is given only to those at risk of the disease. A study examined the cost-utility of adding Hepatitis A vaccine to the doses of hepatitis B vaccine that nine-to-eleven year-olds currently get. It found that the bivalent (hepatitis A and hepatitis B) vaccine, even though it incurs additional health costs, is cost-saving from a societal perspective, which includes consideration of the gain of **quality-adjusted life years** (QALY). The net benefit of using the bivalent vaccine in all nine-to-eleven ten year-olds instead of the monovalent hepatitis B vaccine would be an additional 49 QALYs for Canada.[25]

A cost-benefit analysis

Reducing dietary sodium by 1,840 mg/day (as opposed to not reducing it) would decrease the prevalence of hypertension by 30%, resulting in one million fewer hypertensive patients in Canada. Direct savings to the Canadian health care system would be about $430 million per year because of the reduction of costs due to reduced use of medications, laboratories and physician visits.[26]

Although economic analysis provides ostensibly hard evidence on which to base decisions, results depend on which costs, harms, and benefits are considered, as well as to whom they accrue. Costs can be monetary and non-monetary, such as loss of a limb, loss of a life-year or psychological suffering. There are significant methodological questions over how best to value these non-monetary costs as well as how to compare different health and social outcomes.[24,27-31]

Public health ethics

Public health ethics become relevant to clinicians when they plan projects directed at their practice population?for example, implementation of screening guidelines or vaccination programmes. They can also impinge on the clinical ethical considerations in situations when the health of the population is at risk. Public health ethics can be distinguished from bioethics by differences in emphasis:

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

- Population focus vs. focus on individual
- Community perspective vs. focus on the person
- Social determinants vs. individual agency and responsibility
- Systems of practice vs. individual decision-making
- Distribution of resources vs. patient care.

The central values of clinical ethics, such as autonomy, individual liberty, beneficence and non-malfeasance, relate to the interaction between clinicians and patients. Public health, focussing on the community, has a wider range of values, including:[32,33]

- Solidarity
- Social justice and equity
- The protection of collective interests
- Notions of community (respect for community)
- Common goods and public goods
- General societal well-being and human flourishing
- Reciprocity
- Public trust.

Several frameworks have been developed to assess the ethics of public health action;[33-37] each one highlights a particular perspective. As public health develops, its ethical frameworks are also likely to continue developing. Ethical frameworks help identify the issues to be considered before deciding on an action. They do not identify the ?right' decision;[38] they are just ?frames,' a way of looking at a problem. They aim to assist the decision-maker, but judgment on the decision-maker's part remains essential.

Ethical issues relevant to public health actions

Justifiability: What are the goals of the action? The goals should be publicly justifiable.

Transparency: Is the decision-making process transparent? All those affected by the action should have input into the decision.

Effectiveness: Is the action capable of achieving the goals? "The only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others. His own good, either physical or moral, is not a sufficient warrant" (John Stuart Mill).

Necessity: Must one intervene? Is there another way of achieving the same goals? If there is more than one way, that which causes the least harm should be chosen.

Proportionality: Do the benefits outweigh the harms that the action causes? In the SARS outbreak, civil liberties had to be balanced against the need to quarantine; the protection of health workers had to be balanced against the duty to provide care, and the access to care had to be balanced against the need to reduce spread.

Minimization of harm: Can the harm caused by the action be reduced? Harm includes infringement on general moral principles. For instance, if it is necessary to break patient confidentiality to protect the population, only essential information should be communicated, and those responsible for gathering that information should ensure its security.

Reciprocity: Can individuals be compensated for the harm the action causes? For instance, individuals who suffer secondary effects of vaccination should have free access to the care and rehabilitation they need.

Fair balance of benefits and harm: Is it likely that one sector of the community will benefit while another is harmed, or the harms distributed equally? For instance, people living near a polluting factory are likely to derive health benefits if the factory is closed, but those working in it will suffer the harm of job loss.

Fair implementation: Will the action benefit all sectors of the community? Health campaigns aimed at a general population run the risk of reaching only the best informed section of society that is least in need of them.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Nerd's Corner

Ethics beyond humans[37]

Some recent frameworks, developed to examine global and international moral challenges, include respect for all life and for the environment. Indeed, experts are beginning to call for an expansion of the discourse on ethics from interpersonal relationships to the ethics of relationships between institutions and even to the ethics of relationships between nations. This will require promotion of a deeper understanding of citizenship in an interdependent world, commitment to an extended range of human rights, and new ways of thinking about ourselves, our relationship to others, and to the ecological system. In addition, human rights should be linked to a broader moral agenda embracing the duty to meet essential human needs and to achieve greater social justice within and between nations. It is from this springboard that we move to an expanded discussion of public health ethics and its implications at the international and global levels.

Public health ethics and clinical practice

Certain situations arise in public health that require clinicians to break the code of ethics that governs the care of the individual:

- Reporting a notifiable disease requires a breach of patient confidentiality that, generally, does not benefit the patient. It is required by law in order to assist in the monitoring and control of infectious disease with the aim of benefiting the public.
- Cooperating in epidemiological investigations or intervening to control the spread of disease may also require the physician to carry out an intervention that is of no benefit to the patient.
- Isolation and quarantine restrict patients' freedom.
- Prophylaxis or vaccination of contacts to prevent the spread of infectious disease may not greatly reduce the contact's risk of disease, so the benefit to the contact may not outweigh his risk.

Although public health legislation in the Canadian provinces allows Medical Officers of Health to transgress certain individuals' rights in order to protect the population from illness, the harm to individuals should always be minimized. Reports of notifiable disease are confidential and should be kept under secure conditions. Treatment or isolation of a patient against his will requires legal intervention; the Medical Officer has to present the case before a judge or a magistrate.

A major ethical question for all clinicians concerns how to balance their duty to care for the patient with their duty, as a manager of scarce resources, to ensure equity in allocation of resources. Most health care spending results from decisions made by individual clinicians with individual patients. Inefficient or ineffective clinical intervention in one patient wastes resources that could be used to treat another patient or, indeed, used in another sector?education, housing or employment.

Nerd's Corner

Some ethical perspectives relevant to public health[39]

Utilitarianism

Jeremy Bentham (1748?1832) argued that the rightness of an action is not intrinsic but is determined by the "hedonic calculus" of adding up the pleasure and pain it produces. Translated to public health, policy or action should be judged by its effect on the sum of individual levels of well- or ill-being it produces. However, the measurement of well-being, the outcome of the action, is subject to discussion: whether individuals' perceptions of their own well-being or objective measures should be used; how to weigh ?apples against oranges' (as an example, can dyslexia be compared to the loss of a foot?); or the relative value of harm at different ages. The utilitarian approach is often used in public health, for instance a childhood vaccination programme entails vaccine reactions in some children in order to reduce the larger number of children who would be harmed by the disease, or the quarantine of a person with infectious disease entails restriction of that person's freedom in order to prevent the harm to others that transmission of the disease would cause.

Liberalism

Immanuel Kant (1724?1804) held that human beings ought to be treated with respect, as ends in themselves and not as means to another individual's ends. The approach promotes individual freedom of choice; however, how to achieve it is a matter of debate. Some want minimal government intervention to protect personal liberty. Others argue that the right to choice is meaningless without adequate resources to enable choice, so that health care must be made affordable. Once this happens, health care can be matter of choice. There is also debate over whether health is a prerequisite for choice or a consequence of it. If it is a consequence of choice, governments should provide access to health care and education on how people can improve their health, but allow individuals to decide on whether or not to use it. If it is a prerequisite for choice, governments can be more aggressive, using legislation to coerce people into healthy behaviours, such as using seat belts or not smoking.

Communitarianism

This perspective has a long history. It is found in among the Greek philosophers Plato and Aristotle, as well as in many non-western traditions. It focuses on the creation of a good society. Knowing one's place and fulfilling one's duties are important virtues. The major question of this perspective is: Who decides what is virtuous? One view is that every community defines its own norms so that morality is contextualized; each society must respect the norms of the other. Others believe in a single true form of good society and its associated virtues. With this belief, it is possible to promote cultural patterns, such a female literacy, in all societies. Communitarianism also leads to questions about how much coercion may be used to limit deviancy from norms. Finally, it raises questions about how to define ?a community', that is, when can an individual or group be allowed to opt out to follow a vision different from that of the main community?

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

Further Reading

To read more about public health ethics: Bailey TM, Caulfield T, Ries NM (eds). Public health law and policy in Canada. 2nd ed. Markham (ON): LexisNexis Canada, 2008. Chapter 2 covers public health ethics: pp. 37-59.

Self-test questions

Self-Testing

1. What are the main steps in policy-making?
The main steps of policy-making in theory are:
Agenda-setting *during which priorities are identified,*
Policy formulation *during which policy options are identified,*
Decision-making *to select the preferred policy option,*
Policy implementation *during which policies are put into effect,*
Policy evaluation *during which the implementation and the effect of the policy is assessed.*
In practice, although these steps are useful in analysing how policies are made and implemented, they are rarely followed in an orderly manner.
2. What is the different between policy and politics?
Politics is the art and science of government, or the activities concerned with the acquisition or exercise of authority or power. Policy refers to a course or principle of action. In other words, politics are what makes policy.
3. Define the term healthy public policy.
Healthy public policy includes all policies that are likely to impact health, whether or not health is their primary aim. They include policies relating to transport, the environment, the social system, the economy, and the education system, among many others.
4. What are the main differences between clinical ethics and public health ethics?
The outstanding characteristic of clinical ethics is the unequal power relationship between the patient and the professional. The outstanding feature of public health ethics is the dilemma created by weighing the rights of different individuals in populations. The two are distinguished by virtue of their different emphasis.

Emphasis in public health	Emphasis in clinical medicine
Population focus	Individual focus
Community perspective	Personal perspective
Social determinants	Individual agency and responsibility
Systems of practice	Individual decision-making

The central values of clinical ethics relate to interaction between professionals and patients and deal with issues such as autonomy, individual liberty, beneficence and non-malfeasance. Public health values are broader and include solidarity, social justice and equity, protection of collective interests, common and public goods, societal well-being, reciprocity and public trust.

Discussion Points

- 1. A patient suffering from multiple sclerosis has requested that you write a letter of support so that he can get venous decongestion at the expense of the province. The promoters of this treatment claim that it improves the symptoms to the extent that patients can regain lost functions. The medical establishment, however, considers that the scientific basis of the treatment is lacking as is evidence of its effectiveness. Discuss the ethical issues involved in your decision.
- 2. What are the main health care priorities in the region in which you work or study and what were the criteria for identifying them as priorities?
- 3. Describe a policy in the municipality where you live that affects your health.

References

1. Ackerknecht EA. Rudolph Virchow: Doctor, Statesman, Anthropologist. Madison: University of Wisconsin Press; 1953.

2. Rachlis M. Prescription for excellence: how innovation is saving Canada's health care system. Toronto: Harper Collins; 2004.

3. Barber K editor. Canadian Oxford Dictionary. Toronto & New York: Oxford University Press; 2000.

4. Lasswell H. World politics and Personal Insecurity. New York: McGraw-Hill; 1935.

5. Barker C. The health care policy process. London: Sage Publications; 1996.

6. Frenk J. Dimensions of health system reform. Health Policy 1994 Jan 31;27(1):19-34.

7. Terris M. The complex tasks of the second epidemiologic revolution: the Joseph W. Mountin lecture. J Public Health Policy 1983 Mar;4(1):8-24.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

8. Glouberman S, Kisilevsky S, Groff P, Nicholson C. Towards a New Concept of Health: Three Discussion Papers. CPRN Discussion Paper No. H|03. Ottawa, Ont.: Canadian Policy Research Networks; 2000.

9. World Health Organization. Adelaide Recommendations on Healthy Public Policy. 1988; Available at: http://www.who.int/healthpromotion/conferences/previous/adelaide/en/index.html. Accessed July, 2010.

10. Slackman M. Belatedly, Egypt spots flaws in wiping out pigs. 2009; Available at: http://www.nytimes.com/2009/09/20/world/africa/20cairo.html. Accessed July, 2010.

11. Paddock C. Egypt pig cull not a swine flu measure says government. 2009; Available at: http://www.medicalnewstoday.com/articles/148430.php. Accessed July, 2010.

12. Stevens A, Rafferty J, Mant J. The epidemiological approach to needs assessment. In: Stevens A, Rafferty J, editors. Health care needs assessment Oxford: Radcliffe Medical Press; 1994.

13. Statistics Canada. Consolidated federal, provincial, territorial and local government revenue and expenditures 2005 to 2009. Available at: http://www40.statcan.gc.ca/l01/cst01/govt01b-eng.htm. Accessed February, 2010.

14. Canadian Institute for Health Information. National Health Expenditure Trends 1975 to 2009. 2009; Available at: http://secure.cihi.ca/cihiweb/products/National_health_expenditure_trends_1975_to_2009_en.pdf.

15. Rachlis M, Kushner C. Second Opinion: What's Wrong with Canada's Health Care System. Toronto: Collings; 1989.

16. Russell J, Greenhalgh T, Byrne E, McDonnell J. Recognizing rhetoric in health care policy analysis. Journal of Health Services & Research Policy 2008 Jan;13(1):40-46.

17. Doll R, Hill AB. Smoking and carcinoma of the lung; preliminary report. Br Med J 1950 Sep 30;2(4682):739-748.

18. Howlett M, Ramesh M. Studying Public Policy: Policy Cycles and Policy Subsystems. Toronto: Oxford University Press; 1995.

19. Kingdon JW. Agendas, Alternatives, and Public Policies. 2nd ed. New York: Addison-Wesley Educational Publishers Inc.; 2003.

20. Schofield J, Fershau J. Committees inside Canadian Legislatures. In: Dobuzinskis L, Howlett M, Laycock D, editors. Policy Analysis in Canada: The State of the Art. Toronto: University of Toronto Press; 2007. p. 351-374.

21. Brownell KD, Warner KE. The perils of ignoring history: Big Tobacco played dirty and millions died. How similar is Big Food?. Milbank Q 2009 Mar;87(1):259-294.

22. Saskatchewan Finance. Planning and Budget Cycle. Regina: Government of Saskatchewan.; 2005.

23. Brent RJ editor. Cost-benefit analysis and health care evaluations. Cheltenham, UK; Northampton, MA.: Edward Elgar; 2003.

24. How to read clinical journals: VII. To understand an economic evaluation (part A). Can Med Assoc J 1984 Jun 1;130(11):1428-1434.

25. Bauch CT, Anonychuk AM, Pham BZ, Gilca V, Duval B, Krahn MD. Cost-utility of universal hepatitis A vaccination in Canada. Vaccine 2007 Dec 12;25(51):8536-8548.

26. Joffres MR, Campbell NR, Manns B, Tu K. Estimate of the benefits of a population-based reduction in dietary sodium additives on hypertension and its related health care costs in Canada. Can J Cardiol 2007 May 1;23(6):437-443.

27. How to read clinical journals: VII. To understand an economic evaluation (part B). Can Med Assoc J 1984 Jun 15;130(12):1542-1549.

28. Dolan P. The measurement of individual utility and social welfare. J Health Econ 1998 Jan;17(1):39-52.

29. Dolan P, Stalmeier P. The validity of time trade-off values in calculating QALYs: constant proportional time trade-off versus the proportional heuristic. J Health Econ 2003 May;22(3):445-458.

30. Henderson JW. Health Economics and Policy. Second ed. Cincinnati: South Western Thomson; 2002.

31. Sharpe C. How to conduct a cost-benefit analysis. Alexandria, Va: ASTD; 1998.

32. Dawson A. Resetting the parameters: Public health as the foundation for public health ethics. In: Dawson A, editor. Public Health Ethics: Key Concepts and Issues in Policy and Practice. Cambridge: Cambridge University Press; 2009.

33. Upshur RE. Principles for the justification of public health intervention. Canadian Journal of Public Health.Revue Canadienne de Sante Publique 2002;93(2):101-103.

34. Kass NE. An ethics framework for public health. Am J Public Health 2001 Nov;91(11):1776-1782.

35. Childress JF, Faden RR, Gaare RD, Gostin LO, Kahn J, Bonnie RJ, et al. Public health ethics: mapping the terrain. Journal of Law, Medicine & Ethics 2002;30(2):170-178.

36. Singer PA, Benatar SR, Bernstein M, Daar AS, Dickens BM, MacRae SK, et al. Ethics and SARS: lessons from Toronto. BMJ 2003 Dec 6;327(7427):1342-1344.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse

37. Benatar SR, Daar AS, Singer PA. Global Health Ethics: The Rationale for Mutual Caring. International Affairs 2003;79(1):107-138.

38. Dawson A. Theory and practice in public health ethics: a complex relationship. In: Peckham S, Hann A, editors. Public Health Ethics & Practice. London: Polity Press; 2009.

39. Roberts MJ, Reich MR. Ethical analysis in public health. Lancet 2002 Mar 23;359(9311):1055-1059.

Copyright © The Association of Faculties of Medicine of Canada. Content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ For permissions beyond the scope of this license, visit http://www.afmc-phprimer.ca/termsofuse