IJC
International Journal of Cancer

Special Section Paper

Cancer Incidence in Five Continents: Inclusion criteria, highlights from Volume X and the global status of cancer registration

F. Bray[1], J. Ferlay[1], M. Laversanne[1], D.H. Brewster[2], C. Gombe Mbalawa[3], B. Kohler[4], M. Piñeros[1,5], E. Steliarova-Foucher[1], R. Swaminathan[6], S. Antoni[1], I. Soerjomataram[1] and D. Forman[1]

[1] Cancer Surveillance Section, International Agency for Research on Cancer, Lyon, France
[2] Scottish Cancer Registry, NHS National Services Scotland, Edinburgh, United Kingdom
[3] Registre Des Cancers De Brazzaville, Brazzaville, Congo
[4] North American Association of Central Cancer Registries, USA
[5] Instituto Nacional de Cancerología, Colombia
[6] National Cancer Institute Chennai (WIA)

Cancer Incidence in Five Continents **(CI5), a longstanding collaboration between the International Agency for Research on Cancer and the International Association of Cancer Registries, serves as a unique source of cancer incidence data from high-quality population-based cancer registries around the world. The recent publication of Volume X comprises cancer incidence data from 290 registries covering 424 populations in 68 countries for the registration period 2003–2007. In this article, we assess the status of population-based cancer registries worldwide, describe the techniques used in CI5 to evaluate their quality and highlight the notable variation in the incidence rates of selected cancers contained within Volume X of CI5. We also discuss the *Global Initiative for Cancer Registry Development* as an international partnership that aims to reduce the disparities in availability of cancer incidence data for cancer control action, particularly in economically transitioning countries, already experiencing a rapid rise in the number of cancer patients annually.**

The *Cancer Incidence in Five Continents* (CI5) series serves as a unique source of data on cancer incidence around the world. In the first volume the three editors, Sir Richard Doll, Drs Peter Payne and John Waterhouse, remarked that "*the most valuable data are, undoubtedly, the rates obtained by recording the occurrence of every case of cancer in a defined community over a specified period.*"[1] Fifty years on, and now in its tenth edition,[2] this principle of geographic pathology remains the *raison d'ètre* for the CI5 series. The sheer variability in comparable incidence data observed by place (within volumes) and time (between volumes) provides critical evidence of the environmental, lifestyle and infectious origin of many common cancers. It also points to the impact of interventions linked to increased awareness, prevention and early detection and by extension, the potential avoidability of many cancers.

Since the third volume,[3] CI5 has been published jointly by the International Agency for Research on Cancer (IARC) and the International Association of Cancer Registries (IACR) with the volume editors selected by these two organisations. Comparability has been assured by the sole inclusion of high quality datasets from population-based cancer registries worldwide. The preparation and systematic evaluation of indices of completeness and accuracy has been a hallmark of the CI5 editorial process since its inception, and requires careful attention by the editors to ensure that accepted datasets are of sufficiently high quality to warrant their inclusion.

This article has three principle aims centred on publication of the most recent tenth Volume, comprising incidence data from 424 population-based cancer registries covering 290 populations in 68 countries, mainly over the registration period 2003–2007. The first is to outline the editorial procedures utilised in ensuring a transparent and impartial evaluation of the individual datasets submitted from registries. Secondly, we seek to highlight the notable variation in the incidence rates of selected common cancers that continues to be revealed by this compilation. Lastly, we present the status of population-based cancer registries worldwide, the current disparities in their development within low- and middle-income countries (LMIC), and the *Global Initiative for Cancer Registry Development* (GICR) a partnership to radically improve the situation.

Key words: cancer, incidence, registries, population based, epidemiology
DOI: 10.1002/ijc.29670
History: Received 23 June 2015; Accepted 24 June 2015; Online 1 July 2015
Correspondence to: Dr. Freddie Bray, Cancer Surveillance Section, International Agency for Research on Cancer, 150, Cours Albert Thomas, F-69372 Lyon Cedex 08, France, Tel.: +33-0-4-72-73-84-53, E-mail: brayf@iarc.fr

Material and Methods

Prior to discussing the methods used in presenting geographic variations in incidence, we provide a short

description of the principles of data quality evaluation at the cancer registry, and the derivation of the data compiled in Volume X, including the formal editorial procedures in place to evaluate each registry contribution.

Registries and data quality

Population-based cancer registries (PBCR) collect information on cancer cases from multiple sources within a defined target population. Pathology laboratories, hospital records (from both public and private hospitals) and death certificates are the key sources, while access to records from other facilities (*e.g.*, radiotherapy and oncology departments, hospital discharge records, imaging facilities and haematology laboratories) may augment reporting. In general, the greater the number of sources used for case finding, the more complete and accurate the reporting. Conversely, major omissions of key sources may raise concerns that case finding is incomplete.

There are specific areas of data quality assessment founded on the principles set out by the editors in the first volume of CI5.[4] The practical aspects and techniques of formal evaluation at the cancer registry have been examined most recently in a two-part review[4,5] and in a guidance document emphasising quality assurance in low- and middle-income settings.[6] In the preparation of this volume, the CI5 editorial team assessed the registry datasets using three axes of data quality (more detail in Appendix):

Comparability. Determining the extent of the comparability of a cancer dataset requires consideration of the registry's procedures, including the standards and definitions used in registration, and their adherence to established international standards and guidelines.

Completeness. The methods used to evaluate overall completeness can be considered *semi-quantitative*, as they provide an indication of the degree of completeness of incidence in one population relative to other registries, or incidence in a single registry over time.

Validity. Several indicators of validity (or accuracy), the proportion of cases recorded as having a given characteristic that truly do have that attribute, relate to the precision of the source documents at the registry, and the level of expertise in abstracting, coding, and recoding cases.

CI5 call and editorial process

A call for data was sent to 591 members of the IACR in September 2011. Registries submitted their data and completed an accompanying questionnaire online, *via* the IARC data portal. By early-2012, data were received from 372 cancer registries covering 521 populations in 80 countries.

The editors carried out an extensive process of verifying coding, identifying duplicate registrations, querying unlikely or impossible combinations of codes and converting the data to a standard format. Next, we conducted a detailed assessment of preassembled registry-specific tables including site-specific case numbers, age-specific rates and summary rates (crude, cumulative and age-standardized), the populations at risk by sex and age, and a comparison with the 5-year population data from the previous volume (where applicable). The completed questionnaires provided further details regarding data comparability including the definitions used by each registry.

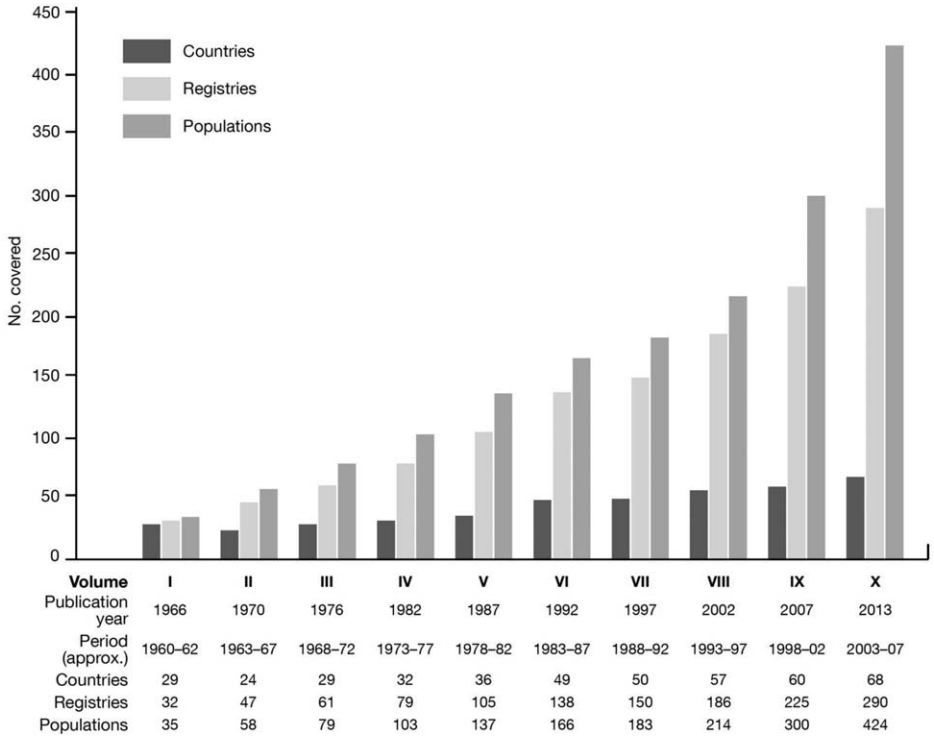
For several of the quantitative indices introduced, a comparison with "standard values" was performed that represented the values from cancer registries in the same region, or where there were a sufficient number of high-quality registries, within a country, using the published data from the two previous volumes of CI5 (Volumes VIII and IX). The use of such regional or national standard values recognises that diagnostic practices (especially with respect to histology and cytology), and the accuracy of recording the underlying cause of death on death certificates vary between populations and regions. In total, 25 regions or countries were defined (Forman et al, reference 2). For each, the mean and variance of values of the site-specific age standardised incidence, proportion of cases microscopically verified and M:I ratios were calculated. In addition, summary statistics of the above indicators were used to identify unusually high or low cancer incidence rates in registries within regions or countries for all sites combined and for major cancers.

Presentation of results

We use two measures of age standardisation in presenting the variability in the incidence: age-standardised rates (ASR) using the world standard population as defined in the first volume of CI5, and cumulative risk, the probability of developing cancer before the age of 75 in the absence of competing causes of death. The resulting bar charts are sorted by cumulative risk for selected cancer types and sex, restricting registries to a maximum of three per country to enable a reasonable visualisation of risk in individual populations worldwide. In describing the relative magnitude of the variability of the ASR across cancer sites by sex, we excluded the upper and lower 10% of registries in terms of the magnitude of the ASR, and estimated the ratio of the 10th and 90th centiles to provide a conservative estimate of the extent of the cancer-specific variability globally. This serves to limit prospects of an overestimation of ratios due to lower than expected rates *via* underascertainment or

Table 1. Numbers of registries submitted and number and proportion finally included in *Cancer Incidence in Five Continents Volume X,* by continent

	Registries submitted	Registries included	Proportion accepted (%)
Africa	18	8	44
Asia	102	63	62
Central and South America	35	25	71
Europe	136	118	87
Oceania	11	10	91
North America	70	66	94
Total	372	290	78



Volume	I	II	III	IV	V	VI	VII	VIII	IX	X
Publication year	1966	1970	1976	1982	1987	1992	1997	2002	2007	2013
Period (approx.)	1960–62	1963–67	1968–72	1973–77	1978–82	1983–87	1988–92	1993–97	1998–02	2003–07
Countries	29	24	29	32	36	49	50	57	60	68
Registries	32	47	61	79	105	138	150	186	225	290
Populations	35	58	79	103	137	166	183	214	300	424

Figure 1. Coverage of registry populations included in CI5 Volumes I-X.

underdiagnosis, or as a result of few cases and a greater inherent randomness in selected populations and/or small population sizes. It may also, to a lesser extent, reduce overestimation of incidence ratios for cancers associated with detection through diagnostic tests and screening.

Results

Registry coverage

Figure 1 shows the overall coverage in terms of the number of countries, registries and populations included in each volume of CI5. Volume X comprises cancer incidence data from 68 countries, 290 cancer registries and 424 populations, regionally depicted by the global map in Figure 2. The proportion of the world's total population covered by the registries included in Volume X is 14%, with the levels of coverage by continent markedly lower in Africa (2%), Asia (6%) and Central and South America (8%) relative to Europe (42%), Oceania (78%) and North America (95%). The number of cancer registries that submitted datasets for Volume X and the number and proportion that were accepted for inclusion (by continent) are presented in Table 1.

Global variations by cancer type

Figure 3 quantifies the relative variability and portrays the absolute differences in the ASR for 27 cancer sites according to the 10th and 90th centiles (rates are for males, other than for breast, cervix uteri, corpus uteri and ovary). There is at least a threefold variation in the rates of each of these cancer types across registry populations worldwide, with the exceptions of ovary cancer (2.1). Male rates vary 45-fold (46.6) for melanoma of the skin and 11-fold for testicular cancer globally, while a further ten cancer types in men vary between fivefold and tenfold across populations, including nasopharyngeal, stomach and liver cancers. Prostate, female breast and male lung cancer rates convey large absolute variability in the ASR with rate differences exceeding 100, 60 and 50, per 100,000 respectively.

Variations in the most common cancers worldwide

To appreciate variations for the four most frequent cancers worldwide, Figures 4. (a-d) convey the differences in risk based on cancer diagnoses for the years 2003–2007: female breast, prostate, lung (females) and colorectum (males). The lifetime risk of breast cancer varies from almost 12% in Belgium to around 1% in the urban population of Blantyre, Malawi and a rural registry population (PROMEC) in South Africa. Rates are high to intermediate in registry populations in Western Europe, as well as in North America and Australia/New Zealand, but relatively low in Asia (including China and India) and Africa. For prostate cancer, the risk circa 2005 is highest amongst US blacks, where cumulative incidence is approaching 20%, with highly elevated lifetime risks also observed in Austria, France and Brazil. In contrast, risks of 1–2% are seen in parts of Asia (including Thai, Indian and Chinese registry populations).

For lung cancer, lifetime risk in females is highest in a number of North American and Nordic populations (Denmark and Iceland), with a cumulative risk of almost 5% in Quebec. Notably, elevated rates of lung cancer in women are also observed in Thailand (Chiang Mai) and India (Mizoram). Risk of female lung cancer is lower than 1% in most African countries

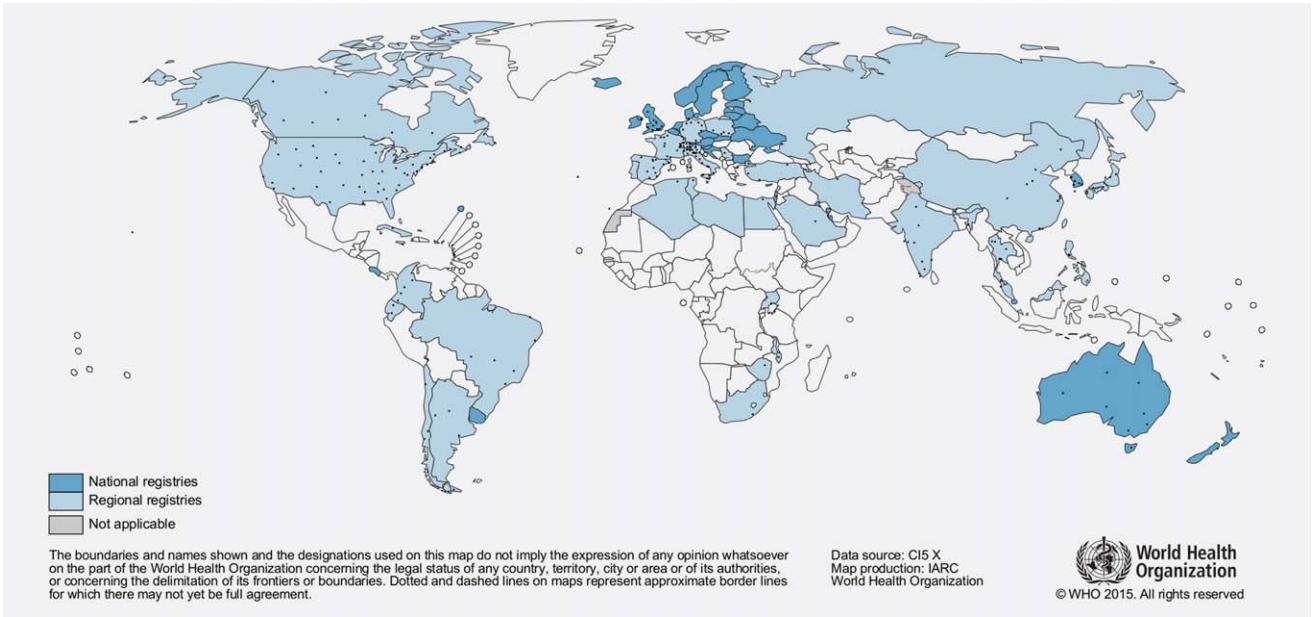


Figure 2. Global map of registries included in CI5 Volume X, with main cities marked. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

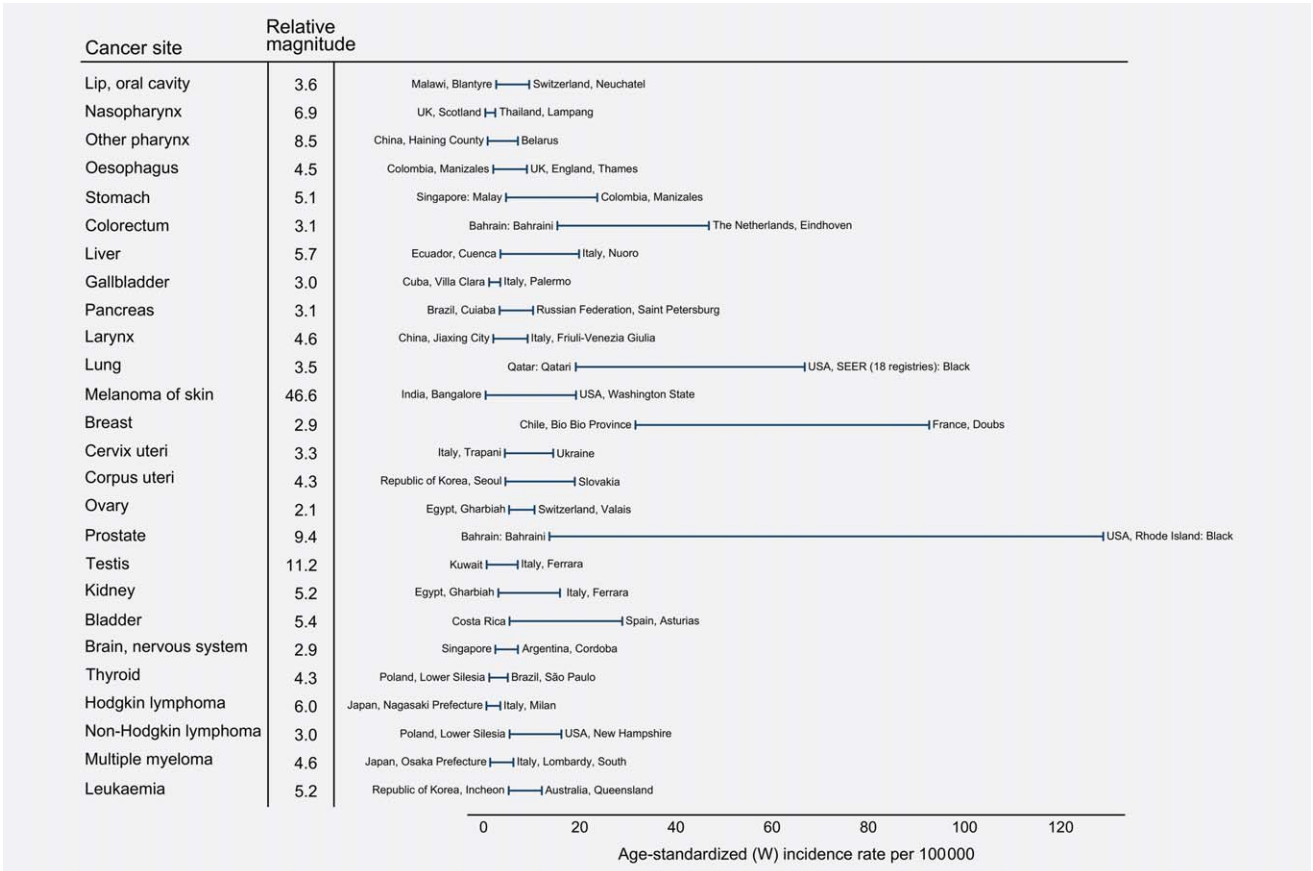


Figure 3. Relative and absolute global variations in age-standardised rates (world) of registry populations included in CI5 Volume X, based on rates within the 10th and 90th percentiles. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

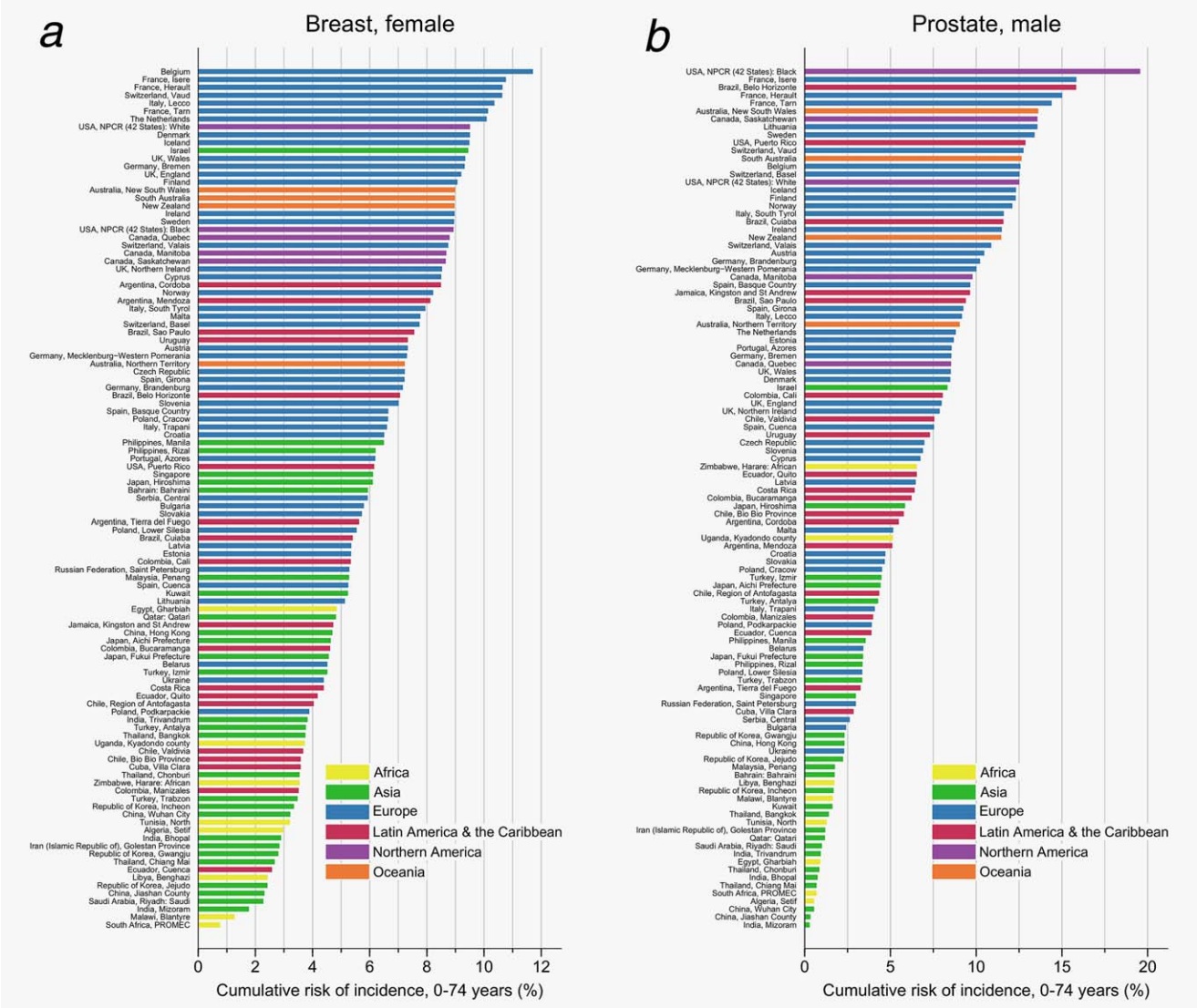


Figure 4. (a–*d*) Bar charts of cumulative risk of cancers of the female breast, prostate, lung (women) and colorectum (men): up to three registries included per country, sorted in descending order by magnitude of risk. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

represented and is very low in a number of North African and Middle Eastern populations. For colorectal cancer (in men), risk varies from around 8% in Japan (Hiroshima), Slovakia and the Czech Republic to 1% or less in many African and Middle Eastern countries including Algeria (Setif), Egypt (Gharbiah), (North) Tunisia and Kuwait, as well as India (Mizoram). Risk tends to be low to intermediate in Latin America.

Variations in cancers linked to infection

Figures 5. (a-c) convey the differences in the risk for the three frequent cancers associated with infection, based on cancer diagnoses in 2003–2007. The cumulative incidence of cervical cancer is very high, between 5 and 10% in the three Sub-Saharan African regional registries compiled in the volume, in Zimbabwe (Harare blacks, where risk approaches 10%), Malawi (Blantyre) and Uganda (Kyadondo country). Risk is intermediate (2–3%) in a number of South American, Asian and Eastern European regions, but low in most registries in high-income settings, including the Middle East.

Stomach and liver cancer in men both display marked variations in risk across the registry populations worldwide (Figs. 5*b* and 5*c*), with lifetime risk of gastric cancer incidence ranging from 5% or above in several regions of Japan (Hiroshima and Fukai Prefecture), Korea (Gwangju, Incheon and Jejudo) and in India (Mizoram). Risk is intermediate (around 3%) in regions of Eastern Europe, South America and certain regions within countries in Asia (parts of China, Golestan province in Iran), and low in many countries, not especially linked to geography, including areas within Africa, Asia and Europe. Very high (around 6%) to intermediate (2%) lifetime risk of liver cancer is seen in regions within Asian countries in Korea, Japan, Thailand, the Philippines and China, with a few exceptions (*e.g.*,

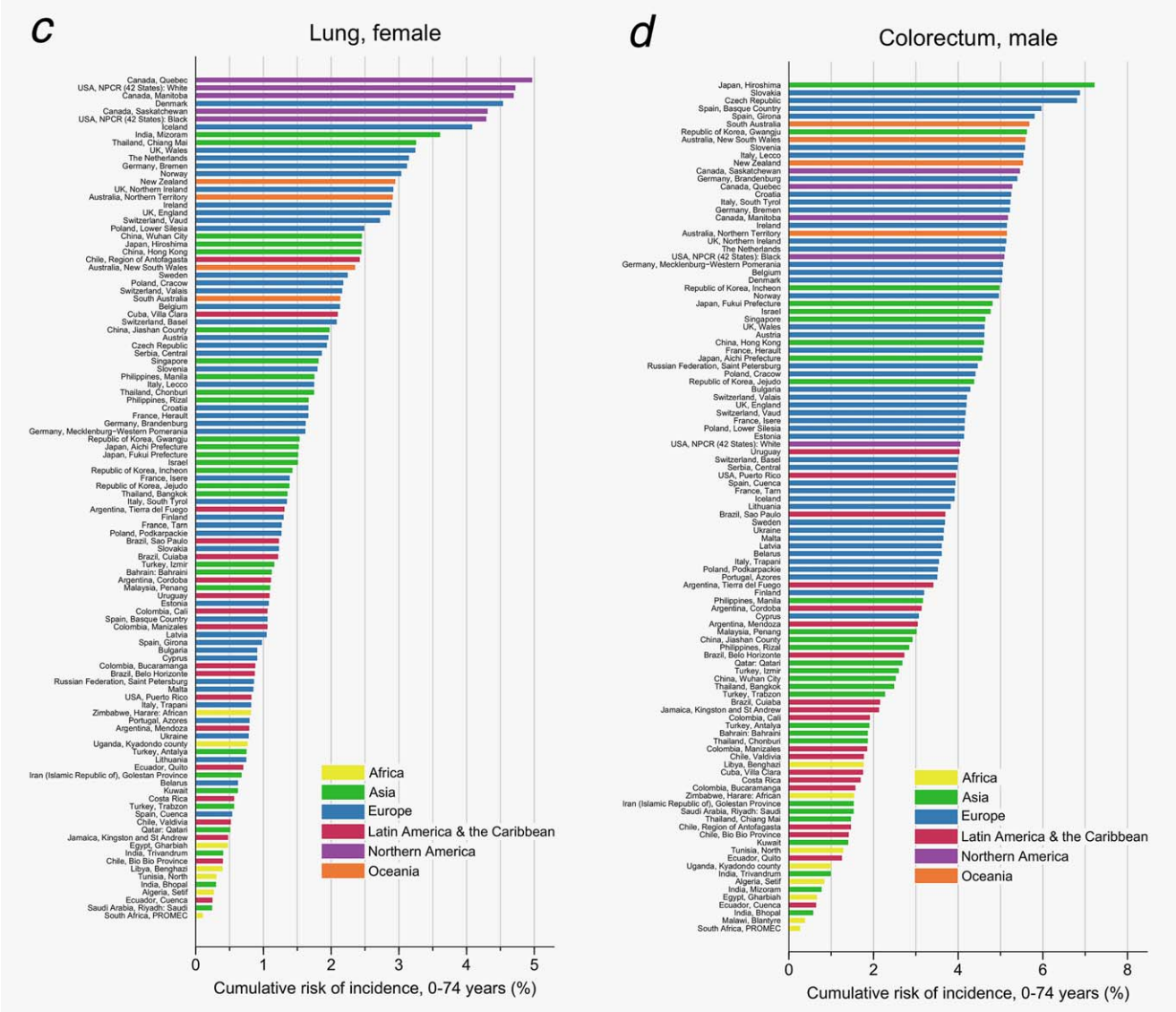


Figure 4. Continued.

Gharbiah in Egypt and Harare blacks in Zimbabwe). Risk is lowest, around 0.25% in regions as populations as Kingston (Jamaica), The Netherlands and Setif (Algeria).

Variations in selected cancers common in well-defined endemic areas

Figures 6. (a-c) show the differences in the risk for selected cancers worldwide that have distinct geographic patterns of incidence. The lifetime risk of malignant melanoma of the skin in males was highest in registry populations in Australia, Northern and Western Europe, and Northern America (whites) where risk of being diagnosed with this cancer was up to 4% (Fig. 6*a*). In contrast, risk was very low in Asian and African populations as well as among Black Americans, with risk as low as 0.15% in Kyadondo county (Uganda) to 0.03% in Riyadh (Saudi Arabia). It is worth noting the 26-fold difference in risk between whites relative to blacks in the United States.

As for oesophageal cancer, risk is very high in sub-Saharan African registry population such as in Blantyre (Malawi) where lifetime risk was almost 4%. Several registry populations in Asia also had elevated rates, *e.g.,* in Hiroshima (Japan), Jiashan county (China) or Golestan province (Iran), ranging from 1.8 to 2.8%. Risks were intermediate in many European and Latin American populations, and generally quite low in most registries in Asia and North Africa or the Middle East.

Finally, Figure 6*c* displays the geographical variation in the lifetime risk of kidney cancer based on diagnoses in the years 2003–2007. The Czech Republic stands out with a cumulative risk of incidence that approaches 3%. Risks were intermediate to high (1–2%) in most European registry populations as well as in Australia and North America alongside higher income Asian and Latin American populations. In contrast, lifetime risk of kidney cancer were uniformly low in the African populations represented (<0.5%).

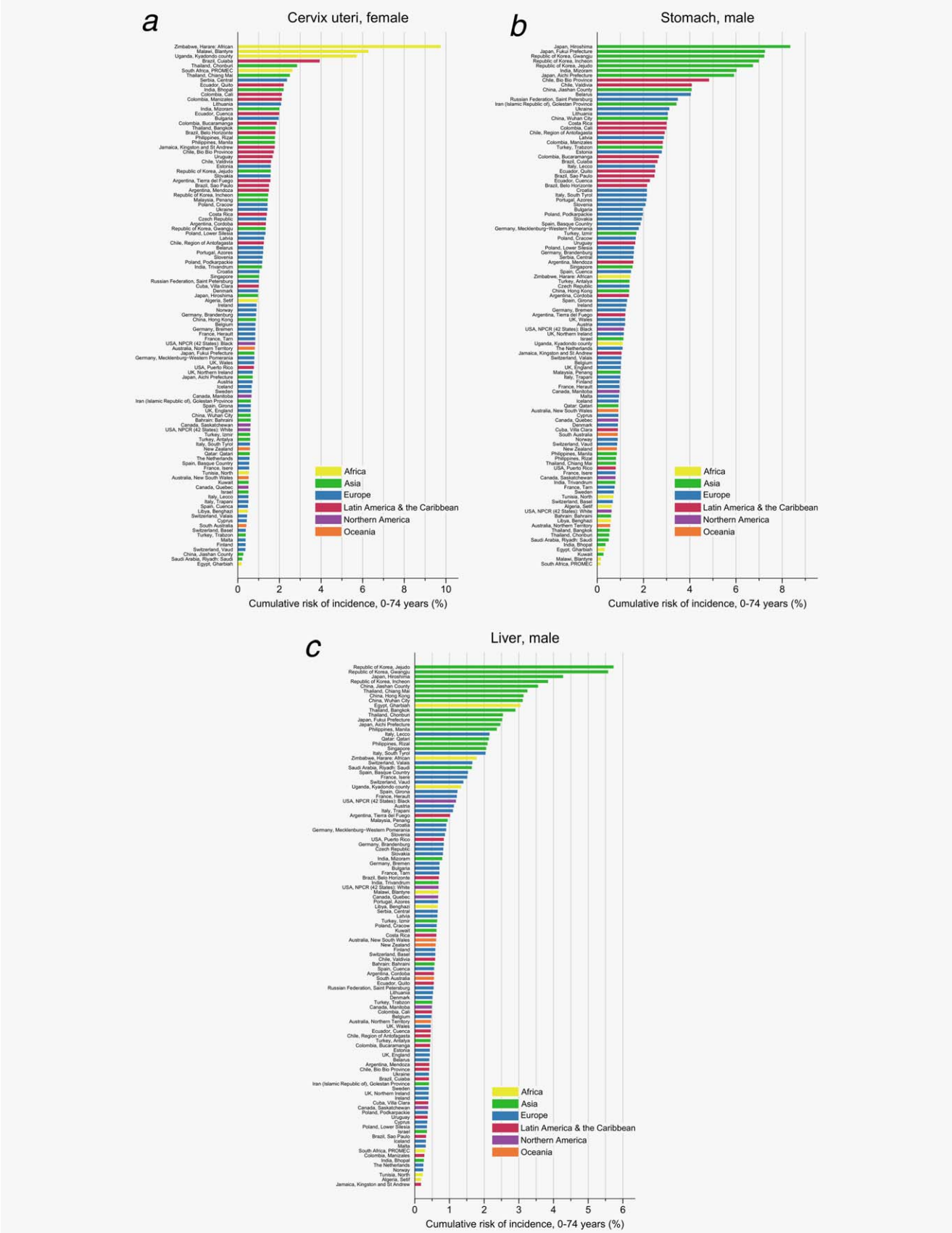


Figure 5. (a–*c*) Bar charts of cumulative risk of cancers of the cervix, stomach (men) and liver (men): up to three registries included per country, sorted in descending order by magnitude of risk. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

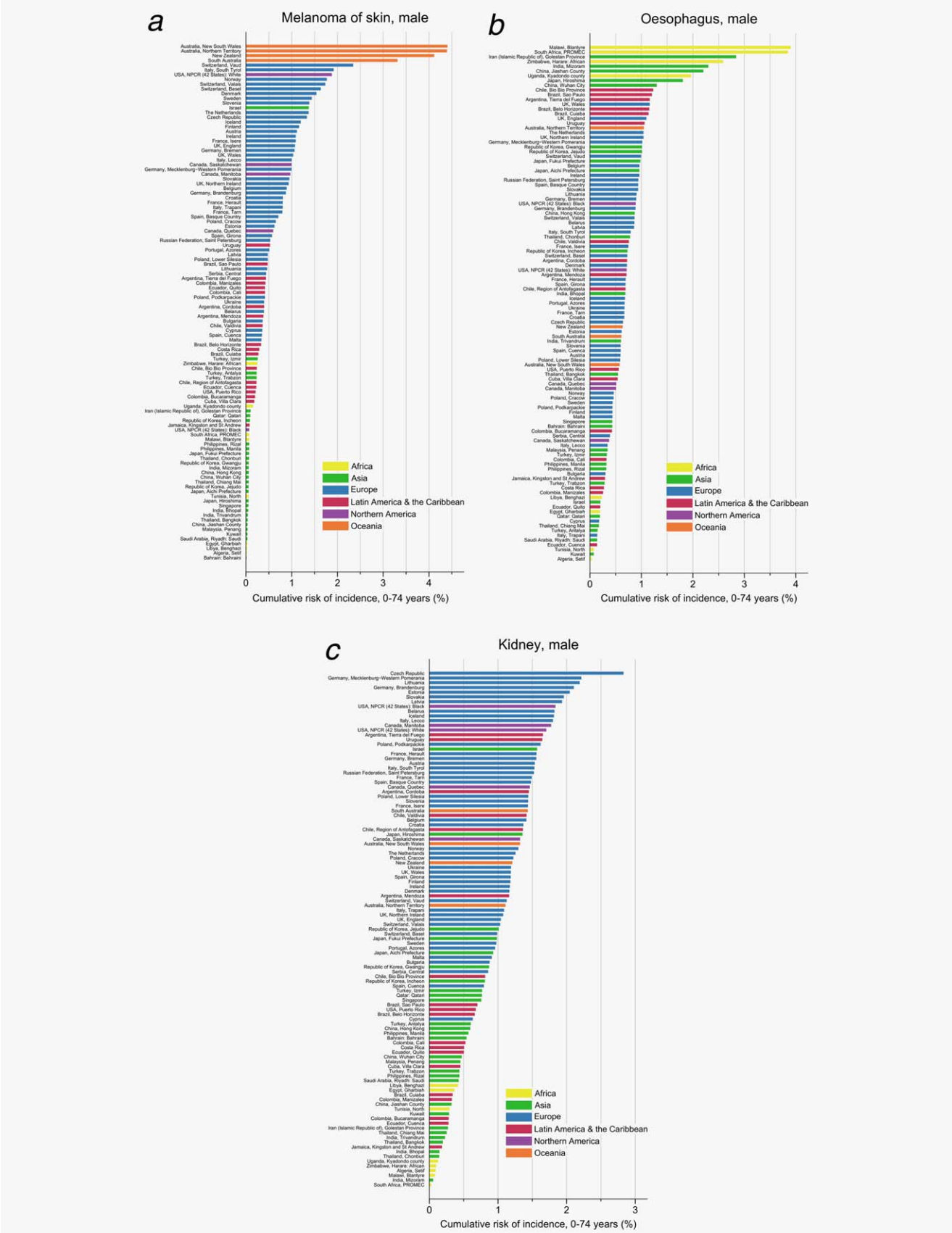


Figure 6. (a–*c*) Bar charts of cumulative risk of melanoma of the skin (men), and cancers of the oesophagus (men), and kidney (men): up to three registries included per country, sorted in descending order by magnitude of risk. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Discussion

Cancer Incidence in Five Continents, a collaboration between IARC, the IACR and population-based registries worldwide, is a major resource for understanding the global patterns and trends of cancers. With the inclusion of an increasing number of high-quality population cancer registries in successive volumes, CI5 has borne witness to consistent and extraordinary variability in the rates of many cancers in different populations between and within continents. The information collected within CI5 is immensely valuable in its own right and also represents one of the key information sources in estimating the global cancer burden.[7] It is important to note that the latter national estimates (GLOBOCAN, http://globocan.iarc.fr) are for a single year (currently for 2012). CI5, on the other hand, comprises continuous recorded national or regional data of high quality from population-based cancer registries essential for the monitoring and evaluation of cancer control activities, planned or implemented at the national or regional level within countries. At present, incidence estimates for 62 countries within GLOBOCAN are based on sources other than registries, often from neighbouring countries, in the absence of this critical information.

We report here on the spectrum of cancer risk observed circa 2005 in registry populations worldwide, noting a ratio of 3–45 in the lowest *versus* the highest rates for specific cancers, even after excluding possible extreme values that comprise rates within the upper and lower 10%. Based on similar comparisons using information from Volume III of CI5,[3] Doll and Peto noted that it is this geographic diversity, along with key evidence from the other principal means of descriptive epidemiology study—time trends in incidence rates and the study of rate in migrants relative to their host populations—that provides unique evidence of the avoidability of cancer.[8] This evidence must cast doubt on the recent assertions that only one-third of the variation in cancer risk is attributable to environmental factors or inherited predisposition.[9]

The variability in the rate ratios and the lifetime risk provides fascinating clues as to what is known and unknown about different cancer types, and the need to elucidate fundamental causes of certain cancers. There are, as an example, marked differences (larger than threefold) in the risk of lung, cervical, stomach and liver cancer and melanoma of the skin that are strongly linked to specific known risk factors (tobacco smoking causing lung cancer, infections with certain viruses or bacteria causing cervix, liver and stomach cancer and exposure to UV among fair skin populations associated with melanoma). This has in turn led to cancer control actions that have successfully reduced morbidity and premature deaths related to these cancers.[10–12] Diagnostic practices, incidental findings and the impact of screening may partly explain geographical differences in breast, prostate[4] and thyroid cancer incidence rates.[13] On the positive side, the elevated years of life lived with disability linked to these cancers[14] have raised awareness on overdiagnosis and overtreatment. For many sites or regions, causes of the reported variations are yet to be discovered such as the high oesophageal cancer risk in Kenya[15] and Iran[16] or the high risk of kidney cancer in the Czech Republic.[17] Nonetheless the mapping of cancer incidence through the CI5 series has increased a global awareness of particular patterns, redirecting international resources to support etiological studies to increase capacity of the health care systems in these areas.

Table 2. Proportion of registry coverage in *Cancer Incidence in Five Continents* Volumes I, V and X, by continent

	Proportion of regional population represented		
	Volume I ≈1965	Volume V ≈1985	Volume X ≈2005
Africa	0.4	0.0	1.9
Asia	0.3	2.0	5.7
Central and South America	5.1	4.9	7.5
Europe	7.3	17.0	41.7
Oceania	19.4	83.0	77.8
North America	7.8	28.2	94.7
Total	2.6	6.3	14.3

Cancer is a major public health issue, representing a leading cause of morbidity and mortality worldwide. The estimated 14 million new cases of cancer diagnosed 2012 is expected to increase to 24 million by 2030.[18] The need for action is especially critical in countries undergoing economic transition, which are currently ill-equipped to provide the necessary patient care and will bear the disproportionate impact of the projected increase in the future cancer burden. The recent global political focus on noncommunicable disease (NCD) provides additional impetus, with governments around the world committed to reporting on cancer as part of the United Nations Resolution on NCD in 2011 and WHO Global Action Plan 2013–2020. As illustrated in Table 2, while substantive increases in the overall number of high quality registries can be seen over the last half-century, the expansion in absolute proportions of registry coverage is still minimal in LMIC regions, with Africa, Asia and Latin America all below 10% coverage in the current volume. Indeed, PBCR remain to be established in many LMIC worldwide. In many instances, PBCR is presently the only source of information on cancer because of the absence of systematic and reliable death certification.

While there has been uniform adherence in successive volumes of CI5 to the editorial philosophy introduced by the editors in the first volume,[1] two modifications within Volume X are worthy of note: firstly, the more rigidly-defined evaluations and groupings (based on ranges of acceptability) that were introduced in Volume IX[19] were discontinued. Although the introduction of these quality "thresholds" had the laudable objective of increasing editorial transparency, the criteria imposed had undesirable consequences, including an automatic exclusion of

registries on the basis of low or high values of a single indicator, without further assessment. Secondly, the expanding number of registries submitting data (many for the first time) required innovation to efficiently manage the editorial review process while maintaining consistency in the review of individual registries; comparative overviews of key quality indicators between registries by region and, where applicable, country were introduced as supplementary tables to augment the standard formats.

The evaluation of data quality of CI5 imposed by the Editors is necessarily strict to maintain a compilation for comparative purposes. However, cancer registries operating in LMIC settings face particular challenges to follow international registration standards, linked to the characteristics of the health systems in place.[6] For example, a lack of coverage by pathology laboratories or difficulty accessing diagnosis records reduces the percentage of microscopically verified (MV%) cases and results in postponement of the incidence date. A low MV% in many registries in lower income settings can reflect the reality, where clinical diagnosis is the vital component in reaching reasonable levels of case ascertainment. It can also be the consequence of the local patterns of cancer, for example, in Western Africa or Asia, where liver cancer incidence predominates. Similarly, the death certificate and M:I ratio methods for estimating completeness depend on the availability of relatively high-quality (complete and accurate) certification of cause of death in the area covered by the cancer registry; alas this is not the case in many LMICs (almost all countries in Africa and many in Asia) and in a few high income countries too.

Noninclusion can be a stigma to registries trying to develop and expand if they are considered (by their programme owners, key stakeholders, *etc.*) of "low quality," because they are not reaching the high standard of operations and data output to be included in the present CI5 Volume. To provide an alternative means of disseminating results and to foster networks, one of the aims is to co-develop continental reports produced by local editorial boards. These reports will compile all available cancer incidence data with minimal quality evaluation within defined regions. Publication will require explanations of data quality issues and limitations in making comparisons.

Working with the IACR in such activities, the IARC-coordinated GICR (http://gicr.iarc.fr) represents an innovative solution to support local planning and advancement of population-based cancer registries, particularly in economically-transitioning countries already experiencing a rapid rise in cancer patients. As a global partnership, the GICR represents a unified plan of action to develop registries and has secured the commitment of leading international organisations to work together to help previously underserved countries to use data to inform their cancer control planning. The WHO has agreed to use the GICR as a tool to support Member States in addressing the morbidity and mortality targets and indicators within the Global Monitoring Framework. A series of IARC Regional Hubs have been established to assist countries to adopt proven global standards to their local context. The key objective of the GICR is to provide measurable improvements in cancer registration in over 20 LMIC by 2020 and a further 30 by 2025.

This article has provided an overview of the evaluation processes in the present volume of *Cancer Incidence in Five Continents,* described some of the contemporary variations in cancer risk observed in registry populations, and commented on the evolving status of population-based registration worldwide linked to new opportunities to ensure their sustained support. It can be foreseen that the compilation of cancer registry datasets of high quality, and a continued expansion of incidence data in LMIC, will ensure CI5 remains as an incredibly rich resource of cancer data for cancer control and cancer research purposes over the next 50 years, as it is at present.

Special Section Paper

APPENDIX

Comparability

Determining the extent of the comparability of a cancer dataset requires consideration of the registry's procedures, including the standards and definitions used in registration, and their adherence to established international standards and guidelines, including:

The system used for classification and coding of neoplasms. Registries were asked to submit their data using ICD-O-3 and verify, and where necessary, correct their data prior to submission, converting coding from other systems to ICD-O-3.

The definition of incidence. What constitutes a cancer case at the registry is especially important in evaluating comparability. The call for this volume requested data on all primary tumours, including, if collected, basal cell and squamous cell carcinomas of skin, nonmalignant tumours of the central nervous system and urinary bladder. Registries were asked on the frequency of incidental diagnosis through the detection of cancers as a result of a screening examination, or at autopsy, and to state whether standard, hierarchal rules or other in-house rules, were being used in defining:

Incidence date. Given the long natural history of cancer from the first mutation to clinical diagnosis, a common definition of cancer is sought, in deciding whether to register the case, and of the actual date when the disease became incident: registries tend to use one of several available algorithms (see Refs. 4, 5).

Multiple primaries. Individuals may develop more than one cancer, and there must be a clear distinction between those that are new cases (and counted as "incident" cancers) and those that represent an extension, recurrence or metastasis of an existing one. Cancer registries have adopted various rules for this purpose (see Refs. 4, 5).

Completeness

If incidence rates are to approximate their true values, completeness—the degree to which *all* diagnosed neoplasms within a registry's catchment population are included in the registry database—must be maximised. The methods used

to evaluate overall completeness can be considered *semi-quantitative*, as they provide an indication of the degree of completeness relative to other registries, or over time. They can be categorised into the following:

Stability of new cases and incidence rates over time. The distribution of new cases registered, by site and year of registration provides a visual check on the stability of numbers of cases (and rates) being recorded each year. Instability suggests potential problems in the registration process (or the source population data) in the period under review.

Comparison of incidence rates over time, or compared with other populations. Significant changes in rates from successive volumes may imply changing levels of case ascertainment at the registry, if they cannot be ascribed to artefact (*e.g.*, major diagnostic interventions or changes in estimation of person-years at risk). Statistical comparisons of incidence rates with regional standards (registries in the same region covering populations similar geographically and/or ethnically) may reflect local variations in the prevalence of risk factors, or the intensity of screening for some cancers. Systematic discrepancies (observed across cancer sites, in both sexes) may point to possible under- or over-registration (the latter due, for example, to the inclusion of duplicate records).

Age-specific incidence curves. Graphical description of rate-by-age profiles for 12 cancer sites by sex may detect abnormal fluctuations in the anticipated patterns, including any drop-off in the rate of increase in incidence in older subjects (suggestive, among other explanations, of under-ascertainment in the oldest age groups). The curves may also reveal problems with the source files used as population denominators for specific age groups.

Childhood cancer incidence rates. Incidence rates (for all types combined) within the childhood age groups (0–4, 5–9 and 10–14) tend to exhibit much less variability than in adults. The possibility of under-enumeration (or duplicate registrations) in this age range can therefore be investigated on comparing these rates with the lowest and highest deciles in the previous volume.

Proportion of cases morphologically verified (MV%). A very high proportion of cases diagnosed microscopically (histology or cytology/haematology), higher than might reasonably be expected (based on a comparison with a local standard), suggests over-reliance on the pathology laboratory as a source of information, and failure to find cases diagnosed by other means.

Mortality:incidence (M:I) ratios. The M:I ratio is an important indicator of completeness and involves comparison of the number of deaths, obtained from a source quasi-independent of the registry (usually, the vital statistics system), and the number of new cases of a specific cancer registered, in the same period of time. As with MV%, the standard used to evaluate the registry M:I ratio must reflect local conditions, and survival and the quality of mortality statistics are somewhat related to the level of socioeconomic development. M:I values greater than expected lead to a suspicion of incompleteness (incident cancers missed by the registry), especially if it is so for several different sites. However, under or over-reporting of tumours on the death certificates will distort the relationship, as will a lack of constancy in incidence and case fatality (the rate of death amongst incident cases) over time. Application of this method does require, however, mortality data of good quality especially with respect to accurate recording of cause of death.

Death certificate methods. Access to death certificates is important to cancer registries as a means of capturing information on cases that escaped the registration process during life. Cases registered from "Death certificate only" (DCO) are those cancers for which no other information other than a death certificate mentioning cancer could be obtained, *e.g.*, the residuum of cases—after all trace-back manoeuvres have been completed. As a proportion, the DCO% is not an indicator of completeness of registration: while a high DCO% may be indicative of an over-reliance on death certificates as a primary source, a low DCO% can result from the efficient trace-back of cases that were first notified by a death certificate.

Validity (or accuracy)

Several indicators of validity, the proportion of cases recorded as having a given characteristic that truly do have that attribute, relate to the precision of the source documents at the registry, and the level of expertise in abstracting, coding and recoding cases. They include internal consistency methods, diagnostic criteria methods (histological verification and death certificate only) and missing information analyses (primary site unknown, age unspecified).

Internal consistency. Registries were asked to verify and correct their data using this or other software tools prior to submission and ensure the ICD-O-3 coding system was used for all relevant variables.

Histological verification. In most instances, the accuracy of the stated diagnosis is likely to be higher if it is based on histological examination by a pathologist. Many cancer registries code diagnoses from exfoliative cytology or haematological examination of peripheral blood in the same category as histological exams, so the index of validity concerns the percentage of cases microscopically verified or MV%.

Death certificate only. DCOs are another measure of validity, since the information on death certificates is well known to suffer from lack of accuracy, or lack of precision, compared with that obtained from clinical or pathology records. The DCO% is sensitive to local circumstances, including availability and accuracy of death certificates, as well as the facility to trace-back cases and success in record linkage.

Other and unspecified/age unknown. The proportion of registered cases with unknown values for various data items

can be an indicator of data quality. Unknown values can result from problems with the data collection system, or access to necessary source documents, the item and code values that are defined, or the misapplication of coding rules. The definitions used will influence the proportion of unknown codes, for instance, when evaluating cases with primary site coded as "Other and Unspecified (O&U)." Other variables for which the proportion of cases with missing values are commonly evaluated include age, ethnicity and stage. A high proportion of cases assigned to these rubrics generally implies poor diagnostic precision (as evidenced by the low MV% observed for this rubric), or failure to specify the site of the primary cancer in cases diagnosed on the basis of tissue obtained from a metastasis.

Prior to formal editorial consideration, the editors carried out an extensive process of verifying coding, identifying duplicate registrations, querying unlikely or impossible combinations of codes and converting the data to a standard format. The review process then routinely applied to the evaluation of the 521 submitted datasets, *via* the scrutiny of preassembled registry-specific tables and other documentation. This included site-specific case numbers, age-specific rates summary rates (crude, cumulative and age-standardized), the populations at risk by sex and age, and a comparison with the previous 5-year population data (where available). The completed questionnaires provided further details including the definitions used by each registry.

For several of the quantitative indices introduced below, a comparison with "a gold standard" was performed. In most cases, the standard is derived regionally, from values from cancer registries within the same region, or nationally, if there are a number of high-quality registries within a country based on published data from the previous two volumes of *Cancer Incidence in Five Continents*. The use of regional values recognises that diagnostic practices (especially with respect to histology and cytology), and the accuracy of recording the underlying cause of death on death certificates varies between populations and regions. One might also reasonably assume that the incidence rates for specific cancers will tend to be rather similar in datasets from the same region. In total, 25 regions or countries were defined. For each, the mean and variance of values of the site-specific age standardised incidence, proportion of cases microscopically verified and M:I ratios were calculated from the contributions to Volume VIII and IX.

The extended regions and countries used to support CI5 editorial decisions are provided in the Chapter 5 of the CI5 Volume.[2] An Annex of the IARC Technical Report also provides tables with standard values of MV%, M:I and ASR for low or middle income countries or regions, selected for use in the CI5 editorial process.[6] In addition, *ad hoc* tables were used to identify unusually high or low cancer incidence rates in specific regions for all sites combined and for major cancers. This allowed the editors to assess the completeness by identifying outliers or unusual patterns.

References

1. Doll R, Payne P, Waterhouse J. Cancer incidence in five continents: a technical report. New York: International Union Against Cancer/Springer-Verlag, 1966.
2. Forman D, Bray F, Brewster DH, et al. Cancer incidence in five continents, Volume X. Lyon; Geneva: International Agency for Research on Cancer; Distributed by WHO Press, World Health Organization, 2014.
3. Waterhouse JAH, Peacham D, Davis W. Cancer incidence in five continents volume III. Lyon: International Agency for Research on Cancer, 1976.
4. Bray F, Parkin DM. Evaluation of data quality in the cancer registry: principles and methods. Part I: comparability, validity and timeliness. *Eur J Cancer* 2009;45:747–55.
5. Parkin DM, Bray F. Evaluation of data quality in the cancer registry: principles and methods Part II. Completeness. *Eur J Cancer* 2009;45:756–64.
6. Bray F, Znaor A, Cueva P, et al. Planning and developing population-based cancer registration in low- and middle-income settings. Lyon: International Agency for Research on Cancer Technical Report No. 43. International Association of Cancer Registries; World Health Organization, 2014.
7. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* 2015;136:E359–86.
8. Doll R, Peto R. The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. *J Natl Cancer Inst* 1981;66: 1191–308.
9. Tomasetti C, Vogelstein B. Cancer etiology. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science* 2015;347:78–81.
10. Montague M, Borland R, Sinclair C. Slip! Slop! Slap! and SunSmart, 1980-2000: Skin cancer control and 20 years of population-based campaigning. *Health Educ Behav* 2001;28:290–305.
11. Chang MH, You SL, Chen CJ, et al. Decreased incidence of hepatocellular carcinoma in hepatitis B vaccines: a 20-year follow-up study. *J Natl Cancer Inst* 2009;101:1348–55.
12. Jha P. Avoidable global cancer deaths and total deaths from smoking. *Nat Rev Cancer* 2009;9: 655–64.
13. Franceschi S, Vaccarella S. Thyroid cancer: an epidemic of disease or an epidemic of diagnosis? *Int J Cancer* 2015;136:2738–9.
14. Soerjomataram I, Lortet-Tieulent J, Ferlay J, et al. Estimating and validating disability-adjusted life years at the global level: a methodological framework for cancer. *BMC Med Res Methodol* 2012;12:125.
15. White RE, Abnet CC, Mungatana CK, et al. Oesophageal cancer: a common malignancy in young people of Bomet District, Kenya. *Lancet* 2002;360:462–3.
16. Islami F, Kamangar F, Nasrollahzadeh D, et al. Oesophageal cancer in Golestan Province, a high-incidence area in northern Iran—a review. *Eur J Cancer* 2009;45:3156–65.
17. Znaor A, Lortet-Tieulent J, Laversanne M, et al. International variations and trends in renal cell carcinoma incidence and mortality. *Eur Urol* 2015;67:519–30.
18. Bray F, Jemal A, Grey N, et al. Global cancer transitions according to the Human Development Index (2008-2030): a population-based study. *Lancet Oncol* 2012;13:790–801.
19. Curado MP. Cancer incidence in five continents, Volume IX. Lyon; Geneva: International Agency for Research on Cancer; Distributed by WHO Press, World Health Organization, 2008.