The Diseases Population Index for Lung Cancer Incidence: How it is Calculated and Applied

Richard Grawath

Sole author:

Richard Grawath

Abstract

The author applies the concept of the Infectious Diseases Index to non-communicable diseases by using lung cancer incidence as an example. The obtained Diseases Population Index (DPI) is calculated in a way that even currently unknown risk factors are considered indirectly since incidence and whole populations are used in the calculation. Lung cancer incidence is understood as a product of risk factors which contribute to the clinical manifestation of the disease. Indirectly the mix and the interactions of the risk factors are also considered.

The calculation leads to values for risk with respect to lung cancer incidence. These are specific definitions and they create a framework in which certain developments of the disease can be monitored, compared and analysed on a global level.

The DPI for lung cancer incidence is calculated for selected countries and for WHO regions. Certain regions and countries build DPI clusters in particular levels which might indicate similarities with respect to lung cancer incidence. For men and for both sexes the European region is part of the DPI cluster in the high level while the African region is in the low level for both groups. This is in line with the age structures.
The proposed risk scenario combines a holistic approach with chaos theory.

Keywords: Diseases Population Index, Diseases Population Index for lung cancer incidence, risk assessment, epidemiology, epidemiological measurements, monitoring tool, setting priorities, public health

Background

The formula of the Infectious Diseases Index (IDI) uses the total number of cases and whole populations [1]. Incidence and whole populations are used in the calculation of the proposed Diseases Population Index (DPI) for lung cancer incidence.

Lung cancer and cancer deaths

In the USA lung cancer constitutes the majority of cancer deaths for both sexes [2]. Worldwide most cancer deaths can also be attributed to lung cancer and annually 1 million lung cancer deaths occur [2]. This is in line with an estimate from Cancer Research UK which states that lung cancer was the major cause of cancer deaths globally in 2008 [3]. In Japan for the first time the Ministry of Health named lung cancer as the leading cause of death in 1998 [4].

According to an estimate from Jemal et al (2008) in the US for women 26 % of all deaths caused by cancer are attributable to lung cancer [5].

Lung cancer incidence in developed and developing countries

In 2008 in developing countries the estimated age-standardised rate (ASR, per 100,000) for lung and bronchus cancer incidence is 27.8 for men and 11.1 for women [6]. For developed countries for men the respective value is 47.4 and 18.6 for women [6].

For 2008 the global incidence of lung and bronchus cancer is estimated to be 1,095,200 for men and 513,600 for women [6]. For the same year in developing countries the estimate for lung and bronchus cancer is 612,500 for men and 272,000 for women while in developed countries the estimate is 482,600 for men and 241,700 for women [6].

An analysis of this situation should consider the fact that the countries categorised as developing countries have a much higher share in the population of the world compared to

developed countries. According to Magner (2005) around 80 % of the world's population lives in developing countries [7]. Therefore a rate may not reflect the real extent of the problem in developing countries since a rate may be low but the absolute number for incidence might be high due to a large population.

Risk factors contributing to the development of lung cancer

In general in the majority of cancer cases it is likely that many factors contribute to the development of cancer [8]. It is not possible to identify the cause of the cancer for the majority of patients [8]. Many contributing factors have been identified for lung cancer [8, 9].

As described above risk factors exist in a complex and complicated environment with no direct path for calculating risk. Populations as well as individuals may be exposed to different risk factors and even if exposure to identical risk factors exist then the extent, duration and response may differ. Even calculating an estimate by summing up known risk factors might locate the risk in an unrealistic place since risk factors act differently on different individuals. There are different mixes of risk factors for different individuals even in the same environment. There might be unknown risk factors and there are also interactions between risk factors.

Interactions of risk factors

Risk factors interacting with other factors may disproportionally magnify the effects of a contributing factor. The effect of smoking can be magnified by other risk factors while these other risk factors also act independently [9]. Interaction between tobacco and alcohol has been studied for certain cancers [8, 10].

Carcinogens such as benzene and nickel may interact with cigarette smoke thus increasing the cancer risk [10].

Synergism between smoking and asbestosis exists and the risk for developing bronchial carcinoma is 5 times its original value compared to smoking alone [11].

A concept about the underlying causes for death has been proposed with the intention of highlighting the role of risk factors which are non-genetic and are major factors for death at a later point in time [12]. The use of tobacco belongs to these major factors which are behavioural and changeable [12].

Smoking

Globally incidence for lung cancer has changed due to changes in prevalence of smoking [13].

According to Mc Ardle and O'Mahony (2008) currently in the UK more than 80 % of all lung cancers may be due to smoking [14]. This is in line with estimates from Cancer Research UK (2013) [15]. Bello et al (2011) state that in general smoking may cause between 80 and 90 % of all lung cancer cases [13]. Macbeth and Morgan (2008) also state that lung cancer is mainly due to smoking and they put the number as high as about 90% [16].

This indicates that smoking is the most potent factor in the development of lung cancer.

Smoking is also implicated in the development of a wide range of other cancers including colorectal cancer, gastric cancer, oesophageal cancer (adenocarcinoma and squamous cell carcinoma), anal canal cancer, gallbladder cancer, pancreatic cancer, cervical cancer, ovarian cancer, vulval cancer, gestational trophoblastic disease, kidney cancer, bladder cancer, head and neck cancer [14].

Smoking may contribute to the development of a wide range of cancers and the degree to which smoking contributes to the developments of these cancers may vary but smoking as such is a very powerful factor in cancer initiation.

An earlier beginning of smoking may cause an increase of lung cancer incidence in the younger age groups. The lower the age at which people start smoking the higher the risk for lung cancer [17]. The Global Youth Tobacco Survey found that compared to boys adult men have a lower smoking prevalence in various African countries [6].

The times at which age boys and adult men start smoking, stop smoking, are passive smokers, may use tobacco in another form and the extent of exposure to other risk factors create an extremely complex situation. It is important to note that lung cancer takes long periods of time to develop and this time may even be several decades [9].

Smoking prevalence as it was 20 to 30 years ago is reflected in current incidence rates for lung cancer [3]. This means that lung cancer statistics may contain cases with vast time differences due to differences in the time at which exposure to risk factors such as smoking started and the cancer diagnosis is made.

The larger the number of years at which smoking ceased the smaller the risk [17]. Aspects of time apart from age, such as beginning to smoke, cessation of smoking and duration of passive smoking are powerful factors in the development of lung cancer.

Socio-economic status, age, race, place of residency and education level affect smoking prevalence [2].

Antismoking campaigns and the ban to promote smoking are steps in the right direction. Many countries have similar tasks in public health and similar solutions may be used [18].

Convergence of heath systems has been observed and globalisation and the internet aid this convergence [18]. The convergence in health systems is supported by the European Union (EU), the Organisation for Economic Co-operation and Development (OECD) and the WHO [18]. This convergence may contribute to improvements in health care data.

Passive smoking

Passive smoking has been identified as a risk factor for non-smokers and it has been underestimated in the past [19]. The risk to develop lung cancer is 20 % higher for healthy non-smokers when they are exposed to Environmental Tobacco Smoke (ETS) [10].
Future decades may show a more pronounced reduction in lung cancer rates due to decreasing smoking prevalence which occurred many years earlier [13]. Past public health activities and declining smoking prevalence have already influenced the rates for men but not for women because of past gender differences in the prevalence of smoking [13].

Risk factors other than tobacco use

Air-pollution, diet and occupational factors also play a role in the development of lung cancer [20]. People working in particular industries and trades have an elevated risk of developing lung cancer whereas so far scientific investigations have failed to provide evidence for risk factors of these particular industries [21]. It may be assumed that there are various unknown risk factors and unknown interactions, however these risk factors and interactions are part of the proposed holistic risk scenario.

Radon exposure greatly increases the risk of developing lung cancer regardless of the smoking status but smokers have a much higher risk for lung cancer to begin with and therefore the effects for smokers are bigger [15].

Based on the degree of radiation exposure the risk to develop lung cancer triples and may go up to 30 times the original value [20]. The average of a certain radiation dose received by members of a particular population during their entire live may be responsible or contribute to certain negative effects of radiation [22]. The detriment of radiation with respect to a population can be understood as the sum of effects and these effects include a reduction in vitality, increased mortality, increased morbidity, decreased life span and damage to genes [22].

The younger age groups

Oeberg et al (2010) list lung cancer as a possible outcome for adult passive smokers and not for children but this should be seen in the context of the long period of time required to develop lung cancer [19]. The mean of latency for lung cancer is between 16 – 17 years [20]. Children who are forced to be passive smokers have an elevated risk for cancer. According to Oeberg et al (2010) in 2004 globally second-hand smoke (SHS) exposure for non smokers was 35 % for women, 33 % for men and 40 % for children [19].

Children are in a different situation compared to adults when it comes to SHS since children have not completed the development of their bodies [23]. For children an elevated level of exposure exists compared to adults even in identically polluted environments [23]. This is due to child specific physiological facts such as increased air volume intake (compared to body volume) and increased frequency of breathing compared to values for adults [23]. Children are not the only group particularly vulnerable to SHS [23]. Embryos may have a higher risk for developing particular diseases later in life when smoking occured during pregnancy [23]. Duration of SHS as well as the dose play an important role when it comes to consequences with respect to health [23]. A specific vulnerability to SHS also exists for the elderly [23].

The vulnerability of children to SHS is of particular importance since children move to the higher age groups as they become older and they may carry what one may call a biological memory with respect to the exposure to risk factors such as SHS and their interactions. Therefore the particular size of the younger age groups as they move along the time line is of importance since the size of these age groups influences incidence data for lung cancer. The importance of the size for the age groups gets lost by creating standardized age groups for the use of about 30 years on a global level. It is an advantage of the DPI for lung cancer incidence that the calculation of the DPI indirectly considers group size and the real historic risk scenario by using whole population size and incidence. The total number of cases is used in the calculation of the Infectious Diseases Index [1].

Many risk factors, such as active smoking, passive smoking and occupational factors are in many cases connected to particular age groups. Kollárová et al (2002) state that lung cancer incidence has an upward trend in the majority of European countries and that the risk for developing lung cancer is on the rise for younger age groups [9].

Age

Age has many connections with other risk factors. During the development of numerous cancers the continuous confrontation with risk factors happens and accumulation occurs as individuals get older [24].

Graphically speaking the imposed age standardisation data are placed on top of the profoundly altered scenario of the remaining risk factors. Assuming that a difference between the age standardised population and the real population structure exists, then a certain degree of bias is introduced with the age standardised data which means that standardised age groups

have been moved either in the direction of smaller or bigger age groups compared to the data of a real population. This has implications for the calculation of measurements for risk since the standardized age groups have unrealistic hypothetical risk scenarios.

All risk factors and the historic mix of these factors are part of a historic reality in which the total number of cases and the incidence data have been produced by the historic mix of all risk factors and all specific interactions. The historic risk scenario includes even currently unknown risk factors and the known or unknown extent to which particular factors act upon each other and with the outer environment.

The assumption of mismatch between real and standardised age groups has even been stated by Ahmad et al (2001) who advocate the new standardised age groups of the WHO [24]. Assuming that the sizes of the real age groups and the standardised age groups do not match, using ASRs means that for each age group errors are created since a particular real age group is either smaller or larger compared to the standardized age group used. These errors add up within the standardized age groups of a country and also between countries when international comparisons are made. The spread of these types of these errors is likely to increase with the number of countries and their differences in age structures. In a discussion paper about the new WHO standard it is stated that the age standard used influences conclusions and comparisons [24]. Globally there are vast differences between the age structures of countries [24].

In 2002 the populations of China and India constitute about 37 % of the global population and they differ significantly in the median ages with 34 years for China and 25 years for India [25]. Adding Japan to this comparison increases the spread even further since Japan has

a median age of 44 years [25]. This means that for these three Asian countries alone a difference of about 20 years for the median age exists and at the same time they constitute a large portion of the global population.

In order to overcome differences and to account for developments over time and for certain data errors an average world population is construct based on estimates, age adjustments and error adjustments [24]. The new WHO standard population is an average calculated for all countries concerned beginning in 2000 and the WHO age standard is assumed to be an average of all the populations in question for the duration of its entire use of 25 or 30 years [24]. It is highly unlikely that any population of any country will match the WHO age standard at any point in time.

Age can be understood as an inseparable part of the quantitative and qualitative character of the risk scenario. This risk scenario is a dynamic system which is changed by removing age differences from the risk scenario. Age is not an independent factor with respect to the risk scenario and the respective age group. This dynamic system has intrinsic properties which change when artificial interference occurs.

The actual mix of risk factors and the concept of a holistic risk scenario

Incidence is influenced by the contributing factors causing the diseases.

The age groups and the risk scenarios can be understood as a holistic system. The holistic system is changing over time with respect to composition and internal and external interactions. This means that the risk scenario is not fixed and it is part of an open system. A holistic approach to understanding the risk scenario requires looking at the risk scenario in its

entirety and complexity. The complexity of this system is its development over time with various driving forces [26].

A system has certain properties such as the number of individuals of the population [26]. Outside influences should be considered since magnification of external influences through sensitivity to initial circumstances may occur in an open system [26]. The state of this system is determined by the values of the variables at a certain point in time [26]. The risk factors are variables. This system may be described as a chaotic system based on a definition of chaos in a scientific way. The processes of the chaotic system are deterministic [26]. The risk scenario fluctuates together with the particular age group as it moves along the time line.

According to Ahmad et al (2001) comparing populations and analysing developments over time may lead to distorted findings when these populations have different age structures and crude rates and specific age groups are used [24].

Besides age structures there are many other large and important differences between populations but these other differences are not considered appropriately by the technique of age standardization.

Large differences in smoking prevalence can be observed for the countries of the world and there are also vast differences in smoking prevalence between men and women [3]. The prevalence of smoking varies not only between different countries, it also changes over time [27, 28].

The prevalence of further risk factors such as obesity, alcohol consumption and exposure to ionizing radiation may also show large differences between populations.

The DPI uses a holistic approach by using historical risk scenarios which is expressed through the use of incidence and whole populations in the calculation.

The Diseases Population Index and other measurements

Different measures exist and these different measures are used for different tasks such as the evaluation of outcomes of interventions in health care, monitoring and assessing health of populations [12]. The burden of cancer can be measured with the help of crude data for incidence and mortality [6].

Rates are widely used in epidemiology but these rates are calculated for sections of populations such as 100,000 people or another specific number of people. Populations with vast differences in population size may have similar or identical disease rates thus the extent of a problem might not be demonstrated through the use of rates.

A high dependence on age can be observed for the majority of rates in epidemiology and this includes rates for incidence, mortality and prevalence [24].

Age adjusted calculations are performed in order to compare countries [29]. In the 1950s Finland had a population structure fairly close to that of the world standard population [29]. In Finland the weighting of the age groups has changed since the 1950s [29]. In 2009 the older age groups have a larger share of the total population [29]. This explains why in 2009 the adjusted rates differ vastly from the actual rates [29]. As a consequence any analysis

about developments over time using data where the real population structures differ significantly from the world standard population may be compromised. The population structures of countries can differ in either having a lower value for a certain age group as a standard population or having a higher value. This may lead to countries having unrealistic high or low values for standardized age groups.

Rates calculated with the help of a standardised age structure are purely theoretical [6]. International comparability is compromised when different standardised age structures such as the European Standard Population, the World Standard Population (1960) or the 2000 US Standard population are used [6].

Conclusions drawn are biased by the respective age standardised population chosen. The use of a standard population where too much emphasis is placed on particular age groups such as the younger groups puts a disproportional weight on events of these particular age groups [24].

Ahmad et al (2001) point out that effects of phenomena such as famine are eliminated when an average population for the world and observational time series are used [24]. In the real world age groups are affected directly and indirectly by unforeseen events and this may have significant implications for the population developments for many decades to come.
In this context it should be mentioned that diet can play an important role in the development of lung cancer [21]. As a consequence nutritional aspects, differences in age structures and their interactions are eliminated regardless of their importance for the historic lung cancer risk scenario. An analysis working with ASRs for lung cancer is left working with profoundly altered and biased risk scenarios where underlying risk factors and interactions have been averaged, put in different contexts or eliminated all together.

Ahmad et al (2001) defend the new WHO world standard population and state that current population standards are far too excessive [24]. Based on this statement one may point to the possibility that the new WHO world standard may also become too excessive at an unknown point in time during its intended period of use of up to 30 years.

European countries have aging populations similar in structure to the Finnish population structure. In Europe 15.2 % of the population is older than 65 while in Africa only 3.3 % of the population reach age groups above 65 [8]. The majority of cancers can be observed in people older than 65 [8]. This may have implications for the age adjusted rates where the world standard population is used since the vast majority of people live in developing countries. Even a low rate may lead to a large number of people in absolute terms.

According to Kendall, Stuart and Ord (1982) a weighted index in economics may be described as an artificial construct [30]. The decision whether or not an index represents reality depends on the interpretation of knowledge [30]. This also applies to the DPI for lung cancer incidence.

The hypotheses

The first hypothesis states that the DPI for lung cancer incidence behaves differently compared to ASRs. It is hypothesized that the DPI for lung cancer incidence has different values for countries which have identical or similar ASRs while these countries have large differences in population size and in incidence.

The second hypothesis assumes that countries with identical or similar population sizes and incidence have similar DPI values. According to this hypothesis the DPI is capable of showing

realistic differences or similarities in lung cancer incidence for men, women and for both sexes.

Methods

A genetic disease such as cancer may serve as an example for calculating the DPI for a non-infectious disease [31].

The DPI for lung cancer incidence is calculated in a similar way as the IDI [1]:
The incidence is squared and the result is divided through the total population. In the final step Log (X) is found for the result of the previous step [1].

For the Log (X) for men the incidence for men for lung cancer and the male population is used in the calculation. The same principle applies for the calculation of the Log (X) for women.

The selected countries in Table 1 can be categorised as developing countries or developed countries and this is based on a definition of economical development as used by the American Cancer Society [6].

Categories are created through the definitions of levels and these categories are summaries of large amounts of data. Ordinal data, such as these categories have the advantage that data are not only put into categories but also arranged in a certain order [32]. In a comparison between categories as such and ordinal data, ordinal data have the advantage of providing more information [32]

Levels for the DPI are clearly defined. The extremely serious, critical and high levels correspond to levels with elevated risk. Level descriptions and the index range for levels are given in Table 2 and Table 4.

Data for incidence in Table 1 and Table 3 are estimates and stem from Globocan2008 [33]. Data about population size stem from Globocan2008 [33].

In Table 1 an international comparison is conducted and in addition a comparison of incidence crude data, ASR (W) and DPI for lung cancer incidence is provided.

<<Table 1>>

Table 2 provides the definitions for the levels of the DPI and places the countries from Table 1 according to their DPI values in these levels.

<<Table 2>>

Table 3 contains a comparison between WHO regions and the world on the basis of incidence crude data, ASR (W) and the DPI for lung cancer incidence.

<<Table 3>>

In Table 4 WHO regions and the world are assigned levels according to the values of the DPI for lung cancer incidence.

<<Table 4>>

Table 2 and Table 4 are designed to provide an overview while the values of the DPI for lung cancer incidence can be used for a more detailed analysis. The DPI and the defined levels are a condensation and abstraction which may aid research, surveillance and formulation of health policy [1].

The Tables 1 – 4 show data for men, women and both sexes thereby enabling analysis about differences and similarities between the sexes.

In this article the calculation of the standard deviation, the mean and the median was performed only for selected countries, and this is a limitation which should be considered.

It is necessary to constantly improve data quality but there is no need for perfection of all data crumbs [34].

Further research is needed. Particular comparisons over time for consecutive years should be done so that trends of lung cancer incidence can be analysed. The capability of the Index to show trends has already been demonstrated for the IDI [1].

Results and discussion

Dancey and Reidy (2004) state that the bulk of scores is located in the vicinity of the standard deviation [35].

The standard deviations of the DPI for Lung cancer incidence in table 1 and in table 3 have similar values. Men have identical values for the standard deviations for the DPI in table 1 and in table 3. Neither in table 2 nor in table 4 are the critical or the extremely serious levels at the higher end of the DPI reached. At the lower end of the DPI the two lowest levels of the DPI are occupied in table 2 and in table 4, however in table 4 only the value for women in the WHO Africa region is in the very low level. Table 3 shows that the standard deviation for the DPI for women has the highest value of all values for the standard deviation of the DPI in table 1 and 3. The DPI values for women in table 3 have a wider spread around the mean compared to the DPI values for men and for both sexes in table 3. The spread of values for

the DPI for women in table 3 is also wider than the spread of the values for men, women and both sexes in table 1.

Taking the scale of the data into account the values of the standard deviations of the DPI are close together. The insensitivity of the median to outliers is advantageous [35]. In the comparison between countries in table 1 the values for the median for the DPI for men and for both sexes are similar while the value of the median for the DPI for women is significantly different. In comparison between WHO regions in table 3 the values for the median are also fairly close for men and for both sexes while again the value for the median for the DPI for women is very different.

In general an index is a measurement which is concise and plainly shows developments over time [36]. The comparison between countries and both sexes in table 1 shows that the DPI is a concise measurement. The DPI relates raw incidence data to the respective population which makes it a more realistic measurement compared to the other measurements in Table 1.

The concept of the DPI allows for obtaining a larger picture about diseases.

The DPI for lung cancer incidence can be understood as a framework which provides a systematic approach to analysis with respect to risk in relation to lung cancer incidence. The work within the DPI framework is based on certain assumptions and definitions.

Incidence alone does not provide a complete picture with respect to diseases for countries since large differences in population sizes exist. Rates calculated for prevalence, incidence, morbidity and mortality refer to samples of populations and might not show the magnitudes of problems.

The age group of 55 years and older has 58 % of the total incidence of cancer in the developing countries while the figure is 78% in the developed countries [6]. The differences in the age structures are the main reason for the discrepancy in estimated incidence between developed and developing countries [6]. The younger age groups are larger in developing countries and the cancer risk is much higher in the older age groups [6].

According to Ferlay et al (2010) globally the incidence rates for women for lung cancer are lower compared to the rates for men and there is the highest rate for women in Northern America while the lowest rate is in Middle Africa [37]. The values for the DPI also show this global pattern.

Women in the AFRO WHO region are placed in the very low level of the DPI and this reflects that smoking prevalence is low for adult women in numerous African countries [6]. Ethiopia and Nigeria have the largest populations in Africa and in these countries smoking prevalence for women is about 2% [6]. For adult men smoking prevalence is below 10% in Africa [6]. This corresponds to the DPI values for men and for both sexes in the AFRO region being placed in the low level of the DPI and women being placed in the very low level.

In their findings Oeberg et al (2010) state that exposure to second-hand smoke has the highest estimates for the European sub-region C (includes Hungary, Russian Federation) with 61 % for children, and 66 % for men and women, the Western Pacific sub-region B (includes China, Viet Nam) with 67% for children, 53% for men and 51% for women [19]. The sub-region B of the South-East Asia region (includes Thailand) with 53% for children, 32% for men and 56% for women [19]. This points in the same direction as the assignment of the

WHO European region and the Western Pacific region in the high level of table 4 and the South-East Asia region in the advanced level of the DPI for lung cancer incidence in table 4. Compared to all other regions the African regions D and E have the lowest estimates for exposure to second hand smoke for children, men and women [19]. This also points in the same direction as the placement of the WHO Africa region being in the low and very low level of table 4.

In North America lung cancer belongs to the cancers with a high frequency [6]. Currently the prevalence for smoking is almost identical for men and women in the US [2]. According to Jemal et al (2008) in the US incidence for lung cancer is 15% of all cancer sites for men, 14% for women and lung cancer is the second most common cancer for both sexes [5]. The values of the DPI for lung cancer incidence for men and women are in line with this smoking prevalence since the values for the DPI for both sexes are very close for the USA in 2008.

In 2008 China has an ASR (W) for men of 45.9 and Italy has an ASR (W) of 45.4 for lung cancer for men [33]. These two ASRs (W) are very similar despite numerous extremely large relevant differences between both countries. For the same year China has an estimated incidence of 351,713 while Italy has only 28,147 which means that China's incidence is about 12.5 times higher than Italy's [33]. China's male population of 697,555,000 is 24 times that of Italy's 28,980,000 males [33]. According to the WHO in 2008 China has 20% people below the aged of 15 while Italy has 14% and in the age group of people over 60 China has 12% while Italy has 26% of its population in this age group [38]. Considering China's extremely large population these differences of 6% and 14 % amount to millions in absolute numbers. Furthermore, in 2008 in Italy the median age is 43 while in China it is 33 thus adding another large difference in the population structures of both countries [38]. In addition

in the Globocan2008 ranking of the top five cancers based on case numbers Italy has lung cancer as number 4 for both sexes and as number 3 for men and for women [33]. In contrast to the Italian ranking in China lung cancer is number one for men, women and both sexes [33]. In 2008 the DPI Log (X) for lung cancer incidence for men in China is 2.2 while Italy has 1.4 for men. The DPI Log (X) in table 1 and table 3 reflects the magnitude the problem with respect to the incidence of lung cancer by considering the sizes of the populations and the actual incidences.

For China increasing incidence rates for lung cancer can be observed [13]. Differences between men and women in these rates can to a certain extent be explained by the delay of women who started smoking and generally this causes a delay of about 25 years for the incidence rates for women compared to the rates for men [13]. Based on Globocan 2002 female incidence in China accounts for a third of all new lung cancer cases for women globally in 2002 [25]. This may indicate a trend since the DPI for lung cancer incidence for women in China is in the advanced level for 2008.

In China there are about 300 million men who smoke, which is a figure almost as large as the entire population of the USA [39]. In China the number of male smokers alone is more than 10 times the entire male population of Italy (calculation based on data from [39] and [33]) This staggeringly high number of male smokers in China has an elevated risk of developing lung cancer. There may also be an increased risk of lung cancer for non-smokers through passive smoking. In table 1 this is expressed in the high DPI lung cancer incidence value for men and for both sexes in China.

According to data from Globocan 2002 there are some values for ASRs (W) which point in the same directions as the respective DPI values for 2008 [25]. A comparison between the USA, China, Singapore, Japan and India shows that the ASR (W) for women has the highest value for women in 2002 in the USA [25]. Table 1 shows that the DPI for women in the USA has the highest value compared to the values for China, Singapore, Japan and India. According to a trend analysis by Toyoda et al (2008) in Japan incidence for lung cancer from 1975 to 2003 in the category of all histological types there is an increase for men and for women [40]. In table 2 Japan is placed in the advanced level of the DPI for lung cancer incidence for men, women and both sexes thus showing a similar overall picture as indicated by Toyoda et al (2008) [40].

Further examples for vast differences in incidence and population size but similar ASR (W) values are Germany, Israel and Viet Nam in the ASR (W) for women with 16.4, 16.1 and 16.4 respectively [33]. Germany has a smaller female population compared to Vietnam and about twice the incidence for lung cancer for women in table 1 [33]. Therefore in table 1 the DPI for women in Germany is higher (0.7) compared to Viet Nam (0.1). In 2008 for both sexes the ASR (W) for France (Metropolitan) is 30.0 and for Iceland it is 30.5 [33]. Table 1 shows that the lung cancer incidence for both sexes for France (Metropolitan) is much bigger compared to the value for Iceland and the DPI for France is 1.2 while it is – 1.2 for Iceland thus taking differences into account.

In 2008 there are cases where countries have noticeable similarities in their population sizes, in their population structures and in lung cancer incidence but there are vast differences in their ASRs (W).

In 2008 in Hungary 15% of the population is under 15 years old and in Greece 14% are younger than 15 [38]. For the same year in Hungary 22% of the population is over 60 and in Greece 24% belong to that age group [38]. Greece has a median age of 41 and in Hungary the median age is 40 [38]. Greece and Hungary are also similar with respect to population size [33]. Hungary and Greece have identical ranks in all three categories, men, women, both sexes in the top 5 cancers ranged by case number in Globocan 2008 [33]. Table 1 shows that in 2008 the values for male incidence for both countries are similar however the ASRs (W) for men show a vast difference with 80.9 for Hungary and 52.2 for Greece [33]. The DPI for lung cancer incidence for Hungary is 0.9 while Greece has 0.8 for the same year. The similar 2008 DPI values for men for Hungary and Greece indicate that the DPI for lung cancer incidence considers the similarities between both countries as described above and that the DPI values provide a realistic scenario.

In table 3 ‘World’ has the highest value for Log (x). Risk is not simply the product of the addition of the single contributing factors. The way the DPI is calculated emphasises that the risk profile is complex and this emphasis is achieved by the particular use of incidence in the calculation of the DPI. Incidence can be understood as the function of an historical situation in which many factors contributed and interacted. The risk profile may be dynamic and this dynamic may apply to individuals and populations.

For both sexes Denmark, Greece, Hungary, India, Thailand and Viet Nam form a DPI cluster in the moderate level. A DPI cluster is a group of countries or regions which fall into the same DPI level. This may indicate similarities with respect to the object under investigation, such as lung cancer incidence.

The European Region, the Region of the Americas, the Western Pacific Region and Global constitute a DPI cluster in the high level. This is in line with the enormous impact of lung cancer in these regions.

Behera and Balamugesh (2004) state that globally almost 70 % of lung cancer incidence occurs in developed countries [41].

The fact that the European Region is in the high DPI level and the African region is in the low DPI level is in line with the different age structures and the higher occurrence of cancer in the higher age groups (Gallagher, Lister, Johnson and Davies, 2009).

Results from Oelberg et al (2010) lend support to the DPI values for Western Pacific, Europe Eastern Mediterranean and African region [19]. An analysis has to take into account that Oelber et al (2010) estimated exposure of non-smokers and possible outcomes such as lung cancer while the DPI values summarise data about the incidence of lung cancer [19]. Nevertheless there are significant similarities in trends.

Conclusion

The first hypothesis can be confirmed since table 1 shows that the DPI for lung cancer incidence behaves differently compared to ASRs. The DPI for lung cancer incidence has different values for countries which have identical or similar ASRs but have large differences in population size and in incidence. This has been demonstrated for varies examples, in particular for China and Italy.

The second hypothesis is confirmed through similar DPI values for countries with identical or similar population sizes and incidence. Greece and Hungary serve as examples in this respect. The risk scenario of the DPI eliminates the disadvantages of the standardised age groups where age is taken out of the risk scenario, the external and internal interactions are disturbed and the sizes of the age groups do not reflect reality.

The holistic risk scenario of the DPI considers the contributing risk factors and the calculation of the DPI considers the total population and incidence. The DPI for lung cancer incidence has the advantage that a realistic mix of risk factors is used in the calculation. The concept of the DPI for lung cancer incidence allows a calculation of risk for populations with respect to lung cancer incidence.

List of abbreviations

AFRO	WHO Africa region
ASR (W)	Age-standardised rate (World)
DPI	Diseases Population Index
EMRO	WHO East Mediterranean region
ETS	Environmental Tobacco Smoke
EU	European Union
EURO	WHO Europe region
IDI	Infectious Diseases Index
OECD	Organisation for Economic Co-operation and Development
PAHO	WHO Americas region
SEARO	WHO South-East Asia region
SHS	Second-hand smoke

WHO World Health Organisation

WPRO WHO Western Pacific region

Competing interests

The author declares that he has no competing interests.

Authors' contribution

RG is the sole author of the article.

Authors' information

RG completed a wide range of university qualifications:

Postgraduate Diploma in Medical Physics, M.Sc., M.B.A., Diploma in Psychology, Certificate in Health and Social Care, Certificate in Managing Care, Certificate in Promoting Health, Certificate in Reflective and Evidence-Based Practice, B.Sc. (Honours) Health Studies, B.Sc. (Honours) Health & Social Care.

These qualifications provide the base for his scientific and professional work as a consultant.

Acknowledgements

RG funded the article independently.

References

1. Grawath R, 2012, The Infectious Diseases Index: Its Calculation and Application, http://www.meadowfordsciencejournal.org accessed 01 Jun 2012

2. Eriksen MP, Green LW (2009) Reducing Tobacco Use. In: Scutchfield, FD, Keck, CW. Principles of Public Health Practice 3rd ed. Delmar, New York

3. Cancer Research UK (2011) Cancer Stats Cancer Worldwide http://publications.cancerresearchuk.org/cancerstats/statsworldwide accessed 27.01.2013

4. Japan Cancer Society, http://www.jcancer.jp/english/cancerinjapan/ accessed 09 Aug 2012

5. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Murray T, Thun. Cancer Statistics, 2008. CA Cancer J Clin 2008; 58:71-96

6. American Cancer Society, Global Cancer Facts & Figures, 2nd edn, 2011

7. Magner LN (2005) A History of Medicine. 2nd edn. Taylor & Francis Group, Boca Raton, FL

8. Gallagher CJ, Lister TA, Johnson MJ and Davies A (2009) Malignant Disease. In: Kumar P, Clark M (eds) Clinical Medicine, 7th edn. Saunders Elsevier, London: 449-498

9. Kollárová H, Janout V, Čížek L, Epidemiology of Lung Cancer. Biomed. Papers 2002, 146(2) 103-114

10. Granger Oberleitner M, Odle TG (2005) Environmental factors in cancer development – A guide to cancer and its treatments. In: Longe JL (ed) Gale Encyclopedia of Cancer 2. 2nd edn. The Gale Group Inc., Farmington Hills: 422-426

11. Holgate ST, Frew A (2003) Chapter 14 Respiratory disease. In: Kumar P, Clark M (eds) Clinical Medicine, 5th edn. Saunders Elsevier, London: 833-915

12. Roy,K Chen, Z & Haddix, A (2009) Prevention Effectiveness. In: Scutchfield, FD, Keck, CW.Principles of Public Health Practice 3rd edn. Delmar, New York

13. Bello B, Fadahun O, Kielkowski D, Nelson G (2011) Trends in lung cancer mortality in South Africa: 1995-2006. BMC Public Health 11:209

14. Mc Ardle O, O'Mahony D (2008) Oncology. Churchill Livingstone Elsevier, London

15. Cancer Research UK (2013) Cancer Stats Key Facts Lung cancer http://www.cancerresearchuk.org/cancer-info/cancerstats/types/lung/cancerstats-key-facts-on-lung-cancer-and-smoking
accessed 27.01.2013

16. Macbeth F, Morgan C (2008) 28 Lung. In: Hanna L, Crosby T, Macbeth F (eds) Practical Clinical Oncology, Cambridge University Press, Cambridge: 313-327

17. Doll R, Peto R (2003) Epidemiology of Cancer. In: Warrell DA, Cox TM, Firth JD (eds) with Benz EJ. Oxford Textbook of Medicine, 4th edn. Oxford University Press, Oxford

18. Blank RH, Burau V (2010) Comparative Health Policy, 3rd edn. Palgrave Macmillan, Hampshire

19. Oeberg M, Jaakkola MS, Woodward A, Peruga A, Pruess-Ustuen A (2010) Wordwide burden of disease from exposure to second-hand smoke: a retrospective analysis of data from 192 countries

20. Tanoue LT, Carter DC, Matthay RA (2005) Lung Neoplasms (Chapter 14) in: George RB, Light RW, Matthay MA, Matthay RA (eds): Chest Medicine, Essentials of Pulmonary and Critical Care Medicine, 5th ed. Lippincott Williams & Wilkins, Philadelphia: 316

21. Redlich CA, Balmes JR (2005) Occupational and Environmental Lung Disease (Chapter 13) in: George RB, Light RW, Matthay MA, Matthay RA (eds): Chest Medicine, Essentials of Pulmonary and Critical Care Medicine, 5th edn. Lippincott Williams & Wilkins, Philadelphia 289

22. Hendee WR, Ibbott GS, Hendee EG (2005) Radiation Therapy Physics,3rd edn. Wiley & Sons, Hoboken

23. Oeberg M, Jaakkola MS, Prüss-Üstün A, Schweizer C, Woodward A (2010) Second-hand smoke: Assessing the environmental burden of disease at national and local levels. WHO Environmental Burden of Disease Series, No. 18. World Health Organization, Geneva

24. Ahmad OB, Boschi-Pinto C, Lopez AD, Murray CJL, Lozano R, Inoue M (2001) Age Standardization of rates: A new WHO standard, GPE Discussion Paper Series; No. 31, EIP/GPE/EBD World Health Organization

25. McDonald M, Hertz RP, Pitman Lowenthal SW (2008) The burden of cancer in Asia, Pfizer Facts

26. Rickles D, Hawe P, Shiell A, A simple guide to chaos and complexity, J Epidemiol Community Health 2007; 61:933-937.

27. Yoshimi I, Ohshima A, Ajiki W, Tsukuma H, Sobue T. A Comparison of Trends in the Incidence Rate ò Lung Cancer by Histological Type in the Osaka Cancer Registry, Japan and in the Surveillance, Epidemiology and End Results Program, USA. Jpn J Clin Oncol 2003; 33(2)98-104

28. Jemal A, Chu KC, Tarone RE. Recent trends in lung cancer mortality in the United States. J Natl Cancer Inst 2001: 93: 277-83

29. Cancer Society of Finland, Finnish Cancer Registry, Cancer in Finland 2008 and 2009 Cancer Statistics of the National Institute for Health and Welfare (THL), http://www.cancer.fi/syoparekisteri/en/search/.
accessed 06 Aug 2012

30. Kendall M, Stuart A, Ord JK (1982) The Advanced Theory of Statistics, vol 3. 4th edn. Charles Griffin & Company Ltd, London

31. Iles RK (2009) Molecular Cell Biology and Genetic Disorders. In: Kumar P, Clark M (eds) Clinical Medicine, 7th edn. Saunders Elsevier, London: 19-52

32. Stewart A (2010) Basic Statistics and Epidemiology, 3rd edn. Radcliffe Publishing, Oxford

33. Ferlay J, Shin HR, Bray F, Forman D, Mathers C and Parkin DM. GLOBOCAN 2008 v2.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 10 [Internet]. Lyon, France: International Agency for Research on Cancer; 2010. Available from: http://globocan.iarc.fr, accessed 09 Sep 2012

34. Grain H, Procter P (2009) Using Health Data. Churchill Livingstone Elsevier, Chatswood

35. Dancey CP, Reidy J (2004) Statistics without Maths for Psychology. 3rd edn. Pearson Education Ltd, London

36. Pass C, Lowes B (1993) Collins Dictionary of Economics, 2nd edn. Harper Collins, Glasgow

37. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM (2010) Estimates of worldwide burden of cancer in 2008: Globocan 2008. Int. J. Cancer: 127, 2893 – 2917

38. World Health Organization. World Health Statistics 2010, http://www.who.int/gho/publications/world_health_statistics/EN_WHS10_Full.pdf

accessed 13 Jan 2012

39. Hecht SS (2011) Ch 13 Tobacco. In: DeVita VT Jr., Lawrence TS, Rosenberg SA, (eds) DeVita, Hellman, and Rosenberg's Cancer: Principles & Practice, 9th edn. University Press, Cambridge

40. Toyoda Y, Nakayama T, Ioka A, Tsukuma H. Trends in Lung Cancer Incidence by Histological Type in Osaka, Japan. Jpn J Clin Oncol 2008; 38(8) 534–539

41. Behera D, Balamugesh T (2004) Lung Cancer in India. The Indian Journal of Chest Diseases & Allied Sciences 46: 269-281

Table 1 Comparing Countries: Incidence, ASR (W) and the DPI for Lung Cancer Incidence in 2008

Country	Incidence (Crude Data) [a]			ASR (W) [a]			DPI Log (X) [b, c]		
	M, F	M	F	M, F	M	F	M, F	M	F
Canada	21,599	11,215	10,384	35.9	40.3	32.3	1.1	0.9	0.8
China	522,050	351,713	170,337	33.5	45.9	21.3	2.3	2.2	1.7
Denmark	4,177	2,203	1,974	38.3	43.3	34.6	0.5	0.3	0.2
France (Metropolitan)	32,430	24,236	8,194	30	47.8	14.7	1.2	1.3	0.3
French Polynesia	101	74	27	43.6	64.6	23.1	-1.4	-1.4	-2.2
Germany	49,869	34,799	15,070	28.1	42.4	16.4	1.5	1.5	0.7
Greece	6,667	5,540	1,127	29.3	52.2	9.5	0.6	0.7	-0.6
Hungary	9,049	5,998	3,051	52	80.9	30.7	0.9	0.9	0.2
Iceland	148	71	77	30.5	31.6	29.4	-1.2	-1.5	-1.4
India	58,567	47,010	11,557	6.6	10.9	2.5	0.5	0.6	-0.6
Israel	2,152	1,345	807	23.8	33.1	16.1	-0.2	-0.3	-0.7
Italy	36,640	28,147	8,493	26.7	45.4	11.4	1.4	1.4	0.4
Japan	86,818	60,531	26,287	24.6	38.7	13.3	1.8	1.8	1.0
Malta	139	116	23	17.9	32.9	5.9	-1.3	-1.2	-2.6
Russian Federation	56,767	46,520	10,247	25.5	55.2	7.0	1.4	1.5	0.1
Singapore	1,748	1,189	559	25.8	37.9	15.3	-0.2	-0.2	-0.9
Thailand	14,129	9,338	4,791	19.0	26.8	12.1	0.5	0.4	-0.2
United Kingdom	40,366	22,392	17,974	31.3	38.2	25.9	1.4	1.2	1.0
USA	215,021	114,691	100,330	42.1	49.5	36.2	2.2	1.9	1.8
Viet Nam	20,659	13,152	7,507	25.7	37.6	16.4	0.7	0.6	0.1
Min [c]	101	71	23	6.6	10.9	2.5	-1.41	-1.50	-2.59
Max [c]	522,050	351,713	170,337	52	80.9	36.2	2.31	2.25	1.80
Mean [c]	58,954.80	39,014.00	19,940.80	29.51	42.76	18.71	0.68	0.63	-0.05
Median [c]	21,129.00	12,183.50	7,850.50	28.70	41.35	16.25	0.80	0.81	0.15
Standard deviation [c]	116,447.80	76,813.08	40,589.21	9.70	14.04	9.74	1.05	1.06	1.13
Confidence intervals [c]	51,034.56	33,664.20	17,788.68	4.25	6.15	4.27	0.46	0.46	0.49

a) Source: Globocan 2008
b) Calculation based on data from Globocan 2008
c) Figures are rounded

M, F: both sexes M: men F: women
ASR (W): Age-standardised rate (World) per 100,000

Table 2 Levels Assigned for the DPI for Lung Cancer Incidence for 2008

Index Range for Levels [a]	Level [a]	Description of Level [a]	Countries in particular Levels [b]		
4 and above 4	**6**	**Extremely serious**	Not reached in the sample		
3 – 3.99	**5**	**Critical**	Not reached in the sample		
2 – 2.99	**4**	**High**	China, USA	China	
1 – 1.99	**3**	**Advanced**	Canada, France (Metro-politan), Germany, Italy, Japan, Russian Federation, UK	France (Metro-politan), Germany, Italy, Japan, Russian Federation, UK, USA	China, Japan, UK, USA
0 – 0.99	**2**	**Moderate**	Denmark, Greece, Hungary, India, Thailand, Viet Nam	Canada, Denmark, Greece, Hungary, India, Thailand, Viet Nam	Canada, Denmark, France (Metro-Politan), Germany, Hungary, Italy, Russian Federation, Viet Nam
-1 – -0.1	**1**	**Low**	Israel, Singapore	Israel, Singapore,	Greece, India, Israel, Singapore, Thailand
below -1	**0**	**Very low**	French Polynesia, Iceland, Malta	French Polynesia, Iceland, Malta	French Polynesia, Iceland, Malta

a: Index Range for Levels, Level, Description of Level as defined by the author

b: Countries located in levels as calculated DPI Log (x)for lung cancer incidence in table 1

Both sexes | Men | Women

Table 3 Comparing WHO Regions and the World: Incidence, ASR (W) and the DPI for Lung Cancer in 2008

	Incidence (Crude Data) [a]			ASR (W) [a]			DPI Log (X) [b, c]		
WHO Region	M, F	M	F	M, F	M	F	M, F	M	F
AFRO	16,923	12,475	4,448	4.1	6.6	2	-0.4	-0.4	-1.3
PAHO	307,482	172,900	134,582	27.4	34	21.9	2.0	1.8	1.6
EMRO	26,552	21,415	5,137	7.4	12	2.8	0.1	0.2	-1.0
EURO	417,554	310,368	107,186	28.2	48.1	12.7	2.3	2.3	1.4
SEARO	151,139	108,979	42,160	11	16.6	5.9	1.1	1.1	0.3
WPRO	688,102	465,653	222,449	31.1	44	19.1	2.4	2.4	1.7
Min [c]	16,923	12,475	4,448	4.1	6.6	2.0	-0.45	-0.41	-1.31
Max [c]	688,102	465,653	222,449	31.1	48.1	21.9	2.42	2.37	1.75
Mean [c]	267,958.6	181,965.0	85,993.7	18.2	26.9	10.7	1.24	1.24	0.45
Median [c]	229,310.5	140,939.5	74,673.0	19.2	25.3	9.3	1.56	1.47	0.86
Standard deviation [c]	236,433.3	161,698.5	78,037.3	10.9	16.0	7.8	1.10	1.06	1.24
Confidence intervals [c]	189,182.6	129,383.4	62,441.7	8.8	12.8	6.2	0.88	0.84	0.99
World	1,608,055	1,092,056	515,999	22.9	33.8	13.5	2.6	2.5	1.9

AFRO: WHO Africa Region
PAHO: WHO Americas Region
EMRO: WHO East Mediterranean Region
EURO: WHO Europe Region
SEARO: WHO South-East Asia Region
WPRO: WHO Western Pacific Region

a: Source: Globocan 2008
b: Calculation based on data from Globocan 2008
c: Figures are rounded
M,F: both sexes
M: men
F: women
ASR (W): Age-standardised rate (World) per 100,000

Table 4 Levels Assigned for the DPI for Lung Cancer Incidence for WHO Regions and the World

Index Range for Levels [a]	Level [a]	Description of Level [a]	Regions in particular Levels [b]		
4 and above 4	6	Extremely serious	Not reached in the sample		
3 – 3.99	5	Critical	Not reached in the sample		
2 – 2.99	4	High	EURO	EURO	
			PAHO	World	
			World	WPRO	
			WPRO		
1 – 1.99	3	Advanced	SEARO	PAHO	EURO
				SEARO	PAHO
					World
					WPRO
0 – 0.99	2	Moderate	EMRO	EMRO	SEARO
-1 – -0.1	1	Low	AFRO	AFRO	EMRO
below –1	0	Very low			AFRO

a: Index Range for Levels, Level, Description of Level as defined by the author

b: Countries located in levels as calculated DPI Log (X) for Lung Cancer Incidence for Regions in table 3.

Both sexes Men Women