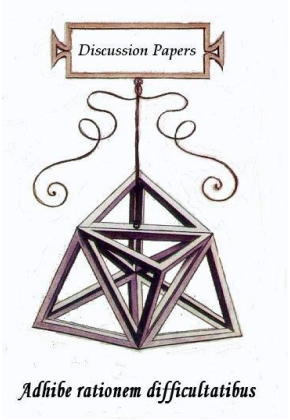
Discussion Papers

Collana di
E-papers del Dipartimento di Economia e Management – Università di Pisa



T. Luzzati* and T. Rughi

Is there a cancer Environmental Kuznets Curve?

Discussion Paper n. 208

2016

Indirizzo dell'Autore:

Tommaso Luzzati

Dipartimento di Economia e Management, via Ridolfi 10, 56100 PISA – Italy

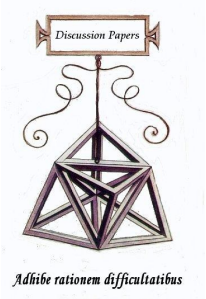
tel. (+39) 050 2216 329

fax: (+39) 050 598040

Email: tommaso.luzzati@unipi.it

web site: www.ec.unipi.it

Discussion Paper
n.



T. Luzzati* and T. Rughi

Is there a cancer Environmental Kuznets Curve?

Abstract

The observed increasing burden of cancer can be considered as good news, being the outcome of better life conditions and higher life expectancy. At the same time, changes in life-styles (e.g. diet, smoking, physical inactivity) and environmental quality brought about by economic development are also important risk factors in cancer. This piece of research aims at empirically assessing the role of economic development in new cancer cases (incidence).

Consistently with the literature on the Environmental Kuznets Curve (EKC), we adopt a coarse-grained approach rather than zooming into the very complicated determinants of the phenomenon under inquiry. The novelty is that we focus on impacts on humans rather than on pressures such as emissions or concentrations.

After reviewing the main statistical evidence and etiological hypotheses about cancer, we run several econometric models to assess the role of per capita income after controlling for life expectancy and diagnostic capacity. We investigated both aggregated cancers, and the most eight important site organ cancers. Data suggest that the increasing cancer incidence is also due to life-styles and environmental degradation.

KEYWORDS: Economic growth, Cancer, Cancer-EKC, environmental degradation, life-styles.

1. Introduction

The relationship between economic development and the environment is usually investigated by employing indicators of anthropic pressure. For instance, the Environmental Kuznets Curve (EKC) debate focuses mainly on emissions or concentrations of pollutants. However, the ultimate purpose of such analyses is envisaging the effects that human societies produce on their environment and, consequently, back on their wellbeing. In this paper, we will focus directly on one possible effect of the anthropic pressure, that is, the incidence of cancer.

Cancer is predicted to grow at alarming rates, particularly in lower- and middle-income countries[1] (e.g. Boyle and Levin, 2008, Stewart and Wild, 2016, Vineis and Wild, 2014). Although data availability on cancer has increased significantly in the last years[2], the possible relationship between the growing burden of cancer and economic development has been not yet fully investigated. Available evidence comes, at the best of our knowledge, mainly from the works by Beaulieu et al. (2009) and by Bray et al. (2012). The first is a report by "The Economist" Intelligence Unit aimed at identifying the health and economic burden of cancer. The data on cancer were taken from the GLOBOCAN[3] 2002 database that was compiled by the International Agency for Research on Cancer and that contains data for 26 unique site-specific cancers and for all sites cancer. The report estimates incidence, that is new cases, in 2009 and 2020. Next, it estimates the economic burden of incidence. Finally, it focuses on the costs associated with cancer control. In its appendix G, which is more relevant to the present paper, Beaulieu et al. (2009) present the results of a multiple regression analysis (OLS) aimed at understanding cross-country variation in 2009 estimated cancer incidence rates and case fatality rates. Regressors include per capita income, per cent of population ages 65+, and regional dummies.

The paper by Bray et al. (2012) focuses on cancer burden too. Based on the GLOBOCAN 2008 dataset, it is aimed at formulating predictions for 2030. It estimates a strong increase in the overall cancer burden and brings support in favour of the so-called cancer-transition, according to which the demographic transition and the economic development will change the composition of the different types of cancers, involving a shift from cancers linked to infections to those associated with non-infectious risk factors and possibly associated with

[1] For some rich Countries incidence is stabilizing or slightly decreasing. Absolute rates, are however high. In the US this occurs since the mid of the 90s, see Siegel et al., 2016.
[2] For an assessment of the status of population-based cancer registries worldwide, see Bray et al, 2015.
[3] See http://globocan.iarc.fr/Default.aspx for the scope, the methods, the data sources, and all details about the GLOBOCAN project.

western lifestyle. Actually, the paper highlights specific patterns of different sites of cancer depending on the level of Human Development Index. In particular, they analysed the data by grouping the countries into the four standard classes of the Human Development Indicator (low, medium, high, and very high).

The specific focus of the present paper is not on cancer burden, rather on the relationship between cancer incidence and economic development, an issue that is not at the core of the two above-mentioned studies. The different focus involves several differences in the methodology and research strategy, as we will see in detail. For this reason, one can interpret our results not only as updating[4] the available evidence but also as strengthening it.
The relationship between economic development and the environment is usually investigated by employing indicators of anthropic pressure. For instance, the Environmental Kuznets Curve (EKC) debate focuses mainly on emissions or concentrations of pollutants. However, the ultimate purpose of such analyses is envisaging the effects that human societies produce on their environment and, consequently, back on their wellbeing. In this paper, we will focus directly on one possible effect of the anthropic pressure, that is, the incidence of cancer.

Cancer is predicted to grow at alarming rates, particularly in lower- and middle-income countries[5] (e.g. Boyle and Levin, 2008, Stewart and Wild, 2016, Vineis and Wild, 2014). Although data availability on cancer has increased significantly in the last years[6], the possible relationship between the growing burden of cancer and economic development has been not yet fully investigated. Available evidence comes, at the best of our knowledge, mainly from the works by Beaulieu et al. (2009) and by Bray et al. (2012). The first is a report by "The Economist" Intelligence Unit aimed at identifying the health and economic burden of cancer. The data on cancer were taken from the GLOBOCAN[7] 2002 database that was compiled by the International Agency for Research on Cancer and that contains data for 26 unique site-specific cancers and for all sites cancer. The report estimates incidence, that is yearly new cases, in 2009 and 2020. Next, it estimates the economic burden of cancer incidence. Finally, it focuses on the costs associated with cancer control. In its appendix G, which is more relevant to the present paper, Beaulieu et al. (2009) present the results of a multiple regression analysis (OLS) aimed at understanding cross-country variation in 2009 estimated cancer incidence

[4] We used the 2012 data.

[5] For some rich Countries incidence is stabilizing or slightly decreasing. Absolute rates, are however high. In the US this occurs since the mid of the 90s, see Siegel et al., 2016.

[6] For an assessment of the status of population-based cancer registries worldwide, see Bray et al, 2015.

[7] See http://globocan.iarc.fr/Default.aspx for the scope, the methods, the data sources, and all details about the GLOBOCAN project.

rates and case fatality rates. Regressors include per capita income, per cent of population ages 65+, and regional dummies.

The paper by Bray et al. (2012) focuses on cancer burden too. Based on the GLOBOCAN 2008 dataset, it is aimed at formulating predictions for 2030. It estimates a strong increase in the overall cancer burden and brings support in favour of the so-called cancer-transition, according to which the demographic transition and the economic development will change the composition of the different types of cancers, involving a shift from cancers linked to infections to those associated with non-infectious risk factors and possibly associated with western lifestyle. Actually, the paper highlights specific patterns of different sites of cancer depending on the level of Human Development Index. In particular, they analysed the data by grouping the countries into the four standard classes of the Human Development Indicator (low, medium, high, and very high).

The specific focus of the present paper is not on cancer burden, rather on the relationship between cancer incidence and economic development, an issue that is not at the core of the two above-mentioned studies. The different focus involves several differences in the methodology and research strategy, as we will see in detail. For this reason, one can interpret our results not only as updating[8] the available evidence but also as strengthening it.

The paper is structured as follows. Section 2 will introduce some basic data and theoretical aspects of cancer. Section 3 will firstly discuss the possible links between cancer and economic development and then test it empirically with a cross-country econometric analysis. Section 4 will discuss the results and conclude.

2. Basic theory and facts about cancer

A useful concept to understand cancer is the "epidemiological transition" described by Omran (2005, 737-738). According to this scholar, three ages of mortality patterns in history are observed, respectively the age of "pestilence and famine", of "receding pandemics", and of "degenerative and man-made diseases". In the first "age" life expectancy at birth is very low, then epidemic peaks become then less frequent or disappear, eventually we enter a phase in which mortality tends to approach stability at relatively low levels and non-communicable diseases, including malignant neoplasms, prevail. The idea of epidemiological transitions can be used also within cancer changing patterns, to highlight a transition "from a predominance

[8] We used the 2012 data.

of cancers linked to infections to cancers associated with risk factors that are mainly non-infectious and possibly related to the so-called western lifestyle" (Maule and Merletti 2012, p. 745).

The identification of this "new epidemiological age" is not only a theoretical construct, but also a relevant empirical fact. Figure 1 highlights that, according to the World Health Organization (WHO 2014) about 52% of worldwide deaths in 2012 were due to Non-Communicable Diseases (NCDs) and, among them, about 27% were associated with Malignant Neoplasm. For the same year, the absolute level of cancer incidence rate (new cases) is 14.1 million (GLOBOCAN 2012), confirming an increasing temporal trend.

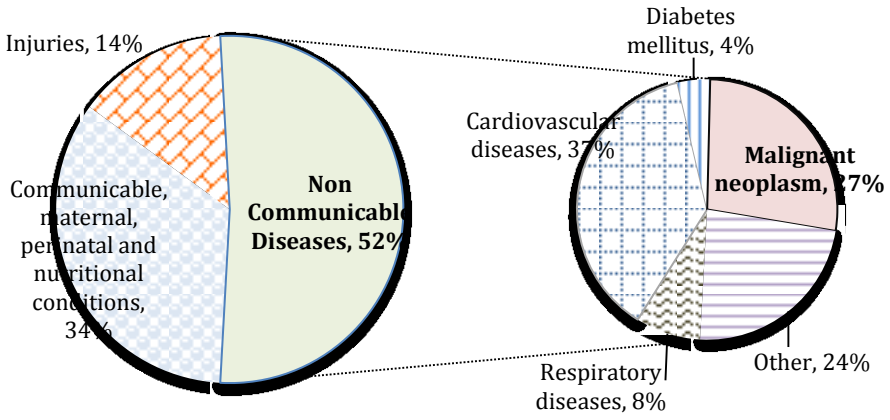


Figure 1. Causes of death under the age of 70 in 2012, and disaggregation by the different non-communicable diseases (Data: WHO 2014, p. 10).

For its intrinsic characteristics (see below), cancer increases with population ageing. Consistently, incidence is higher in countries with higher life expectancy. However, this holds true also when looking at standardised incidence rates, that is, rates that are recalculated[9] to compare populations with different age profiles, as is shown, e.g., by the pictures about trends in incidence of "cancer facts" of the GLOBOCAN 2012 webpage[10]. Hence, population ageing cannot fully account for the increasing cancer incidence rates. Actually, there is evidence that many risk factors are involved, among which also an increasing role of environmental factors. For instance, according to Prüss-Üstün et al. (2016, 16), 20% of cancer incidence is attributable to environmental factors.

[9] For details about standardization see section 3
[10] http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx

The role of an ageing population is easily explained by the simplest version of the so-called Somatic Mutation Theory (SMT)[11] according to which "random mutations in the genes which control proliferation or apoptosis are responsible for cancer" (Bertram 2001, 170). Many broader versions of SMT have been elaborated (see, e.g., Nowell 1976[12], or Burgio and Migliore 2015[13]). The best-known version is by Hanahan and Weinberg (2000, 2011) who highlight the basic features in a neoplastic cell due to accumulation of sequential mutations. In extreme synthesis, according to the SMT, cancer is due to stochastic (relevant) mutations that occur in oncogenes and tumour suppressor genes. The older a person, the higher is the number of accumulated of stochastic mutations, which ultimately leads higher probability of cancer occurrence.

However, SMT is only part of the story since it does not fully explain available evidence, among which one can mention the increase of standardised incidence rates, the significant differences at regional level, the increase of neoplastic pathologies in highly polluted areas, and the increase in childhood cancers (e.g. Steliarova-Foucher et al. 2004). The SMT has some increasingly acknowledged theoretical problems, among which Burgio and Migliore (2015) highlighted

- the lack of consideration by SMT of cell loss of differentiation, which seems one of the key points of cancer aetiology,
- the absence of correlation between a particular mutation and a specific ending neoplastic type,
- the "enormous prevalence of neoplastic processes affecting tissues persistently exposed to pollutants (skin, pulmonary, gastrointestinal, uterine epithelia) and/or solicited by continual exposure to these agents to react (above all immune and neuroendocrine systems)". (Ibid. 779)

Hence, cancer is acknowledged as stemming from many complex interacting factors, that is, from mutations in oncogenes and tumour suppressor genes, from genetic inheritance, from work and living environment, and from lifestyles (e.g. Belpomme et al. 2007a and 2007b, or Stewart and Wild, 2016).

[11] "The death of cells which occurs as a normal and controlled part of an organism's growth or development. Also called programmed cell death",
Oxford dictionary online http://www.oxforddictionaries.com/definition/english/apoptosis accessed 08/02/2016.

[12] "It is proposed that most neoplasms arise from a single cell of origin, and tumor progression results from acquired genetic variability within the original clone allowing sequential selection of more aggressive sublines. Tumor cell populations are apparently more genetically unstable than normal cells, perhaps from activation of specific gene loci in the neoplasm"(Nowell 1976)

[13] " In SMT carcinogenesis is generally conceived as a multistep process, including initiation, promotion and progression: a multifactorial pathology characterized by the accumulation of a multitude of genetic and cytogenetic alterations leading to malignancy" Burgio and Migliore (2015, p. 778).

The heritable factors have an important, but not exclusive, role. For instance, using data from Swedish, Danish and Finnish twin registries, Lichtenstein et al. 2000 reported that the genetic influence upon the incidence of cancer explains not more that 42% of the variance in incidence rate, depending on the cancer site. The estimates obtained by the authors are presented in Figure 2 that shows for each cancer site, the proportion of susceptibility to cancer attributable to heritable factors, to causes that are due to common family experiences and habits of the twins, and to non-shared risk factors.

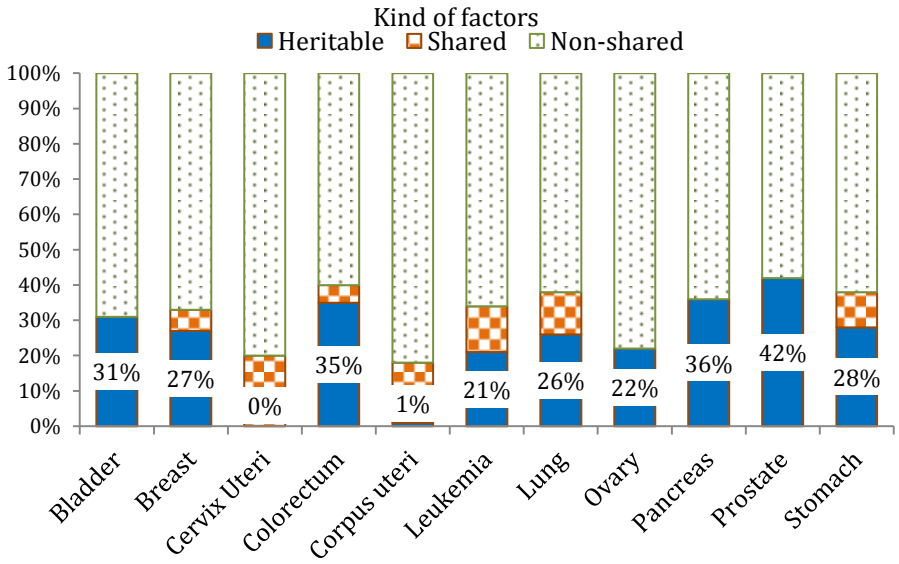


Figure 2: Effects of heritable, common environment and unique factors according to Lichtenstein et al. (2000)

Many studies attempt to estimate the differential contribution of non-genetic risk factors to cancer incidence (see, for example, Danaei et al. 2005). The factors on which the literature has focused are mainly overweight and obesity, diet, physical inactivity, smoking, alcohol overconsumption, unsafe sex, infections, occupational factors, urban air pollution, ionizing radiation and UV rays' exposition, and indoor smoke from household use of solid fuels.

Research also focuses on strictly environmental factors (see, e.g., Alavanja et al. 2003, Boffetta 2006). However, precise estimates of the role of the environment is very difficult since environmental pollution is a complex phenomenon. It is characterized by the multiplicity and variability of polluting agents, the presence of many complex interactions and synergies among agents, the persistence and ubiquity of pollution, the bioaccumulation and biomagnification of pollutant all along the food chain, the multiplicity of biological actions performed by each single agent.

To summarise, cancer is increasingly seen as the break of a complex equilibrium, that is, an evolutionary process in which random genetic mutations have to face the selection of environmental pressures, without forgetting intrinsic epigenetic plasticity, clonal evolution and high cellular adaptability (see Greaves 2016).

3. Empirically investigating the links between economic development and cancer

3.1. The theoretical framework

The previous section has summarised the basic facts about cancer, and the possible risk factors, among which the quality of the human environment. The focus of this paper is about the empirical relationship between income and cancer. Figure 3[14] sketches the causal chain from income growth to cancer.

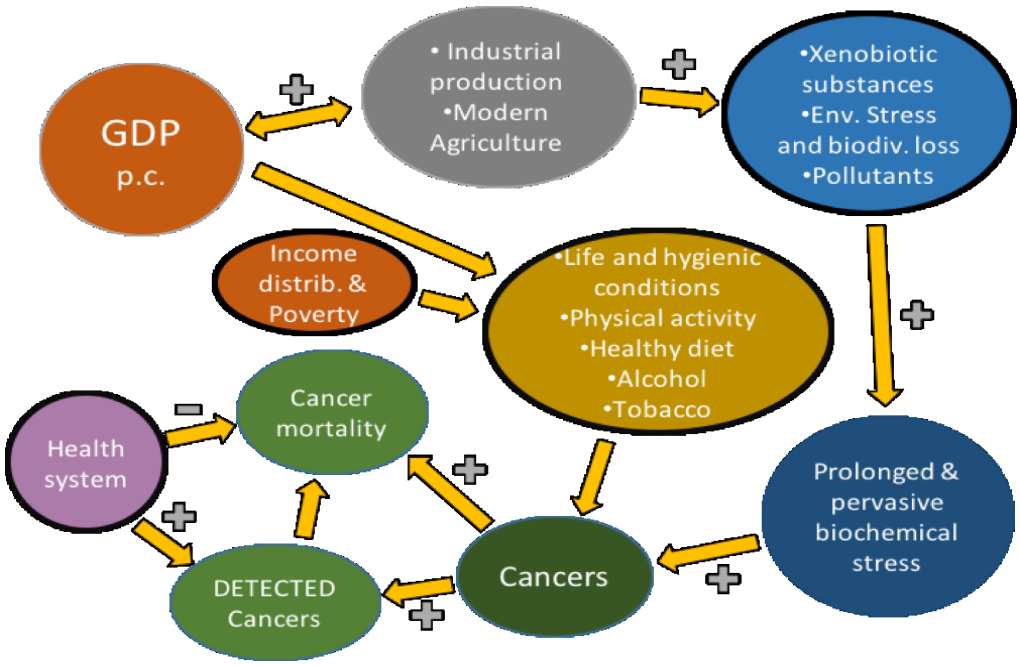


Figure 3: From income to cancer: possible links

Income growth has started with the industrial revolution, which was literally fuelled by fossil fuels. The availability of unprecedented quantity of energy radically transformed every industry, also agriculture that became highly mechanized and heavily relying on chemicals, pesticides, and resource use. A huge amount (and number) of pollutants has been populating the places where we live in, involving prolonged and pervasive biochemical stresses that have been found to be risk factors for several diseases, including cancer. Along

[14] Black borders of some ovals in Figure 3 indicate the domains where policy can intervene.

with the economic development process, life-styles evolved towards habits that have been proved as risk factors for many non-communicable diseases, including cancer. At the same time, life conditions have generally improved causing, on one side, reductions in cancers related to some infectious diseases, and on the other, increase in cancer due to higher life expectancy. Health policies (and expenditure) are obviously crucial in reducing mortality rates and can cause incidence and mortality rates radically to diverge, as occurs in high-income countries. However, health policies have also a role in cancer detection. In poorly developed health systems, both cancer statistics collection is poorly organized, and the causes of death remain often undiagnosed. For instance, Fallah and Kharazmi (2008) report under-registration of cancer deaths in developing countries. In other words, part of the increase in cancer incidence could be an artifact of improved diagnostic scrutiny (e.g. Li et al. 2013, Moynihan et al. 2012).

As stated in the introduction, our aim is to shed some light on the "GDP → environmental deterioration → cancer" casual chain. To do this, we focused on incidence rates, which exclude the highly complicated issue of the efficacy in fighting cancer of the health systems. Moreover, we controlled for life expectancy and for the capacity of diagnostic scrutiny. Data availability did not allow us to control for the risk factors associated with life conditions and life-styles. To sum up, our empirical analysis investigated the relationship between income per capita and cancer incidence rates after controlling for population ageing and for improved diagnostic capacity. Hence, our analysis falls within the EKC framework[15], involving a reduced model that looks only at the end of complicated casual chains.

3.2. Variables

Cancer incidence

At a global level, data are becoming increasingly reliable due to the diffusion of national cancer registries (see e.g. Parkin 2006). Still, national distinctions in coverage and quality are quite pronounced, resulting in high variability of coverage and reliability of the collected data. For a worldwide comparison, perhaps the most relevant project is GLOBOCAN[16], managed by IARC (an agency of WHO specialized in cancer) that produced the most recent estimates (2012) of incidence, mortality and prevalence. GLOBOCAN reports both

[15] The EKC literature investigates the overall effect of income growth on the environment and human well being and not at the various factors underlying the relationship (for a discussion of the EKC narrative see, e.g., Luzzati 2015).

[16] http://globocan.iarc.fr/Default.aspx (access 30/07/2016). The project now is incorporated inside Cancer Today.

the crude numbers of cancer and their rates. Since we study populations with different age profiles, the Average Standardized Rates (Weighted), *ASR(W)*, has to be used. The standardization procedure (for details see, e.g., Boyle and Parkin, 1991) adjusts observed age-specific rates to a reference population, commonly referred to as the Standard Population usually the world one[17]. This allows controlling for the differences in cancer figures that are merely due to different population. We calculated[18] *ASR(W)* using the population weight of World Standard Population[19] and the population data of the United Nations[20]. We used different incidence rates referred both to all cancers, and to the most common specific cancers. To exclude any effect of ageing, we did a further check by using as dependent variable not only the ASR(W) of the whole population, but also the ones of the population in the range 40-60 years,

Income per capita

Income per capita was the main explanatory variable. Data were taken from the World Bank online database. Since we perform a cross-country analysis, we used GDP per capita in Power Purchasing Parity (PPP2011)[21]. A three-years average of the GDP was taken in order to mitigate the effect of the business cycle. We used the 20 years lagged GDP to take into account of the long genesis of cancer[22].

Physician density

Physician density (Physician per 1000 inhabitants), taken from the World Bank online database[23], was used as the proxy for the diagnostic potential of a Country. While early diagnosis requires a high level (and expensive) health system, access to a doctor is often sufficient to detect a cancer. We will further discuss this issue below.

To summarize, the variables used in this empirical analysis are

- the *ASR(W)* of incidence for

[17] The term weighted refers to standard weights taken from the population adopted as a standard.
[18] GLOBOCAN 2012 gives already ASW(R) rates. Using the data available online and implementing the procedure described by the Glossary section of GLOBOCAN 2012 (http://globocan.iarc.fr/Pages/glossary.aspx) we got slightly different figures.
[19] http://seer.cancer.gov/stdpopulations/world.who.html. World Standard Population is used also in GLOBOCAN 2012.
[20] http://data.un.org/DataMartInfo.aspx Access
[21] PPP GDP is gross domestic product converted to international dollars using purchasing power parity rates. An international dollar has the same purchasing power over GDP as a U.S. dollar has in the United States. Data are in current international dollars based on the 2011 ICP.
[22] It might be claimed that a larger lag would be better (maybe 25 years or 30) to simulate the epigenetic nature or cancer, but data constraints forced us to find a compromise between the proper lag and sufficient data availability.
[23] We took the most recent data. Except for a few exceptions, data are recent, range from 2010 to 2013.

a. different site-specific cancers
   - *AllC* : all cancers, excluding non-melanoma skin cancer
   - *Name of site*: site specific cancers

b. different set of population
   - the whole population (no suffix)
   - population in the range 40-60 years (suffix *"_40-60"*)

- *Y_yr* , the three-year average, centred on the year "*yr*", of GDP p.c. (PPP2011)
- *PhysD* , the physician density in 2012 (number of physician every 1000 inhabitants)

3.1.Countries

The Globocan 2012 dataset covers 184 Countries. In our dataset we did not included those countries (33) for which data were estimated by merely inputing the data of neighbouring countries or registries in the same area.[24] Of the 151 remaining Countries, we had to exclude five that are not included in the World Bank online database[25]. We had also to exclude 18 countries for which 20 years lagged income or other data were not available. Finally, seven other countries have been considered as outliers since their performances cannot be mimicked by other countries. They have a disproportionate income levels due to very peculiar economies (based on oil or financial services) and/or their size is very small[26] (see also Figure 4). Hence, we ended up with 121 Countries[27].

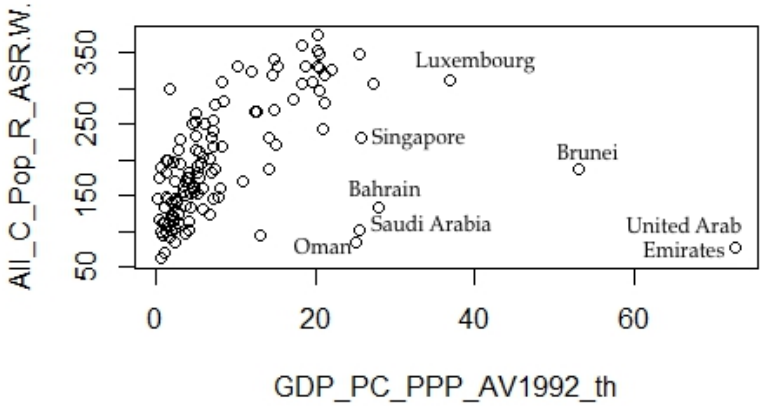


Figure 4: Outliers

[24] According to the metadata, this estimation method had the lowest "priority".

[25] State of Palestine, France: Guadeloupe, La Reunion, Martinique, and Guyana.

[26] They are Bahrain, Brunei, Oman, Saudi Arabia, United Arab Emirates, Luxembourg, and Singapore. Similarly to Luxembourg and Singapore, also Macao and Hong Kong are not included in our sample. In any case, they are not included in the GLOBOCAN 2012 countries.

[27] The appendix reports the list of Countries.

3.2. Data descriptive statistics

A preliminary overview of the data is given by Table 1, which reports the main descriptive statistics for the variables. Cancer figures refer to yearly *ASR(W)* incidence, that is, new cases on 100,000 inhabitants, GDP are expressed in dollars (PPP2011). Table 2 reports the correlation matrix for all-sites cancers, all-site cancers for the population in the 40-60 age class, income (and its lagged values), and physician density (also as natural logarithm). As expected the autocorrelation of income is remarkably high.

Table 1. Descriptive statistics for the variables used in the analysis

	Min	Max	Median	Average	Stand Err.
Incid_AllC	61.8	374.1	182.2	196.1	79.2
Incid._AllC_40_60	18.3	115.2	59.9	64.7	22.8
Lung	0.2	55.4	15.2	17.2	13.2
Lung_40_60	0.1	22.8	4.2	5.1	4.2
Breast	5.4	118.5	45.1	51.3	27.4
Breast_40_60	2.7	57.7	22.9	25.2	12.8
Colorectum	1.2	48.8	12.9	17.9	13.3
Colorectum_40_60	0.3	16.6	3.9	5.0	3.2
Prostate	1.3	144.4	31.1	44.4	37.1
Prostate_40_60	0	67.5	4.1	7.6	9.2
Stomach	0.8	45.4	7.3	9.7	7.6
Stomach_40_60	0.3	15.4	2.0	2.8	2.4
Liver	1.1	89.1	5.2	7.7	9.5
Liver_40_60	0.2	29.7	1.4	2.5	3.4
Cervix	2.3	86.7	17.1	20.9	15.3
Cervix_40_60	1.2	49.1	8.4	9.9	7.7
Oesophagus	0	27.9	2.6	4.3	4.9
Oesophagus_40_60	0	10.3	0.8	1.3	1.5
Y_92	250.8	27352.2	4767.5	7113.1	6675.5
Y_97	372.0	31530.0	5265.7	8604.7	8291.4
Y_02	490.0	38372.3	6695.0	10663.1	10389.2
Y_07	611.5	57992.8	9600.2	14335.4	13241.7
Y_12	770.5	64967.6	12086.9	16655.8	14474.9
PhysD	0.02	6.72	1.59	1.80	1.50

Table 2: Correlation matrix

	AllC	AllC 40-60	*Y_92*	*Y_97*	*Y_02*	*Y_07*	*Y_12*	*Phys*	*Ln. Phys*
AllC	1								
AllC 40-60	0.93	1							
Y_92	0.8	0.65	1						
Y_97	0.81	0.65	0.99	1					
Y_02	0.82	0.67	0.98	0.99	1				
Y_07	0.84	0.68	0.96	0.97	0.99	1			
Y_12	0.85	0.7	0.95	0.96	0.98	0.99	1		
Phys	0.71	0.67	0.63	0.6	0.62	0.66	0.67	1	
Ln.(PhysD)	0.68	0.63	0.59	0.57	0.59	0.62	0.64	0.85	1

3.3. Estimates

3.3.1. An overview of the estimates

Firstly, we focused on cancer aggregated over all sites, and then we moved to the eight most common site-specific cancers. For each cancer, we compared the OLS regressions for the incidence over all population with the ones over the population in the age class 40-60. This was done to check whether the age standardisation is sufficient to take into account the increase of incidence due to population ageing[28].

We used as regressors the 20 years lagged income and the density of physicians. Income entered the regressions also as squared and cubed in order to allow better functional forms (see, e.g., Van Alstine, Neumayer 2010, while a different estimation strategy is employing non-parametric or semi-parametric methods, as in Luzzati and Orsini 2009). We also checked for different time lags of income since, due to its high autocorrelation, income lags were not expect to affect substantially the estimates. As a preliminary step, for all-sites cancers incidence we also run a "pure" EKC regression, that is, without controlling for physician density.

[28] This might occur since in poor countries the older age classes might be too small to consider incidence rates reliable.

3.3.2. "Pure EKC"

The OLS regression of GDP per capita in 1992 on the incidence rate of all types of cancer for all age classes gave the following estimate[29]:

All_C_asrw	=	117	+	12.735	(*Y_92*)	-	0.0069	(*Y_92*)^3
	t =	14.82		8.59			-2.47	
	p	*<0.001*		*<0.001*			*<0.015*	

n=121, AdjRsq.=0.6457

Calculated turning point: y92=24.643$; min(y92)=251$, max(y92)=27.352$, mean(y92)=7.133$

When checking for all the different time lags, results do not change substantially[30]. Here we show only the estimates that one gets when using the 2012 GDP p.c.

All_C_asrw	=	109	+	5.756	(*Y_12*)	-	0.000478	(*Y_12*)^3
	t	15.61		11.12			-2.52	
	p	*<0.001*		*<0.001*			*0.013*	

n=121, AdjRsq.=0.7268

Calculated turning point: *y12*=63.343$; min(*y12*)=771$, max(*y12*)=64.968$, mean(*y92*)=16.656$

As expected, the strong autocorrelation of GDP makes the estimates with lagged and not lagged income very similar one another. To see this one has also to consider that GDP has increased over time, as reported below the regression results and in Table 1. The scatterplots (Figures 5a and 5b) showing the actual values and the estimated curve of both regressions confirm the irrelevance of changing the time lag.

3.3.3. Controlling for diagnostic improvements: physician density

The "pure" EKC framework was only a preliminary step since cancer incidence could be seriously affected by country differences in diagnostic and statistics collection efficacy. Actually, one can expect that, along the process of development, the improvement both in cancer statistics and diagnostic capacity involve higher incidence rates. A full assessment of the extent of this phenomenon goes far beyond the scope of this paper. To find a proxy of it is sufficient for the purpose of the present analysis. A possible candidate is the per capita health expenditure, taken in a concave specification because of its plausible "diminishing returns" in cancer incidence reporting. However, the almost perfect linear correlation between income and health expenditure does not allow including health expenditure among the regressors. According to the literature that investigates early cancer detection, physician density has proved to be very relevant (e.g. Ananthakrishnan et al 2010, Fleisher et al. 2008, Li et al. 2013,

[29] Regressions shown in the paper include only significant terms. Full results are available upon request.
[30] We checked the relevance of changing the time lag for all the regressions of our analysis.

Sundmacher and Busse, 2011). This holds for many other care issues, such as infant mortality (e.g. Farahani et al. 2009), and generally for health outcomes (e.g. Friedberg et al. 2010, Macinko et al. 2007, Mondal and Shitan 2014, Shi 2012).

As shown in Table 2 and by the scatterplot in Figure 6, the correlation between Physician density and GDP p.c. is not strong enough to prevent using both variables as regressors, particularly if one takes the income in 1992 and the natural logarithm of the physician density. Taking density in logarithm has also a theoretical motivation, that having easy access to a physician is presumably enough for cancer detection (and statistics collection) while further density increases over some thresholds do not contribute relevantly.[31]

When controlling for the density of physicians (in natural logarithms) the quadratic and the cubic terms are significant. The shape, however, is more or less the same as the "pure" EKC, and the turning points are close, respectively 23.542 and 24.643 $ 1992 (PPP2011). Figure 5c shows the actual values and the estimated curve of this regression, to be compared with the "pure" EKC estimate (Figure 5a). The OLS estimate is the following:

All_C_asrw	=	159.7	+	0.786	$(Y\_92)^2$	-	0.02226	$(Y\_92)^3$	+	18.97	Ln(*PhysD*)
	t =	26.812		5.238			-3.487			5.716	
	p	*<0.001*		*<0.001*			*<0.001*			*<0.001*	

n=121, AdjRsq.=0.7008 , Calc.. TP=23.542

3.3.4. Incidence in the age class 40-60

Focusing on the age class 40-60 (and controlling for physician density) does not change the outcome qualitatively. Again, the linear term is not significant, as shown by the following OLS estimate:

All_C_40_60	=	57.44	+	0.1539	$(Y\_92)^2$	-	0.004182	$(Y\_92)^3$	+	6.21	Ln(*Phys.D*)
	t=	26.394		2.806			-1.793			5.120	
	p	*<0.001*		*0.00588*			*0.07561*			*<0.001*	

n=121, AdjRsq.=0.5166. Calc.. TP=24.528

Figure 5d shows the actual values and the estimated curve for incidence in the 40-60 age class, to be compared with the estimate for all age classes (Figure 5c)

[31] Results do not change qualitatively if the physician density is used rather than its natural logarithm.

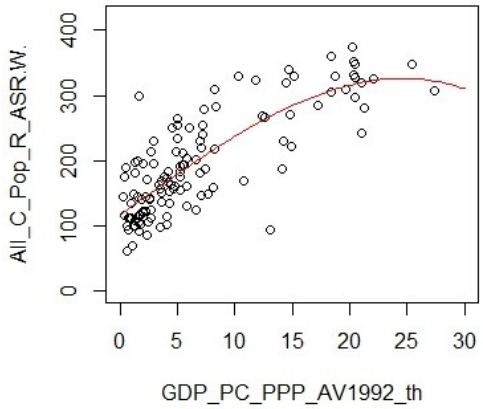


Figure 5a: ALL age classes vs. GDP p.c. in 1992

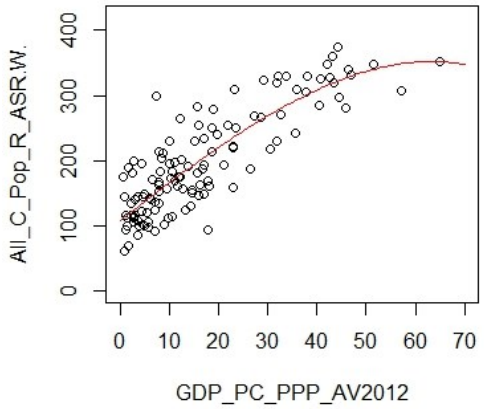


Figure 5b: ALL age classes vs. GDP p.c. in **2012**

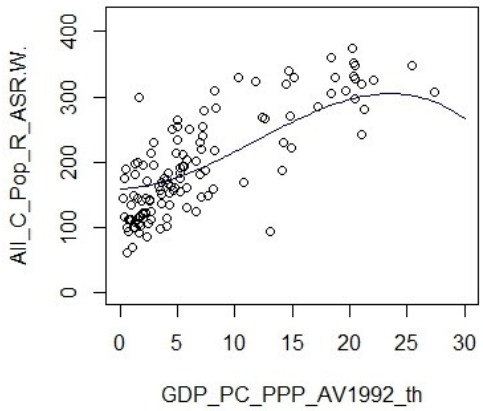


Figure 5c: ALL age classes vs. GDP p.c. in 1992, **controlling** for the density of physicians

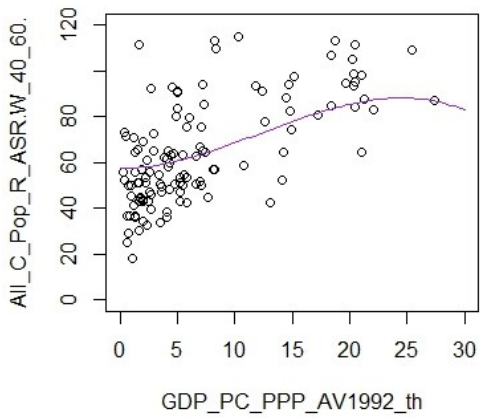


Figure 5d: **Age class 40-60** vs. GDP (PPP2011) p.c. in 1992, **controlling** for the density of physicians

Figure 5: All-sites cancer incidence rates vs. GDP (PPP2011) p.c.

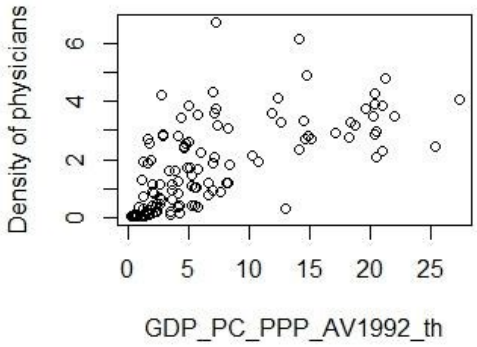


Figure 6: Physician density in 2012 vs. GDP per capita in 1992

3.3.5. Specific site cancers in the age class 40-60

We replicated the analysis shown so far for the most common site-specific cancers. Table 3 summarises the results. Cancers (second column) are ordered according to their frequency, which is reported in the first column. The third column indicates whether the row refers to all age classes or to the 40-60 one. The fourth and the fifth columns refer to the type of cancer diffusion, whether typical of low/medium income countries (L) or high ones (H). The fourth column indicates what is expected from the literature while the fifth what results from our estimates. The sixth column indicates whether the estimated relationship has a linear or an inverted-U shape. For the EKC shaped relationships see also Figures 7 and 8. The seventh column shows the estimated coefficients for income (labelled as '*x*'). The eighth column reports the sign of the coefficient of physician density[32] and the last column indicates the adjusted R-squared.

Table 3. Summary of the OLS estimates for the 8 most common organ-sites cancer

% of all cancers[33]	Organ Site	Age	Expe c-ted	Estim ated	Shape	Role of GDP (x)	Phys. Dens	Adj RSq
13.0	Lung	All	H	H	Linear (+)	$y=13.6+0.6272x$	+	0.57
		40-60		No	n.a.	N.S.	+	0.57
11.9	Breast	All	L & H	H	EKC	$y=36.7+0.33x^2-0.0098x^3$	+	0.65
		40-60		H	EKC	$y=19+0.147x^2-0.0045x^3$	+	0.58
9.7	Colo-rectum	All	H	H	EKC	$y=11.2+0.177x^2-0.00591x^3$	+	0.71
		40-60		H	EKC	$y=3.66+0.036x^2-0.0012x^3$	+	0.60
7.9	Prostate	All	H	H	Linear (+)	$y=14.5+4.21x$	n.s.	0.57
		40-60		H	Linear (+)	$y=1.9+0.79x$	n.s.	0.32
6.8	Stomach	All	L (?)	L	Linear (-)	$y=13.4-0.49x$	+	0.24
		40-60		L	Linear (-)	$y=4.17-0.18x$	+	0.23
5.6	Liver	All	L	L	Linear (-)	$y=9.8-0.303x$	n.s.	0.037
		40-60		L	Linear (-)	$y=3.4-0.13x$	n.s.	0.053
3.7	Cervix uteri	All	L	L	Linear (-)	$y=23.6-0.43x$	-	0.38
		40-60		L	Linear (-)	$y=11.6-0.24x$	-	0.41
3.2	Oeso-phagus	All	L	No	n.a.	n.s.	n.s.	-
		40-60		No	See Figure 9	$y=2.4-0.45x+0.0353x^2-0.000786x^3$	n.s.	0.064

[32] Rather interestingly, physician density is negatively correlated with the cervix uteri cancer. The increase in physicians not only helps improving cancer statistics but also is also associated with improvements in prevention, which is crucial for reducing the incidence of cervix uteri cancer.

[33] Data exclude non-melanoma skin cancers. See http://www.wcrf.org/int/cancer-facts-figures/worldwide-data

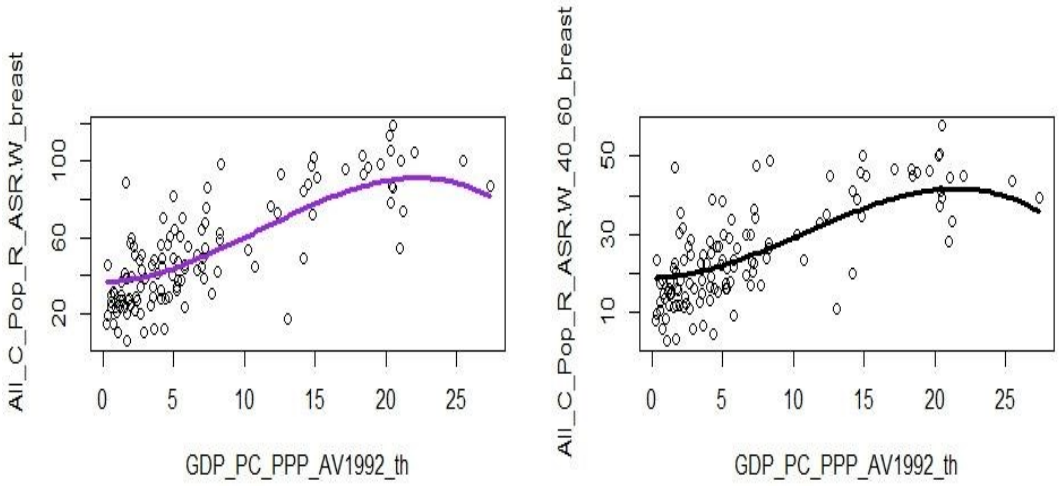


Figure 7. Breast cancer incidence vs. income respectively for all age classes and for 40-60 (controlling for physician density)

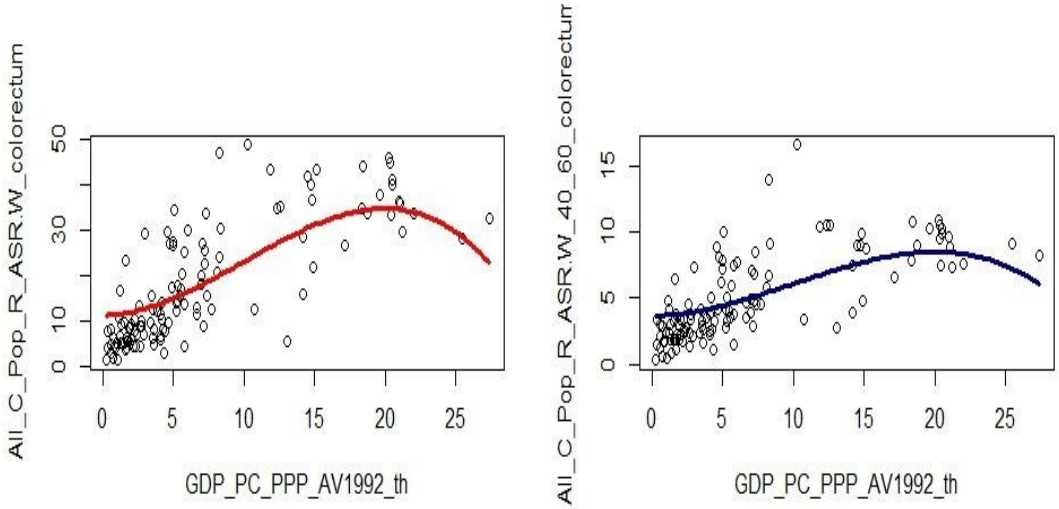


Figure 8. Colerectum cancer incidence vs. income respectively for all age classes and for 40-60 (controlling for physician density)

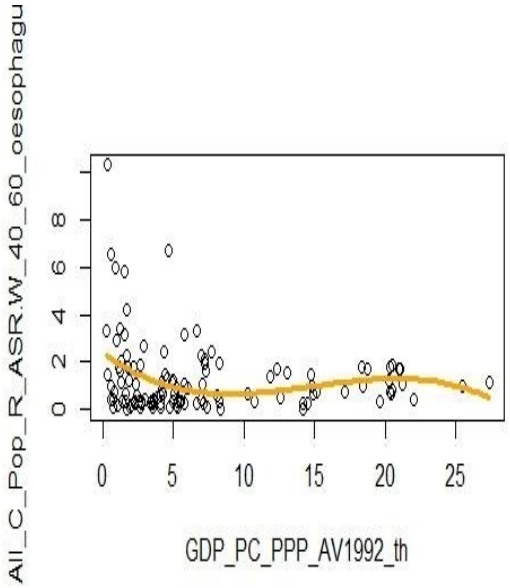


Figure 9. Oesophagus cancer incidence for 40-60 (controlling for physician density)

4. Discussion and conclusion

When looking at all-sites cancers, the estimates show that incidence rates sharply increase with per capita income, almost doubling from low to high-income countries. This result does not depend on the increase in incidence that accompanies population ageing since age standardised rates have been used. As a further check, we run regressions on the age classes 40-60, which did not change the evidence. By controlling for the number of physicians per person (physician density), it emerged that the increase in incidence with income does not seem an artifact deriving from the improvement in cancer diagnostics and statistics produced by economic development. Since the genesis of cancer is long, we used 20 years lagged values of income. Changing the time lags, however, does not change the evidence because of the strong autocorrelation of income.

There is also evidence that the relationship is concave, EKC shaped. The turning point, however, is towards the upper end of the income range and the incidence after the turning point remains very high. Moving to site-specific cancer allows understanding how a "week" EKC emerges at the aggregate level. We focused on the eight most frequent site-specific cancers, among which some are typical of high income and other of low-income countries. The first ones are positively correlated with income (or a "week" EKC shape), while the opposite holds for the second ones. The higher incidence rates of the "high income" cancer-sites make the overall relation increasing for most of the income range.

The results presented in this paper confirm the evidence already available from older data. Bray et al. (2012) derived their evidence from grouping the Countries into the four standard categories (low, medium, high, and very high) of the Human Development Indicator, which is very different from estimating a relationship from a continuous variable. The report by Beaulieu N. et al. (2009) tackles the issue of the relationship between development and cancer by using income and by performing a regression analysis. However, they used their estimated cancer data and controlled for the effect of population ageing by including the per cent of population ages 65+.[34] Being the focus of this paper the relationship between cancer incidence and income development, rather than cancer burden as the just mentioned studies, we hope it will not be deemed presumptuous for us affirming that the results of this paper not only confirm but also strengthen previous evidence.

[34] The report does not explicitly state whether the regressions (appendix G) were run with age standardised incidence rates.

The empirical positive relationship that has been highlighted in this paper between income and cancer incidence is not due to population ageing and to improvements in statistics involved by income growth. This confirms what is maintained by much theoretical and epidemiological literature, that an important role in cancer occurrence has to be attributed both to the changing lifestyles and to the deterioration of environmental quality.

5. References

Alavanja M. C., Samanic C., Dosemeci M., Lubin J., Tarone R., Lynch, C. F., Knott C., Thomas K., Hoppin J.A.,Barker J., Coble J., Sandler D. P. , Blair A. (2003). Use of agricultural pesticides and prostate cancer risk in the Agricultural Health Study cohort. *American Journal of Epidemiology*, *157*(9), 800-814.

Ananthakrishnan, A. N., Hoffmann, R. G., & Saeian, K. (2010). Higher physician density is associated with lower incidence of late-stage colorectal cancer. *Journal of general internal medicine*, *25*(11), 1164-1171.

Beaulieu N., Bloom D., Bloom R., and Stein, R. (2009). Breakaway: The global burden of cancer—challenges and opportunities. *The Economist Intelligence Unit*.

Bertram J. S. (2001). The molecular biology of cancer. *Molecular aspects of medicine*, *21*(6), 167-223.

Belpomme, D., Irigaray, P., Hardell, L., Clapp, R., Montagnier, L., Epstein, S., & Sasco, A. J. (2007a). The multitude and diversity of environmental carcinogens. *Environmental research*, *105*(3), 414-429.

Belpomme D, Irigaray P, Sasco AJ, Newby JA, Howard V, Clapp R, Hardell L. (2007b) The growing incidence of cancer: role of lifestyle and screening detection, Int J Oncol.; 30:1037-1049, 2007

Boffetta, P. (2006), "Human cancer from environmental pollutants: the epidemiological evidence." Mutation Research/Genetic Toxicology and Environmental Mutagenesis 608.2 (2006): 157-162.

Boyle P, Levin B (2008). *World Cancer Report 2008*. World Health Organisation.

Boyle, P., and D. M. Parkin. (1991) "Statistical methods for registries." *Cancer registration: principles and methods* 95, 126-158.

Bray F., Jemal A., Grey N., Ferlay J., and Forman, D. (2012). Global cancer transitions according to the Human Development Index (2008–2030): a population-based study. *The lancet oncology*, *13*(8), 790-801.

Bray F., Ferlay J., Laversanne M., Brewster D.H., Gombe Mbalawa C., Kohler B., Piñeros M., Steliarova-Foucher E., Swaminathan R., Antoni S. and Soerjomataram I. (2015). Cancer incidence in five continents: inclusion criteria, highlights from volume X and the global status of cancer registration. *International Journal of Cancer*, *137*(9), pp.2060-2071.

Burgio E., Migliore L., 2015, "Towards a systemic paradigm in carcinogenesis: linking epigenetics and genetics." *Molecular biology reports* 42.4 pp. 777-790.

Danaei G., Vander Hoorn S., Lopez, A. D., Murray, C. J., Ezzati, M., & Comparative Risk Assessment collaborating group (Cancers) (2005). Causes of cancer in the world: comparative risk assessment of nine behavioural and environmental risk factors. *The Lancet*, *366*(9499), 1784-1793.

Fallah M and Kharazmi E (2008) Substantial under-estimation in cancer incidence estimates for developing countries due to under-ascertainment in elderly cancer cases. Cancer Letters 2008; 264: 250–255.

Farahani M, Subramanian SV, and Canning D (2009). The effect of changes in health sector resources on infant mortality in the short-run and the long-run: a longitudinal econometric analysis. *Social Science & Medicine*, *68*(11), 1918-1925.

Fleisher JM, Lou JQ, and Farrell, M. (2008). Relationship between physician supply and breast cancer survival: a geographic approach. *Journal of community health*, *33*(4), 179-182.

Friedberg MW, Hussey PS, and Schneider EC (2010). Primary care: a critical review of the evidence on quality and costs of health care. *Health Affairs*, *29*(5), 766-772.

GLOBOCAN 2012, Estimated cancer incidence, mortality and prevalence worldwide in 2012, WHO, http://globocan.iarc.fr/Default.aspx

Greaves M, 2016. "An evolutionary foundation for cancer control" in International Agency for Research on Cancer." in Stewart and Wild (2016), 337-342.

Hanahan D, Weinberg RA (2000). "The hallmarks of cancer." (2000)*, Cell* 7, 100: 57-70.

Hanahan D, Weinberg RA (2011) Hallmarks of cancer: the next generation. Cell 4(144):646–674.

Li N, Du XL, Reitzel LR, Xu L and Sturgis EM (2013). Impact of enhanced detection on the increase in thyroid cancer incidence in the United States: review of incidence trends by socioeconomic status within the surveillance, epidemiology, and end results registry, 1980–2008. *Thyroid*, *23*(1), pp.103-110.

Lichtenstein, P., Holm, N. V. Verkasalo, P. K. Iliadou, A. Kaprio, J. Koskenvuo, M., E. Pukkala,A. Skytthe, K. Hemminki. (2000). Environmental and heritable factors in the causation of cancer—analyses of cohorts of twins from Sweden, Denmark, and Finland. *New England journal of medicine*, *343*(2), 78-85.

Luzzati T., (2015) "Kuznets Curves", in Wright J.D., International Encyclopedia of the Social & Behavioral Sciences, 2nd edition, Vol 13., 144-149, Elsevier.

Luzzati T, and Orsini M. (2009), "Investigating the energy-environmental Kuznets curve." *Energy* 34.3: 291-300.

Macinko, J., Starfield, B., & Shi, L. (2007). Quantifying the health benefits of primary care physician supply in the United States. *International journal of health services*, *37*(1), 111-126.

Maule M., and Merletti F (2012). Cancer transition and priorities for cancer control. *The lancet oncology*, *13*(8), 745-746

Mondal,MNI., & Shitan, M. (2014). Relative importance of demographic, socioeconomic and health factors on life expectancy in low-and lower-middle-income countries. *Journal of Epidemiology*, *24*(2), 117-124.

Moynihan R, Doust J, and Henry D (2012). Preventing overdiagnosis: how to stop harming the healthy. *Bmj*, (e3502).

Nowell, P. C. (1976). The clonal evolution of tumor cell populations. *Science*, *194*(4260), 23-28.

Omran, AR "The epidemiologic transition: a theory of the epidemiology of population change." *Milbank Quarterly* 83.4 (2005): 731-757.

Parkin, Donald M. (2006) "The evolution of the population-based cancer registry." *Nature Reviews Cancer* 6.8: 603-612.

Prüss-Üstün A, Wolf J, Corvalán C, Bos R, Neira M (2016), *Preventing disease through healthy environments*. World Health Organization, Geneva, Switzerland.

Shi L. (2012). The impact of primary care: a focused review. *Scientifica*, *2012*.

Siegel, R. L., Miller, K. D., & Jemal, A. (2016). Cancer statistics, 2016. *CA: a cancer journal for clinicians*, *66*(1), 7-30.

Steliarova-Foucher, E., Stiller, C., Kaatsch, P., Berrino, F., Coebergh, J. W., Lacour, B., and Perkin, M. (2004). Geographical patterns and time trends of cancer incidence and survival among children and adolescents in Europe since the 1970s (the ACCIS project): an epidemiological study. *The Lancet*, *364*(9451), 2097-2105.

Stewart B. and Wild, CP (eds.) (2016). *World cancer report 2014*. World Health Organisation, Geneva Switzerland.

Sundmacher L, and Busse R (2011). The impact of physician supply on avoidable cancer deaths in Germany. A spatial analysis. *Health Policy*, *103*(1), 53-62.

Van Alstine J and Neumayer E (2010) “The environmental Kuznets curve”. In: Gallagher KP (ed.) *Handbook on trade and the environment,* E. Elgar.

Vineis P, and Wild CP (2014). Global cancer patterns: causes and prevention. *The Lancet*, *383*(9916), 549-557.

WHO (2014). *Global status report on noncommunicable diseases 2014*. World Health Organization, Geneva, Switzerland.

6. Appendix: list of countries

Albania
Algeria
Armenia
Australia
Austria
Azerbaijan
Bahamas
Bangladesh
Barbados
Belarus
Belgium
Belize
Bhutan
Bolivia
Botswana
Brazil
Bulgaria
Burkina Faso
Cameroon
Canada
Chile
China
Colombia
Congo Republic of
Costa Rica
Cote d'Ivoire
Cuba
Cyprus
Czech Republic
Denmark
Dominican Republic
Ecuador
Egypt
El Salvador
Ethiopia
Fiji
Finland
France metropolitan
FYR Macedonia
Gabon
Georgia
Germany
Ghana
Greece
Guatemala
Guinea
Guyana
Honduras
Hungary
Iceland
India
Indonesia
Iran, Islamic Republic of
Iraq
Ireland
Israel
Italy
Jamaica
Japan
Jordan
Kazakhstan
Kenya
Korea Republic of
Kyrgyzstan
Lebanon
Malawi
Malaysia
Mali
Malta
Mauritius
Mexico
Mongolia
Morocco
Mozambique
Namibia
New Zealand
Nicaragua
Niger
Nigeria
Norway
Pakistan
Panama
Papua New Guinea
Paraguay
Peru
Philippines
Poland
Portugal
Republic of Moldova
Romania
Russian Federation
Samoa
South African Republic
Spain
Sri Lanka
Sudan
Suriname
Swaziland
Sweden
Switzerland
Tajikistan
Tanzania
Thailand
The Gambia
The Netherlands
Togo
Trinidad and Tobago
Tunisia
Turkey
Turkmenistan
Uganda
Ukraine
United Kingdom
United States of America
Uruguay
Uzbekistan
Vanuatu
Venezuela
Vietnam
Yemen
Zambia