PART 1

General principles of cancer diagnosis and management

1 Cancer epidemiology

Rajesh Dikshit[1] and Preetha Rajaraman[2]

[1]Tata Memorial Centre, Centre for Cancer Epidemiology, Mumbai, India
[2]Centre for Global Health; Division of Cancer Epidemiology and Genetics, U.S. National Cancer Institute, New Delhi, India

Introduction

Epidemiology is the study of the distribution and determinants of disease in specified populations, and the application of this information to the control of health-related problems. Cancer epidemiology thus encompasses understanding the distribution of cancer morbidity and mortality, identifying the causes of cancer and evaluating preventive measures and use of health services.

Distribution of disease refers to the identification, description and interpretation of the patterns of cancer among different populations over different time periods. This branch of cancer epidemiology, often referred to as *descriptive cancer epidemiology*, has provided insights into the disease burden of cancer and trends of cancer over time, and has also helped in generating hypotheses for aetiological research. Key to accurate descriptive cancer epidemiology is the availability of high-quality population-based cancer registries in many areas of the world.

The term *determinant of disease* refers to the study of disease aetiology. Knowledge about the causes and preventive strategies for cancer has largely arisen from carefully conducted epidemiological studies, often referred to as *analytical epidemiology*. The search for causes in an epidemiological setting is not limited to lifestyle factors, but also includes infectious agents and genetic factors. Over the past few decades, the field of epidemiology has evolved with the use of biomarkers, including genetic markers, to deepen our understanding of both exposure and outcome. This particular approach to understanding disease aetiology is often termed *molecular epidemiology*.

Along with the identification of the causes of cancer, epidemiological studies have been used to evaluate the success of primary and secondary prevention strategies in controlling the burden of cancer. These studies are often conducted as *field intervention trials* and assess the feasibility and success of primary preventive measures and screening programmes in different population settings.

Study designs in cancer epidemiology (Box 1.1 and Table 1.1)

Descriptive studies

A well-functioning cancer surveillance system which provides accurate information on cancer incidence, mortality and time trends is a key feature of an effective cancer control programme. The task of cancer surveillance is undertaken by population-based cancer registries (PBCRs). These registries collect cancer data by age, cancer site and date of diagnosis

Box 1.1 Introduction to cancer epidemiology

- Descriptive cancer epidemiology is the study of the distribution of cancer in different geographical areas over different time periods. This information is mainly derived from population-based cancer registries (PBCRs)
- Cancer epidemiology is also concerned with identifying the causes of cancer: this is mainly achieved through observational study designs such as case–control and cohort studies
- The branch of epidemiology that uses biomarkers of exposure and disease to understand the aetiology of cancer is referred to as 'molecular epidemiology'
- Since cancer is a relatively rare disease, case–control studies are often the study design of choice. Given the potential for selection and recall bias, these studies should be conducted with utmost care to minimize bias for meaningful interpretation of study results
- Although expensive and requiring longer follow-up, cohort or longitudinal studies are powerful study designs to understand the aetiology of cancer
- Experimental studies which involve randomization of individuals into two or more study groups are commonly used to study treatment efficacy ('clinical trials'). This design can also be used in the field to study the effectiveness of interventions such as vaccination (by randomization of individuals into vaccinated and non-vaccinated groups)

UICC Manual of Clinical Oncology, Ninth Edition. Edited by Brian O'Sullivan, James D. Brierley, Anil K. D'Cruz, Martin F. Fey, Raphael Pollock, Jan B. Vermorken and Shao Hui Huang.

Table 1.1 Summary of the scope of cancer epidemiology

Main goal	Approach	Statistical measures
Distribution of cancer	Population-based cancer registries (PBCRs)	Incidence rates, mortality rates, cumulative risk
Determinants of cancer (lifestyle, environmental, infection, genetic)	Case–control studies, cohort studies, molecular epidemiology (including genetic markers)	Odds ratio, relative risk, attributable risk
Public health (screening, primary prevention measures)	Field intervention trials	Hazard ratio, mortality ratio

for populations in well-defined geographical areas. PBCRs can also provide data on population-based survival for different cancer sites in different populations. The quality and completeness of the data depends upon the availability and utilization of health services by cancer patients, and proper documentation by various facilities in which cancer is diagnosed and treated. Key requirements for a well-functioning PBCR are: accurate census data for the population covered by the registry by age and gender; access to various sources of cancer diagnosis and treatment; and support from local policy-makers as well as leading cancer diagnostic and treatment centres. While the numbers of high-quality PBCRs have been increasing worldwide, the large majority is still located in high-income countries. In the absence of PBCRs, many countries depend upon data from hospital-based cancer registries (HBCRs) and pathology-based cancer registries to estimate cancer burden. Data from these registries, however, are not suitable for cancer control planning, given the potential biases due to referral patterns and underestimation of incidence for cancer sites where histology is uncommon. In order to promote descriptive epidemiology research and increase the number of cancer registries in low resource settings, the International Agency for Research on Cancer (IARC), in collaboration with a number of global partners, including UICC, the U.S. National Cancer Institute and Centers for Disease Control and Prevention (CDC), has launched the Global Initiative for Cancer Registry Development' (GICR; www.gicr.iarc.fr). The GICR functions with the help of IARC regional hubs which provide assistance in the establishment of cancer registries.

Analytical studies

Analytical studies fall into two main categories: *observational* and *experimental*.

Observational studies

An observational study is a non-interventional investigation of disease causation in a human population. Such studies are based on observing associations between the exposure(s) and disease(s) of interest. Measurement of exposure is usually based on some combination of lifestyle data from questionnaires, external monitoring of exposure (e.g. for air pollutants) and biomarkers of exposure. New advances in microchip technologies and informatics are being used to understand the role of genetics as well as the interaction between environmental exposures and genes. Observational studies yield measures of association between exposure and disease, but interpretation of causality requires further information, including the following considerations:

- *Temporality*. Does exposure precede disease? For a factor to be causal for a disease, it must occur before the disease.
- *Strength* of the association, measured by the relative risk or odds ratio. The stronger the association, the more likely that the relationship is causal.
- Existence of a *dose–response* relationship. If the risk of disease increases with exposure dose, this provides further evidence for causality.
- *Replication* of findings. If an association is observed in various studies and settings, this provides further support for causality.
- *Biological plausibility* of the association. A relationship is more likely to be causal if the existing biological literature supports the finding.
- *Ruling out alternative explanations*. Alternative explanations for the observed results should be considered to rule out the possibility of spurious associations due to confounding or bias (see below).

As individuals are not randomly assigned to exposure groups in observational study settings, these designs can lead to non-causal associations, particularly due to:

- Confounding
- Selection bias
- Misclassification.

Confounding occurs when a variable that is not part of the disease causal pathway is associated with both disease and outcome. Confounding can be addressed by appropriate study design, data collection and analysis. In order to statistically control for confounders during analysis, it is essential to obtain information on potential confounding variables during data collection. For example, in a study measuring the effect of alcohol intake on lung cancer, results could be confounded by smoking, as smoking is associated with both the exposure under investigation (alcohol intake) and is also independently a risk factor for the disease (lung cancer); therefore, the confounding effect of smoking needs to be addressed either at the design stage (e.g. matching for smoking status or restricting the study to non-smokers) or at the analysis stage (by adjusting for smoking information in statistical models).

Most observational studies rely on data collected from accurate reporting of information, e.g. by study participants, physician records or laboratory procedures. Errors in classification of exposure or disease can occur if this information is not properly

provided or recorded. This type of bias is called *misclassification bias* or *information bias*. In case–control studies particularly, the exposure or disease frequencies among study participants may not be representative of the target population, resulting in *selection bias*, which can produce inaccurate measures of association.

Careful interpretation of results from observational studies should consider the study design (including selection of cases and controls), potential biases and confounding to rule out alternative explanations. In order to conclude if the observed association is causal, further considerations include the strength of association, temporal relation between exposure and disease, dose–response gradient, biological coherence and consistency of results across studies.

Common study designs for analytical observational studies include *cohort* and *case–control* studies.

Cohort studies

In a cohort study, a group of individuals free of the disease(s) of interest is enrolled and followed up to ascertain different endpoints such as premalignant conditions, occurrence of cancer or death. Exposure measurements are ideally collected at the time of enrolment (prospective cohort study), but can also be collected in subsequent questionnaires or be historically reconstructed (retrospective cohort study). Disease risk is then compared between groups classified based on their exposure. Cohort studies resemble clinical trials in that both study designs compare disease risk between exposure groups. However, since the allotment of exposure is based on natural variation between the groups rather than random allocation by the investigator, more care needs to be taken in interpreting observed associations. Cohort studies allow estimation of the *incidence rate* (the instantaneous rate of occurrence of new disease events) as well as the *cumulative risk* (the cumulative probability of the disease during a given time interval). The ratio of the incidence rates in groups based on different categories of exposure is termed the *relative risk*. While cohort studies are very effective in determining disease aetiology, they can be expensive and difficult to implement logistically as they require long-term follow-up to obtain disease endpoints, particularly for rare diseases like cancer.

Case–control studies

To investigate aetiological factors for relatively rare diseases, the case–control study is often the design of choice for reasons of speed and efficiency. In a study to investigate cancer aetiology, individuals diagnosed with the cancer of interest are recruited from a defined population in a defined time period. A similar group of cancer-free individuals is sampled as a 'control' from the same study base from which the cases arise. The distribution of exposure among cases is then compared with that among controls, and the *odds ratio*, which approximates the relative risk when the disease prevalence is low, is computed as a measure to identify the strength of association between the exposure and disease. Case–control studies appear easy to conduct, but can give misleading results if cases and controls are not properly selected (leading to *selection bias*), or if the information on exposure is not properly collected, e.g. because of poor questionnaire design or administration, or improper collection, storage or analysis of biomarkers (leading to *exposure misclassification*). Additionally, if cases are more likely to report or recall exposure to a given factor than controls, this can lead to spurious results due to *reporting* or *recall* bias.

Experimental studies

In experimental studies, the investigator randomly allocates study subjects to exposure or no exposure. Study subjects are then followed up to observe the outcome(s) of interest. The random allocation makes these studies less susceptible to many of the biases that can be present in observational studies. Nonetheless, experimental studies are susceptible to selection bias if subjects being enrolled in the study are a selective group of individuals (e.g. because of high refusal rates for participation) or if there is considerable drop-out due to incomplete follow-up. Experimental study designs are less commonly used to study disease causation given that it is often unethical and/or logistically difficult to randomize subjects. These designs are thus most commonly used to study the efficacy of treatment, and then are most commonly referred to as 'clinical trials'. Experimental designs are also used in field settings to study the effectiveness of interventions such as the introduction of vaccines or vitamin supplementation, and are commonly known as 'field intervention trials'.

Global burden of cancer (Fig. 1.1)

Cancer is becoming the major cause of death worldwide. The World Health Organization (WHO) statistics for the year 2011 indicate that 7.9 million deaths worldwide were due to cancer (followed by 7 million deaths from ischaemic heart disease and 6.2 million deaths from stroke). An estimated 14.1 million new cases and 8.2 million deaths from cancer (excluding non-melanoma skin cancer) occurred in 2012, with corresponding age standardized incidence and mortality rates of 182 and 102 per 100,000 respectively. More than 60% of the world's cancer cases occur in Africa, Asia and Central and South America. According to GLOBOCAN estimates, the 5-year prevalence of cancer was 32.6 million for both sexes combined in 2012. The numbers of new cancer cases and new cancer deaths were slightly higher in males than females. The five most commonly occurring cancers worldwide among males in 2012 were lung (16.7% of the total), prostate (15.0%), colorectal (10.0%), stomach (8.5%) and liver (7.5%). Among females, the most common sites were breast (25.2% of the total), colorectal (9.2%), lung (8.7%), cervix (7.9%) and stomach (4.8%).

While the above estimates reflect overall global patterns, there are stark differences in cancer patterns across the globe. One way to understand these differences in burden and type of cancer is to classify countries according to their human development

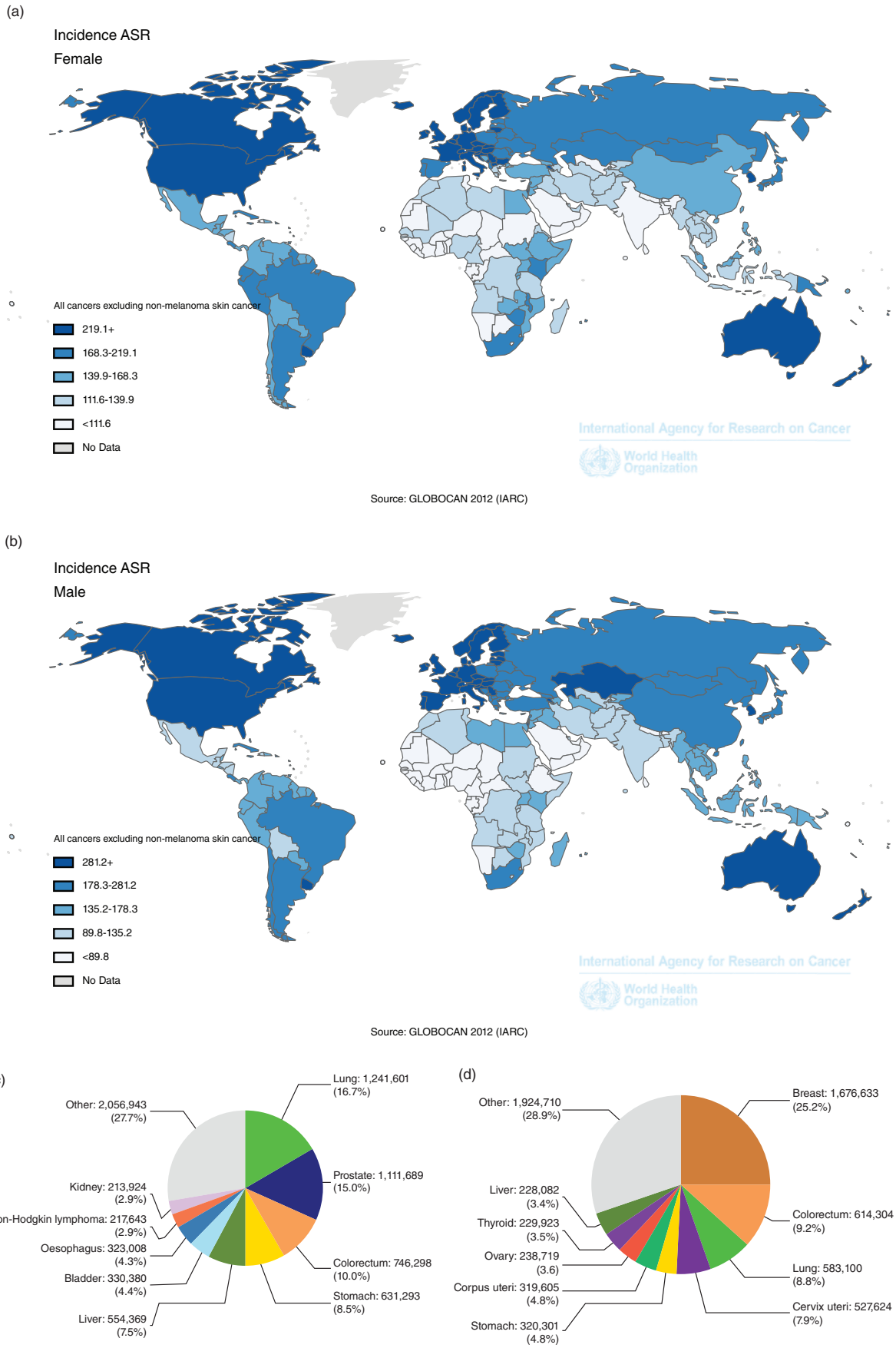

Figure 1.1 (a, b) World Cancer Incidence (WHO 2012) for males and females. (c, d) Percentage distribution of common cancer sites among males and females.

ASR, age-standardized rate. Data source: GLOBOCAN 2012. Reproduced with permission of WHO.

index (HDI). The United Nations Development Programme developed this composite index of three basic dimensions of human development: long and healthy life, level of education and standard of living as measured by gross national income per capita. Lung, breast, prostate and colorectal cancers are the most frequent cancers in countries with high or very high HDIs. On the other hand, countries with low or medium HDIs have a higher burden of infection-related cancers such as stomach, liver and cervical cancers. In recent years, countries with low or medium HDIs have also been witnessing a rise in cancers of the breast, colorectum and lung, indicating that they are undergoing a transition in economy and human development.

It is estimated that there will be >20 million new cancer cases by the year 2025. Countries with medium HDIs will experience the greatest increase in cancer burden, largely due to demographic changes with increases in life expectancy. Adoption of the higher-risk behaviours and lifestyle of more affluent countries (e.g. use of tobacco, higher fat diets) is also responsible for the changing profile of cancer in these settings.

Risk factors for cancer (Table 1.2)

Current knowledge indicates that lifestyle and/or environmental factors are the major contributors to the aetiology of the majority of cancers, although a small proportion can be explained by inherent susceptibility. The *attributable fraction* of a given risk factor is the proportion of the disease of interest that is thought to be due to that risk factor.

Table 1.2 Major risk factors for cancer

Risk factors	Cancer type
Tobacco use (smoking and chewing)	Oral cavity, pharynx, oesophagus, stomach, liver, pancreas, nasal cavity, larynx, lung, cervix, ovary, uterus, kidney, bladder, myeloid leukaemia
Alcohol	Mouth, nasopharynx, oropharynx, oesophagus, colorectum, liver, larynx, female breast
Chronic infection with human papillomavirus (HPV)	Cervix, oropharynx
Chronic infection with hepatitis B and C virus (HBV, HCV)	Liver
Chronic infection with *Helicobacter pylori*	Stomach
Obesity and physical activity	Colon, breast (postmenopausal), kidney, endometrium, oesophagus (adenocarcinoma), pancreas
Diet	Colon, breast, prostate
Reproductive and hormonal factors	Breast, ovary, endometrium
Occupation (exposure to asbestos, heavy metals, diesel exhaust)	Lung, urinary bladder
Pollution (air and indoor)	Lung, bladder, skin
Genetic susceptibility	All
Chemical compounds Aflotoxin (naturally occurring) Aspirin	 Liver Protective effect on colon cancer

Tobacco

Tobacco use in any form (including smoking or chewing) is the single largest cause of cancer worldwide. It has been associated with cancers of the oral cavity, pharynx, oesophagus, stomach, liver, pancreas, nasal cavity, larynx, lung, cervix, ovary, uterus, kidney and bladder, and with myeloid leukaemia. At least 16% of all cancers are estimated to be related to tobacco use, with a higher proportion of tobacco-related cancers among men (25%) than women (4%). In general, the risk of cancer related to smoking and smokeless tobacco use increases with the duration and amount of tobacco smoke/chewed. Even involuntary or passive smoking (the inhalation of second-hand smoke by non-smokers) has been shown to cause cancer, with an estimated 25% increase in lung cancer risk compared to non-smokers. Successful quitting of tobacco smoking has been associated with decreased risk, but risk still remains higher than for never-smokers. Tobacco smoke contains >7000 chemical compounds and smokeless tobacco products >3000, of which many are known carcinogens. Broad classes of carcinogens in tobacco smoke include polycyclic aromatic hydrocarbons, N-nitrosamines and aromatic amines. Similarly, smokeless tobacco products contain at least 28 carcinogens including tobacco-specific nitrosamines, N-nitrosoamino acids and volatile aldehydes such as formaldehyde and acetaldehyde, as well as metals including cadmium, lead, arsenic, nickel and chromium. The major pathways by which tobacco use produces cancer are thought to be DNA binding and consequent mutation, as well as inflammation and epigenetic mechanisms.

Alcohol

Alcohol intake is associated with increased risk for cancers of the oral cavity, hypopharynx, oropharynx, oesophagus, colorectum, liver, larynx and female breast. In addition to these cancer sites, there is some suggestive (but inconclusive) evidence for increased risk of cancers of the stomach, pancreas, prostate, kidney and bladder. There appears to be a positive dose–response relationship with the amount of alcohol consumed. Evidence suggests that the risk of head and neck cancer decreases with time since cessation of drinking. To date, no conclusive differences in carcinogenicity among alcohol beverages have been noted. There does seem to be a synergistic effect between

tobacco smoking and alcohol consumption on risk of cancers of the oral cavity, pharynx, larynx and oesophagus, whereby the risk of consuming both tobacco and alcohol is greater than the individual risk of each of these factors. Approximately 4.2% of all cancer deaths have been attributed to alcohol use.

Alcoholic beverages contain several carcinogenic compounds such as ethanol, acetaldehyde, aflatoxin and ethyl carbamate. The main mechanisms by which alcohol is thought to act as a carcinogen include: the genotoxic effect of acetaldehyde; the induction of cytochrome P450 2E1 and associated oxidative stress; increased oestrogen concentration; acting as a solvent for tobacco carcinogens; and altering folate metabolism and DNA repair.

Infections

There is growing epidemiological evidence that chronic infection with viruses, bacteria and macroparasites are strong risk factors for specific cancer sites. Overall, about 2 million (16%) of the total 12.7 million new cancer cases in 2008 are thought to be attributable to infection. The attributable fraction of cancer due to infection varies widely by geographical region, with the lowest rates in North America, Australia and New Zealand (<4%), and highest in sub-Saharan Africa (33%). Table 1.3 lists the infectious agents that have confirmed associations with cancer.

Infection with *Helicobacter pylori* is associated with gastric adenocarcinoma in the non-cardiac part of the stomach. Chronic infection with the hepatitis B (HBV) and/or hepatitis C (HCV) viruses has been consistently associated with hepatocellular carcinoma. Worldwide, the fraction of hepatocellular carcinoma attributable to infection with these viruses is estimated to be 77%. Infection with the human papillomavirus (HPV) is a necessary (but not sufficient) cause of cervical cancer. Generally, 13 high-risk HPV types are classified as carcinogens (HPV 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59 and 68). The two most common oncogenic HPV types for cervical cancer are HPV 16 and 18. In addition to cervical cancer, HPV is also related to the risk of anogenital and oropharyngeal cancers. Individuals infected with the human immunodeficiency virus (HIV) have increased risk of acquired immune deficiency syndrome (AIDS)-defining cancers (Kaposi sarcoma, non-Hodgkin lymphoma [NHL] and cervical cancer) and other virus-related cancers. The fact that only a portion of infected individuals develop cancer suggests that infection is not sufficient to produce cancer. Further, there is geographical variation in both infection rates and rates of infection-related cancer, suggesting that much of the infection-related cancer is preventable.

Table 1.3 Associations between infectious agents and human cancers (World Cancer Report 2014)

Cancer site	Infectious agent
Stomach	*Helicobacter pylori*
Liver	Hepatitis B, C virus, *Opisthorchis viverrini, Clonorchis sinensis*
Cervix	Human papillomavirus
Anogenital (penis, vulva, vagina, anus)	Human papillomavirus
Nasopharynx	Epstein–Barr virus
Oropharynx	Human papillomavirus
Non-Hodgkin lymphoma	*Helicobacter pylori*, Epstein–Barr virus, hepatitis C virus, human T-cell lymphotropic virus type 1
Kaposi sarcoma	Herpes virus with or without human immunodeficiency virus (HIV)
Hodgkin lymphoma	Epstein–Barr virus
Bladder	*Schistosoma haematobium*

Obesity and physical activity:

Overweight and obesity are associated with increased risk for cancers of the colon, breast (postmenopausal), kidney, endometrium, oesophagus (adenocarcinoma) and pancreas. There is also evidence that obesity increases risk of cancers of the gall bladder (in women), ovary and thyroid, non-Hodgkin lymphoma, multiple myeloma and leukaemia. Overweight and obesity are generally determined from the body mass index (BMI), which is calculated from the weight in kilograms divided by the square of the height in metres. Individuals with a BMI of over 25 kg/m^2 are considered to be overweight, while individuals with a BMI of over 30 kg/m^2 are considered as obese. The association of BMI with cancer risk generally demonstrates a dose–response relationship. However, BMI provides no information about body fat distribution. Other measures of body composition include the waist-to-hip circumference ratio (WHR) and skinfold thickness. Some studies indicate that measures of central adiposity may be better measures of cancer risk in certain populations than overall BMI.

Physical activity contributes to reduction in the risk of all obesity-related cancers, probably at least partially due to protection against weight gain. There is also evidence that physical activity has an independent effect on incidence as well as survival of patients with breast or colon cancers, possibly by acting through hormonal mechanisms.

Diet

There are many methodological challenges in measuring dietary intake, and therefore the exact role of dietary factors in causing human cancers remains unclear. Initial studies on diet and cancer revealed that diets rich in fruits, vegetables and whole grain protect against cancer. However, it is now clear that this protection might not be as strong as previously thought. Similarly, the role of fat in the development of many cancer types was

thought to be important. However, recent prospective studies have demonstrated little or no relationship between fat intake and the risk of breast, colon and prostate cancer.

Higher consumption of meat is associated with increased risk for colon cancer and possibly breast and prostate cancer. In addition to the food itself, the method of food preparation (e.g. grilling) and differences in metabolism are also important. Grilled and barbecued meat and fish contain carcinogenic polycyclic aromatic hydrocarbons and hetrocyclic amines, and high intake of these foods has been associated with increased risk of colorectal and stomach cancer. On the other hand, high intake of calcium, vitamin D and folate has been associated with a reduced risk of colorectal cancer. A Mediterranean-style diet which is high in cereals, fruits and vegetables, and low in animal products has been effective in weight loss (and associated with reduced cancer risk). Similarly, reduced consumption of soda and other sugar-sweetened beverages has resulted in weight loss.

Further research is required to fully understand the role of diet in cancer, including understanding the role of dietary behaviours during childhood and in early adult life.

Reproductive and hormonal factors

Epidemiological evidence for the carcinogenic effect of reproductive and hormonal factors is strongest for breast cancer, and moderately convincing for endometrial and ovarian cancer. Risk of developing breast cancer is almost double in nulliparous women compared to parous women. Women with lower age at first childbirth are at lower risk of developing breast cancer, and the risk increases linearly with later ages at first childbirth. While the protective effect of full-term pregnancy on breast cancer is well established, there is little evidence for the relationship with short-term pregnancies, including miscarriages and abortion. Early age at menarche and late age at menopause are also associated with increased risk of breast cancer. Women with early surgical menopause before the age of 40 years have approximately half of the risk of those who have natural menopause. Uses of hormone replacement therapy (HRT) and oestrogen–progesterone oral contraceptives have been associated with increased risk of breast cancer.

Nulliparity, early age at menarche and late age at menopause also increase the risk of ovarian and endometrial cancer. Use of oral contraceptive is associated with reduced risk of ovarian and postmenopausal endometrial cancer. Prostate and testicular cancer are also linked to hormonal factors; however, further studies are required to fully understand the relationship.

Occupation

The most comprehensive source of occupational exposures associated with cancer is maintained in a series of monographs published by the IARC. According to the IARC monographs, 32 occupational agents and 11 exposure circumstances have been identified as human carcinogens. Some of the identified agents (e.g. mustard gas) are now of only historical interest, while other workplace exposures such as asbestos, polycyclic aromatic hydrocarbons, heavy metals, diesel engine emissions and silica are still widespread. Due to the widespread existence of mixed exposures in the occupational setting, it is sometimes not possible to identify the exact agent responsible for carcinogenesis, and the occupational groups themselves are labelled as carcinogenic (e.g. painters and workers engaged in aluminium production or in rubber manufacture). It is not the occupation itself that confers a risk, rather the exposure or conditions at work that are responsible for the carcinogenic effects. Although the overall burden of occupational cancer is small, this burden may be substantial among the exposed occupational group. Research on exposure biomarkers has substantially contributed to the understanding of occupational cancer aetiology. For example, ethylene oxide was identified as a human carcinogen after the detection of specific protein adducts in exposed workers.

Pollution

Large numbers of people are exposed to environmental pollution from air, water and soil. Many of these pollutants are known or possible carcinogens. Emissions from multiple sources pollute the ambient air, but mainly stem from vehicle emissions, power generation and a range of carcinogenic compounds from industrial waste, including diesel emissions, polycyclic aromatic hydrocarbons and compounds containing asbestos, arsenic and chromium. Complex air pollution mixtures are characterized in terms of summary indicators such as $PM_{2.5}$, which is the mass concentration of fine particulate matter of <2.5 μm in diameter. Ambient $PM_{2.5}$ exposure has been estimated to contribute to 223,000 deaths from lung cancer globally.

In many countries (particularly in East, South and Central Asia), people are exposed to indoor air pollution from the burning of solid fuels such as coal or biomass for household cooking or heating. Exposure to indoor burning of coal has been strongly associated with lung cancer, and indoor burning of biomass may also be responsible for lung cancer.

Another major source of non-occupational carcinogen exposure is asbestos from the installation, degradation and repair of asbestos-containing products during house maintenance. Exposure to asbestos results in an increased risk of mesothelioma and may cause lung cancer, particularly among smokers.

Consumption of drinking water contaminated with arsenic causes cancers of the skin, lung and bladder, and possibly of other organs such as the liver and kidney. Drinking water may also be polluted with carcinogenic organic compounds, e.g. chlorinated solvents and pesticides. High nitrate levels in drinking water have been associated with stomach cancer.

Although the role of pollution in cancer causation is small in terms of attributable risk, pollution presents an important cancer hazard in certain geographical areas, despite the fact that cancer due to pollution is amenable to primary prevention.

Radiation

Exposure to ionizing radiation from both natural and man-made sources has been consistently linked with increased risk of several cancers, including non-chronic lymphocytic leukaemia (CLL) and cancers of the female breast, lung and thyroid. Evidence for increased risk, particularly at moderate-to-high doses, also exists for basal cell carcinoma of the skin, and cancers of the bone, brain, rectum and bladder.

While the main source of exposure to ionizing radiation in the general population remains natural radiation (cosmic rays and radionuclides originating from the Earth's crust), exposure from medical procedures is becoming an increasingly common source in many countries. Exposure to indoor radon is the main source of elevated exposure to natural ionizing radiation, while exposure to man-made ionizing radiation occurs mainly in the course of medical care due to diagnostic procedures (e.g. radiography and computed tomography) or treatment (e.g. radiotherapy). Exposure to ionizing radiation causes multiple types of DNA damage, including single-strand and double-stand breaks. Exposure to non-ionizing radiation in the form of ultraviolet radiation from both the sun and tanning devices causes all types of skin cancer, including melanoma. Exposure to extremely low-frequency magnetic fields has been associated with childhood leukaemia in some studies, but interpretation of this increased risk is difficult given the potential biases. Exposure to a radiofrequency electromagnetic field (including use of mobile phones) has been classified as possibly carcinogenic to humans (group 2B) in the IARC monograph series, largely based on reports of an association between heavy use of mobile phones and risk of glioma and acoustic neuroma. However, due to possible biases such as the self-reporting of mobile phone use, these associations have been difficult to interpret, particularly in the absence of strong supporting biological data. At present, there is insufficient evidence to assess cancer risk due to environmental exposure from transmitters of low-frequency electromagnetic fields, including from television, radio and mobile phone networks.

Genetic susceptibility

Familial aggregation has been shown for many cancers, including those of the breast, colon and prostate. Individuals with high-penetrant gene mutations have greatly increased risk of certain cancers (e.g. *BRCA 1* and *BRCA 2* for breast cancer). However, such high-penetrant mutations are very rare in the general population and likely account for a very small proportion of cancer cases globally.

With advances in microarray technologies it is possible to study more common genetic variants which confer smaller relative risk. Initial studies to identify the risk of common variants in the population using a candidate gene approach (in which variants were selected by investigators based on probable function) were mostly unsuccessful, mainly because the prior probabilities for these candidates was very low.. In the last decade, the genome-wide association study (GWAS) approach (which does not assume prior functionality of any the genotyped variants) has been successfully used to identify risk loci for several cancers. As the effect size for these variants is usually small, very large sample sizes are required to detect true associations. By design, GWAS studies examine a high volume of markers, or 'tag SNPs' (single nucleotide polymorphisms) across the genome, and thus detected associations are not necessarily the 'causal' variant. To correct for multiple comparisons and the associated probability of detecting chance findings, GWAS studies impose a much more stringent threshold for statistical significance ($P = \leq 5 \times 10^{-8}$), and replication of results is essential to establish a conclusive finding. To date, nearly 400 distinct genetic loci have been conclusively identified for common cancer types (breast, colon and prostate), as well as rare cancer type (e.g. Ewing sarcoma, neuroblastoma and paediatric cancers). Notable examples of common variants increasing the risk of cancer include *NAT2* variants with bladder cancer, variants in alcohol dehydrogenase genes with aerodigestive cancers, and *FGFR2* variants with breast cancer. Given the lack of dependence of genetic markers on disease development, the case–control study design is well-suited to identify new genetic loci.

Other factors

Poor oral hygiene, ill-fitting dentures and use of mouth wash with high alcohol content are all factors that have been associated with oral cancer. Additionally, some pharmaceutical drugs used in cancer treatment or as immunosuppressant or hormonal agents have also been associated with cancer development. The *Aristolochia* plant (used in traditional Chinese medicine as an antirheumatic and diuretic) has been associated with increased risk of cancer of the renal pelvis and ureter. Some naturally occurring chemicals from plants, fungi, lichens and bacteria are also carcinogenic (e.g. aflotoxins, ochratoxin or sterigmatocystin are associated with liver cancer). A protective role of aspirin and other non-steroidal anti-inflammatory drugs has been demonstrated for colorectal cancer. Epidemiological studies have indicated that metformin, a widely used oral anti-diabetic drug, may reduce the risk of several cancer types; however, this requires further confirmation.

Prevention and cancer control

Prevention of cancer comprises three stages: primary prevention (by avoiding exposure to carcinogenic agents); secondary prevention (by early detection of premalignant or early stage disease); and tertiary prevention (by providing effective treatment). A successful cancer control plan requires the proper integration and implementation of these activities at a population level, as well as continuous surveillance to evaluate the success of these activities.

The knowledge gained regarding modifiable risk factors for cancer has made it possible to achieve primary prevention of cancer by changing lifestyle and avoiding exposure to known carcinogens. The key preventable exposure is to tobacco in its

many forms. Tobacco control is not only the top priority for cancer control but also for many other chronic diseases. The WHO launched the Framework Convention on Tobacco Control (FCTC) to stimulate international efforts to reduce tobacco-related harms. There are two main approaches to tobacco control: one is directed towards the tobacco industry (by regulating price, availability and packaging), and the second towards current users or populations vulnerable to initiation of tobacco use (by education and restrictions). Ongoing epidemiological surveillance efforts have been implemented to monitor the progress of the WHO framework convention on tobacco control.

Obesity and physical inactivity are also largely modifiable risk factors for cancer which can be controlled by changing lifestyle behaviours. Behavioural weight loss programmes focusing on reduced caloric intake and participation in moderate-intensity physical activity have been shown to result in weight loss and reduced incidence of diabetes. Ongoing trials are evaluating the benefits of weight loss on reduction in cancer incidence and improvement in cancer survival.

Yet another means for the primary prevention of cancer is vaccination against infections clearly linked to cancer. Thanks to the widespread introduction of HBV vaccination, the incidence of liver cancer has decreased dramatically in the past few decades. Following the recent development of prophylactic vaccines for the control of HPV-related cancer, vaccination of adolescent girls is being recommended and implemented in most parts of the world.

Secondary prevention of cancer can be achieved by implementing population-based early detection programmes for cancers of the cervix, colorectum and breast. These programmes aim to detect premalignant or early-stage disease when effective treatment is available. Common methods of cervical cancer screening include Pap smear testing, testing for HPV DNA and visual inspection of the cervix. For colon cancer, both faecal occult blood testing and faecal immunohistochemical testing have been shown to be effective. Screening with mammography permits detection of early-stage breast cancer. Studies are currently underway to evaluate the effectiveness of clinical breast examination in reducing breast cancer mortality. Population-based screening/early detection programmes require political commitment, engagement of civil society and the ability to mobilize a large number of healthcare professionals. Implementation of primary and secondary prevention measures requires legislative and regulatory initiatives as well as population-wide campaigning. Monitoring and evaluating the success of these efforts using epidemiological study designs is key.

Conclusion

Despite being a relatively young field, cancer epidemiology has successfully quantified patterns of disease and identified several key causes of cancer, thus paving the way for the development of cancer control programmes. Epidemiological surveillance systems have been set up to monitor the burden of cancer and to evaluate the effectiveness of preventive measures. Newer technologies and methodologies are being used to improve the precision of epidemiological studies by improving the classification of exposure and outcome. In addition to traditional data on risk factors from questionnaires, large epidemiological studies are now using complex methods of exposure assessment and collecting various biospecimens (such as blood, tumour tissue and saliva) to examine the role of various biomarkers (e.g. variations in DNA, proteins, metabolites and microbiome) in the hope of gaining a better understanding of the complex aetiological role of the environment and genes in the development of cancer.

More detailed information on the epidemiology of individual cancers is given in the relevant chapters.

Recommended reading

A Catalog of Published Genome-wide Association Studies. Available at http://www.genome.gov.gwastudies

Adami H, Hunter D, Trichopoulos D, eds (2008) *Textbook of Cancer Epidemiology*. New York: Oxford University Press.

Burton PR, Clayton DG, Cardon LR *et al.*; Wellcome Trust Case Control Consortium (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3000 shared controls. *Nature* 447:661–678.

De Martel C, Ferelay J, Franceschi S *et al*. (2012) Global burden of cancers attributable to infection in 2008: a review and synthetic analysis. *Lancet Oncol* 13:607–615.

Ferlay J, Soerjomataram I, Ervik M *et al*. (2013) GLOBOCAN 2012 v1.0, Cancer incidence and mortality worldwide: IARC Cancer Base no. 11. Lyon, IARC.

Global Initiative for Cancer Registry Development in Low and Middle income Countries. Available at http://gicr.iarc.fr/

IARC (2007) Smokeless tobacco and some tobacco specific N-nitorosamines. *IARC Monogr Eval Carcinog Risks Hum* 89:1–592.

IARC (2013) Non-ionizing radiation, part 2: radiofrequency electromagnetic fields. *IARC Monogr Eval Carcinog Risks Hum* 102:1–460.

Pukkala E (2010) Nordic biological specimen bank cohorts as a basis for studies of cancer causes and controls: quality control tools for study cohorts with more than two million samples donors and 130,000 prospective cancers. In: Dillner J, ed. *Methods in Biobanking*. New York: Springer, pp. 61–112.

Renehan AG, Tyson M, Egger M *et al*. (2008) Body mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. *Lancet* 371:569–578.

Stewart BW, Wild C, eds. (2014). *World Cancer Report 2014*. Lyon: IARC.

WHO (2011) *WHO Report on the Global Tobacco Epidemic, 2011. Warning about Dangers of Tobacco*. Geneva: WHO.

World Cancer Research Fund/American Institute for Cancer Research (2007) *Food, Nutrition, Physical Activity, and the Prevention of Cancer; A Global Perspective*. Washington, DC: American Institute for Cancer Research.