



# Cardiovascular and cancer risk factors analysis for 2001–2020 from the global research output and European newspapers

Elena Pallari<sup>1</sup> · Grant Lewison<sup>2</sup> 

Received: 10 November 2021 / Accepted: 30 June 2022 / Published online: 5 August 2022  
© The Author(s) 2022

## Abstract

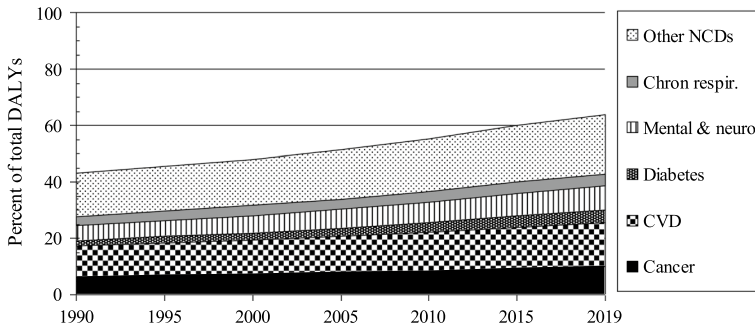
Cancer and cardiovascular disease (CVD) are now two of the leading components of the global burden of disease, especially in high- and upper-middle-income countries. Causes of the diseases that are amenable to intervention are multiple: tobacco control closely followed by obesity treatment, including promotion of a healthy diet and physical exercise, remain the global priorities. We interrogated the Web of Science (WoS) from 2001 to 2020 to determine the numbers of papers describing research into 14 different possible risk factors causing the two diseases. These ranged in relative importance from tobacco and being overweight to the consumption of excessively hot drinks (linked to oesophageal cancer), pollution (linked to lung cancer particularly) and also non-interventional genetic risks. The risks varied between different continental regions, and obesity has increased as a risk factor for CVD in some of these regions. Because many of these factors are subject to human behavioural choices, we also investigated how such research was being presented to the European public through newspaper reportage. About 40% of the factors that influence the cancer burden can be attributed to particular causes, and more than 85% of those factors influencing CVD can also be so attributed. They are led by tobacco use as a risk factor for cancer, but this is slowly declining in most high-income settings. For CVD, the major risks are metabolic, such as high systolic blood pressure and high body-mass index, but also from tobacco use. Research outputs on some of these different factors in the continental regions correlated positively with their influence on the disease burdens. The selection of European newspaper stories was biased towards those risk factors that could be considered as being under the control of their readers. Reports of research in the mass media have an important role in the control of both cancer and CVD, and should be regarded by public health authorities as a useful means to promulgate health education. This paper is based on one presented at the ISSI conference in Leuven in July 2021 (Pallari and Lewison, in: Glänzel et al (eds) Proceedings of the 18th international conference on scientometrics and informetrics, 2021), but has been extended to cover CVD as well as cancer. The geographical analysis of risk factors and research publications has also been modified.

**Keywords** Cancer · Cardiovascular disease · Risk factors · Research publications · Newspaper stories

---

✉ Grant Lewison  
grantlewis@aol.co.uk

Extended author information available on the last page of the article



**Fig. 1** Contributors of different non-communicable diseases (NCDs; World Health Organization group II) to the total global disease burden, measured in Disability-Adjusted Life Years (DALYs), 1990–2019. CVD cardiovascular diseases. Data from the Institute for Health Metrics and Evaluation (IHME), University of Washington, Seattle

## Introduction

As public health improves worldwide, life expectancy increases, and there is demographic ageing. This means that cancer, and other non-communicable diseases (NCDs) and cardiovascular disease (CVD) in particular, have become ever larger contributors to the global disease burden, as shown in Fig. 1. World-wide, their share of the burden has gone up from 45% in 1990 to 65% in 2019 because of the decline in communicable, maternal, neonatal, and nutritional diseases (group I).

As with other NCDs, prevention is cheaper and more effective than cure, so there is an increasing effort world-wide to understand context-specific causes of cancer and CVD and to intervene (Ginsberg et al., 2009; Hurry et al., 2020; Zimmermann et al., 2017). Many of these cancer and CVD risk factors are in theory under our control, such as our behaviour, our diet, and some aspects of our environment, but others are not, such as our genetics. In this study, we sought to determine which of these causative factors were being researched, and in which continental regions. We also sought to compare the global research agenda with the distribution of the different factors, see Table 1, and with how such research is being presented to the European public through the mass media. This is because a large part of the causal risks of these diseases is contributed by activities that we choose to undertake such as whether we smoke, what and how much we eat, and where we live and work. This applies much more to CVDs than to cancer, where more than half the total risk is from our genome.

The table shows that although controllable risks have slightly decreased, and those from tobacco have also reduced slightly, metabolic risks, especially those caused by high body-mass index (BMI), have increased for both diseases. They are the main source of risk for CVD, while smoking remains the principal risk for cancer.

Some of these causative factors are also constrained by government regulation. This is seen most obviously in the way that the tobacco industry has been circumscribed over recent decades, with restrictions on where people may smoke, the price of cigarettes, and controls on packaging (Flor et al., 2021). One consequence is that smoking has, almost uniquely, decreased as a cause of cancer over the last three decades in many (but not all) settings, when others have increased, notably high BMI, that is obesity, and the use of alcohol.

**Table 1** Main causes of CVD and cancer risk, 1990–2020

	CVD			Cancer		
	1990	2005	2020	1990	2005	2020
Total (all ones listed by the IHME)	83.3	85.0	85.8	39.8	41.3	41.8
Genetic (by subtraction from 100%)	16.7	15.0	14.2	60.2	58.7	58.2
Environmental	31.6	31.1	29.6	3.6	3.7	3.7
Occupational	0.0	0.0	0.0	2.7	2.7	2.8
Behavioural	54.5	54.4	52.4	35.2	35.7	34.9
Tobacco	24.1	23.3	21.0	24.5	24.7	23.5
Alcohol	2.6	3.5	3.0	4.8	5.0	5.2
Dietary	40.8	39.7	38.9	5.7	5.8	5.6
Metabolic	68.9	71.4	74.1	4.7	6.1	7.8
High systolic BP	51.7	52.9	54.4	0.0	0.0	0.0
High BMI	15.5	18.9	22.3	2.7	3.6	4.5

Data from the Institute for Health Metrics and Evaluation (IHME), University of Washington, Seattle [Data for 2020 extrapolated from 2019. *BP* blood pressure, *BMI* body-mass index]. Because of interactions, the various behavioural risk factors sum to more than the total

There is an abundance of research on the epidemiology of both CVD and cancer, and many papers have looked at the risk factors associated with particular cancers, such as breast (Zhang et al., 2011), cervical (Zhang et al., 2020), colorectal (Ma et al., 2014), gastric (Klingelhofer et al., 2021), lung (Yu et al., 2015) and melanoma (Mirza et al., 2021). However, there is a paucity of bibliometric studies on the epidemiology of cancer. The only one that we could find that related to all cancers was published 15 years ago (Ugolini et al., 2007), and was reviewed later that year (Boffetta, 2007). There the matter rested until November of last year (Lin et al, 2020) when a team from Xiamen University examined Chinese publications, but only ones in the Surveillance, Epidemiology, and End Results (SEER) database, so their coverage was inevitably very incomplete. In CVD, there has been a series of studies on the genetics of cardiovascular disease, beginning in 2009 (Irvin et al., 2016; Merino et al., 2019; Morrison et al., 2014; Psaty et al., 2009; Smith et al., 2010). However, we could find only the paper by Levois and Layard (1995), which discussed possible publication bias in the environmental tobacco-smoke and coronary heart disease literature. Therefore it appears that a comprehensive and up-to-date bibliometric examination of risk factors in cancer and cardiovascular research has not appeared in the literature. Our study attempts to fill this gap.

In this study, we examined the world output of papers on CVD and cancer research that involved epidemiology as applied to cancer risk factors, and classified them by 14 subject areas, some of them corresponding to the modifiable cancer risk factors listed in Table 1. They are listed in Table 2. Most of the papers concerned an increase in the risks, but some described means of lowering it, such as taking more exercise or eating a good diet (e.g., one rich in fruit and vegetables). These papers were classified by the continents of their authors, see Table 3. They are based on the continental and regional divisions used by the IHME.

We also analysed the research papers behind the stories in 31 European newspapers reporting the epidemiology of CVD and cancer risk factors. These were all classified by inspection of the reported story. We classified the stories by the countries of the citing

**Table 2** List of the 14 causes used for classification of the cancer epidemiology papers

Code	Subject area	Code	Subject area
ALCO	Alcohol use	INFE	Infections
DENT	Dental health and procedures	LIFE	Lifestyle choices, including exercise
DIET	Food and drink consumed	OBES	Obesity and its <i>sequelae</i>
DRUG	Medication and substance use	POLL	Pollution, and occupation
ECON	Socio-economic factors	RADI	Radiation, incl. medical use
GENE	Genetics	TOBA	Tobacco, including passive smoking
HEAT	Hot drinks and environment	UVRA	Ultraviolet radiation, tanning

**Table 3** Nine continental regions used for partition of the world for disease risk factors and for outputs of risk research papers in the Web of Science

Code	Continental region	Code	Continental region
ANZ	Australia and New Zealand	NAM	N. America (Canada & USA)
CEE	E. Europe, incl. Russia and Ukraine	SAS	South Asia, not Sri Lanka
EAS	East Asia, incl. China and Oceania	SSA	Sub-Saharan Africa
LAC	Latin America and the Caribbean	WEU	W. Europe, incl. Cyprus, Israel
MEM	Middle East and North Africa		

newspapers rather than those of their researchers, and by the risk(s) identified. The mass media are important as a way to provide public health information, and may draw attention to lesser-known causative factors for CVD and cancer. Several of these differ from the usual suspects, such as age (Macdonald et al., 2018), alcohol (Christensen et al., 2019; Pratt et al., 2014), and meat (Kowall & Stand, 2019). Thus the media have been used on several occasions to warn the public about the risks of cancer (de Vito et al., 2014; Donell et al., 2004; Lee et al., 2001; Stryker et al., 2008, 2009). However, there do not seem to be any equivalent studies for heart disease or stroke.

## Methodology

The papers about the risks of CVD and cancer were taken from the Web of Science (WoS, © Clarivate Analytics) from the 20 years, 2001–2020. They were limited to articles and reviews, but with no language restriction. First, we selected all papers on CVD and cancer research by means of two complex filters, which had been developed for two earlier projects. The one for CVD included 68 specialist cardiology journals, and 137 title words or phrases, but with an exclusion of the following title words: *Alzheimer\** or *antimalaria\** or *dementia* or *Falciparum* or *malaria\** or *optic\** or *retina\**. This filter had a precision, *p*, of 0.95 and a recall, *r*, of 0.90 (Huffman et al., 2013). The filter for cancer was based on 323 title words and 185 specialist cancer journals, and had a precision, *p*, of 0.95 and a recall, *r*, of 0.98 (Begum et al., 2018).

We then applied an epidemiology filter, consisting of 25 title words, such as:

associat\*, cohort-study, determinant, exposure, incidence, predict\*, risk.

and a set of 35 epidemiology journals. They were selected from the list of all the journals used for CVD and cancer research where the papers had the selected title words, and whose names contained the strings *EPIDEM* or *PREVENTION* or *RISK*. This third filter was calibrated with reference to the outputs of researchers in epidemiology departments (Lewison, 1999), and had a precision of 0.78 and a recall of 0.89, so it over-estimated the numbers of epidemiology papers by 12%. The application of this third filter identified a true total of 150,307 papers, or 13.2% of the total in CVD research. The actual number of CVD epidemiology papers identified in the WoS was 168,344, and these formed the database and were the ones whose characteristics were determined. For cancer, the true total was 171,366 epidemiology papers, or 10.5% of all cancer research papers in 2001–2020.

In order to select the main causes, or risk factors, we used the “compare” tab tool on the IHME website (IHME, 2020) to identify the ones that were responsible for many CVD or cancer DALYs (Disability-Adjusted Life Years) or could be clearly defined by means of title words in their papers. [The settings were as follows: Display=Risk; Location=Global; Year=2010; Age=All ages; Sex=Both; Level=2]. We then inspected a large number of the two sets of epidemiology papers and made a provisional list of the distinctive title words of the papers that could be used to identify those on each selected cause. This process was iterative, as the lists of title words were inevitably initially incomplete, and so it was repeated until we had created a comprehensive list of the title words to be used, some of which are reported in Table 4, below. We proceeded to download the long list of journals that had been used to publish the epidemiology papers. Journals whose names indicated that they were in one of the 14 selected subject areas (see Table 2)

**Table 4** List of five of the 14 risk factors chosen for analysis of the cancer epidemiology papers, with the title words used to identify them

Diet	Genetics	Infection	Lifestyle	Obesity
Carbohydrate	Allele	AIDS	Exercise	Bariatric surgery
Consume	BRCA	E-coli	Fitness	BMI
Consumption	Chromosom*	Escherichia	Interval-training	Body-fat*
Dairy	Deletion	Helicobacter	Lifestyle	Body-mass-index
Deficiency	Familial	HIV	Phone	Body-size
Diet*	Gene	Infect*	Physical activity	Body-weight
Food*	Genetic	Papillomavirus	Running	Lorcaserin
Fruit	Hereditary	Papilloma-virus	Screen-time	Obes*
Grain	Inherited	Vaccin*	Sedentary	Overweight
Meat	Loci	Virus	Sexual-activ*	Silhouette
Nutrient	Locus		Sexual-behav*	Waist
Nutrition	Mutation		Sex-with-men	Weight gain
Vegetable	Nucleotide		Shift-work	
Vitamin	Polygenic		Sitting	
	Polymorphism		Sports-activity	
	Repeat		Telephone	
	Triple-negative		Walking	
	Variant		Weight-training	

were then listed and used to select a second set of papers. We examined those papers that were so identified, but which were *not* identified by the provisional list of title words. This helped us in two ways. First, it suggested title words that could be added to the provisional list. And second, it showed which of these specialist journals were not appropriate because many of the papers that they generated were not relevant. [They were often about the increased risks of *other* diseases because the patients already had CVD or cancer. However, we retained papers where the risk of disease was increased because the patients were suffering from other diseases]. The list of title words was then finalised for each subject area, complemented by the specialist journals, and applied to the two epidemiology files so as to generate sets of papers, both in the world and in nine continental regions, see Table 3.

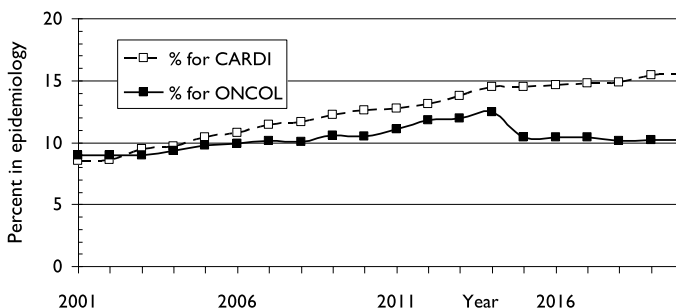
Table 4 lists the title words for five of the selected subject areas. They were designed to be used with the WoS software, and the hyphen indicates that the words must be next to each other in the titles. The WoS automatically includes plurals as well as singular forms, e.g., *vitamins* as well as *vitamin* but wild cards (\*) are needed to cover different forms such as *infected*, *infection*, *infects*. Some papers would have been classed in more than one of the 14 causes, and others in none of them. The \* denotes any character.

## Results

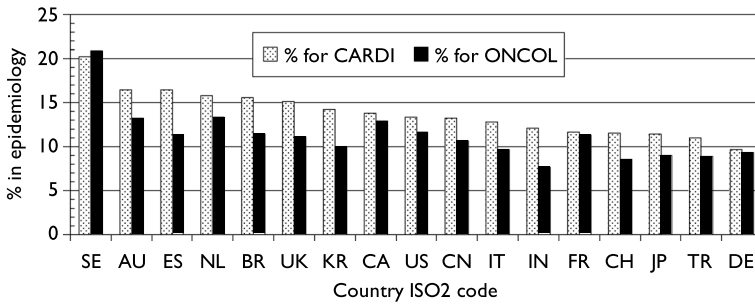
### Research papers on epidemiology

Epidemiology has increased markedly with time as a percentage of CVD research papers, see Fig. 2. It has almost doubled, from 8.5% in 2001 to 15.5% in 2020. However, the corresponding percentage for cancer research has hardly changed. It increased only from 8.9% in 2001 to 10.2% in 2020, although it reached 12.4% in 2014. [The subsequent drop was probably an artefact, because the WoS increased its journal coverage in 2015 to include more from countries outwith Western Europe and North America].

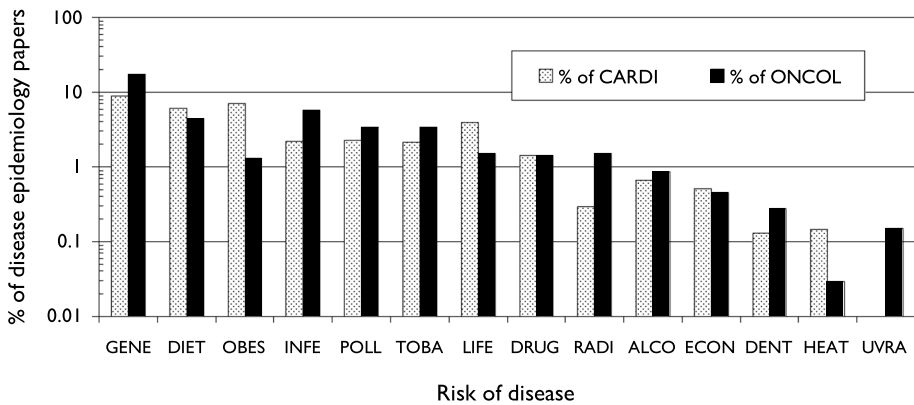
There was a big variation in the relative commitment to epidemiology as a subject for research within CVD and cancer. Figure 3 shows that, among the 17 countries with at least 20,000 CVD research papers, Sweden ranked highest and Germany the lowest. However, if we extend the coverage to countries with > 300 CVD research papers, then Iceland becomes the highest-ranked, with 38% in both disease areas, followed by some African countries—Uganda with 33% and 25% in CVD and cancer, Ethiopia with 32% and 23%,



**Fig. 2** Percentages of CVD and cancer research papers that were in the epidemiology sub-field, 2001–20, variation by year



**Fig. 3** Percentages of CVD and cancer research papers that were in the epidemiology sub-field, 2001–20, for leading countries (with > 20,000 CVD research papers). *SE* Sweden, *AU* Australia, *ES* Spain, *NL* Netherlands, *BR* Brazil, *KR* South Korea, *CA* Canada, *CN* China, *IT* Italy, *IN* India, *FR* France, *CH* Switzerland, *JP* Japan, *TR* Turkey, *DE* Germany



**Fig. 4** Percentages of epidemiology of CVD and cancer research papers in each of 14 sub-fields (for codes, see Table 2), 2001–2020

and Cameroon with 32% and 12%. Bangladesh is also prominent in the epidemiology of CVD with 32% of its total research output. At the lower end of the scale is Russia, with percentages just below those of Germany.

Figure 4 shows the numbers of papers in each of the 14 subject areas, plotted on a logarithmic scale. There is a difference of almost three orders of magnitude between the largest (genetics) and the smallest (hot drinks and heat, and Ultra-Violet radiation). Hot drinks seem to be locally or regionally important (mainly for oesophageal cancer) but do not feature among the main causes in Table 1. UV radiation (UVRA) is primarily a problem in Australia and New Zealand, see Tables 5 and 6, and attracts research on its effects on cancer, but not on CVD. On the other hand, the effects of obesity (OBES) on CVD are more frequently researched than their effects on cancer, despite claims by Cancer Research UK, a leading funder of cancer research, that obesity is the second most important risk factor (after smoking) (Cancer Research UK, 2021).

The next two tables show the relative commitment of the nine continental regions to research on the different risk factors for the two diseases. Australia and New Zealand (ANZ) are primarily concerned with the effects of UV radiation on both diseases, of heat

**Table 5** Relative commitment of different continental regions to different types of epidemiological research on CVD, 2001–20, relative to the world as a whole

	ANZ	CEE	EAS	LAC	MEM	NAM	SAS	SSA	WEU
ALCO	0.80	1.14	1.21	0.90	0.80	0.86	<b>1.65</b>	<i>0.44</i>	1.07
DENT	<i>0.61</i>	0.76	1.13	1.17	1.28	0.94	1.28	<b>1.94</b>	0.95
DIET	1.16	0.82	0.99	1.38	1.28	1.03	1.06	0.89	0.94
DRUG	1.01	<i>0.69</i>	0.90	<i>0.65</i>	<i>0.70</i>	1.31	0.84	<i>0.30</i>	0.92
ECON	<b>1.61</b>	<i>0.64</i>	<i>0.45</i>	0.87	<i>0.64</i>	1.22	<b>1.54</b>	<b>2.17</b>	1.31
GENE	<i>0.69</i>	1.35	<b>1.48</b>	0.93	1.30	0.79	<b>1.42</b>	<i>0.62</i>	0.93
HEAT	<b>2.15</b>	<i>0.67</i>	<b>1.68</b>	<i>0.60</i>	<i>0.42</i>	0.90	<i>0.32</i>	<i>0.28</i>	0.98
INFE	0.81	<i>0.62</i>	0.78	1.17	0.99	1.14	0.90	<b>2.84</b>	0.97
LIFE	<b>1.78</b>	0.71	<i>0.61</i>	<b>1.87</b>	<i>0.70</i>	1.11	<i>0.56</i>	1.09	1.15
OBES	1.14	1.18	0.79	<b>1.79</b>	1.29	1.01	0.79	1.12	0.94
POLL	0.98	0.81	1.06	1.06	0.75	1.26	0.93	<i>0.60</i>	0.91
RADI	0.90	1.01	0.81	<i>0.30</i>	<i>0.42</i>	1.20	1.13	<i>0.14</i>	1.13
TOBA	0.88	1.01	0.97	0.80	1.28	1.05	1.31	<i>0.49</i>	0.97
UVRA	<b>6.78</b>	<b>1.59</b>	<i>0.32</i>	<b>1.62</b>	<i>0.00</i>	0.87	<i>0.00</i>	<i>0.00</i>	0.82

Ratios > 1.414 shown in **bold** type, ratios < 0.707 shown in *italic* type

**Table 6** Relative commitment of different continental regions to different types of epidemiological research on cancer 2001–20, relative to the world as a whole

	ANZ	CEE	EAS	LAC	MEM	NAM	SAS	SSA	WEU
ALCO	1.28	1.12	1.02	<b>1.44</b>	<i>0.69</i>	1.09	1.19	0.78	1.21
DENT	<i>0.64</i>	1.41	0.97	<b>2.10</b>	<b>1.44</b>	0.81	<b>3.29</b>	1.11	0.98
DIET	1.21	0.72	0.80	1.20	0.94	1.27	<i>0.66</i>	<i>0.77</i>	1.13
DRUG	1.04	<i>0.70</i>	0.81	<i>0.65</i>	<i>0.61</i>	<b>1.42</b>	0.79	<i>0.68</i>	1.01
ECON	<b>1.77</b>	<i>0.60</i>	<i>0.35</i>	1.04	<i>0.57</i>	1.32	1.21	0.94	1.19
GENE	1.05	<b>1.54</b>	1.29	1.01	1.13	0.88	1.36	<i>0.60</i>	0.89
HEAT	1.38	<b>1.94</b>	1.14	<i>0.53</i>	<b>1.45</b>	0.90	<i>0.66</i>	<b>10.3</b>	0.93
INFE	0.77	0.84	1.07	<b>1.96</b>	1.16	0.93	1.06	<b>4.14</b>	0.92
LIFE	<b>2.22</b>	<i>0.71</i>	<i>0.42</i>	0.88	0.73	<b>1.46</b>	<i>0.38</i>	<i>0.77</i>	1.19
OBES	<b>1.56</b>	<i>0.70</i>	0.80	0.96	<i>0.68</i>	<b>1.42</b>	<i>0.46</i>	<i>0.57</i>	1.00
POLL	1.24	<b>1.47</b>	<i>0.67</i>	1.00	0.83	1.22	1.11	1.05	1.22
RADI	0.78	<b>1.94</b>	0.81	<i>0.49</i>	0.72	1.21	0.73	0.73	1.05
TOBA	0.89	1.08	0.83	0.86	0.76	1.32	<b>2.64</b>	<i>0.60</i>	0.91
UVRA	<b>5.54</b>	<i>0.48</i>	<i>0.17</i>	<i>0.52</i>	0.86	1.38	<i>0.26</i>	<i>0.41</i>	1.29

Ratios > 1.414 shown in **bold** type, ratios < 0.707 shown in *italic* type

on CVD, and of lifestyle choices on cancer. Radiation (RADI) is of major interest in Eastern Europe for its effects on cancer. There is relatively little research in East Asia (EAS) on lifestyle choices (mainly the effects of exercise, or lack of it). Their effects on CVD, but not on cancer, are however a focus for research in Latin America (LAC). Sub-Saharan

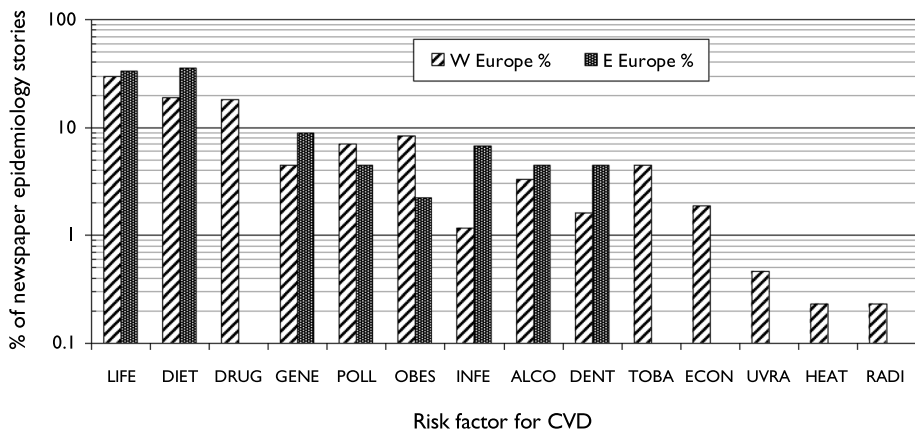


Africa’s research output is quite low, but is focussed on infection (INFE), especially the link between AIDS and cancer, and the effects of heat on cancer. The effects of smoking (TOBA) on both CVD and cancer are well-established, and the heaviest relative concentration of research is in South Asia (SAS), principally India. This region is also researching the effects of dental problems (DENT), and their treatment, on cancer. Although this is a relatively small subject area, it is growing rapidly (at over 10% per year over the study period), and is also of research interest in Latin America, especially in Brazil. The effects of poverty and economic inequality (ECON) are primarily researched in Australia and New Zealand, and in North America (NAM), both on CVD and on cancer.

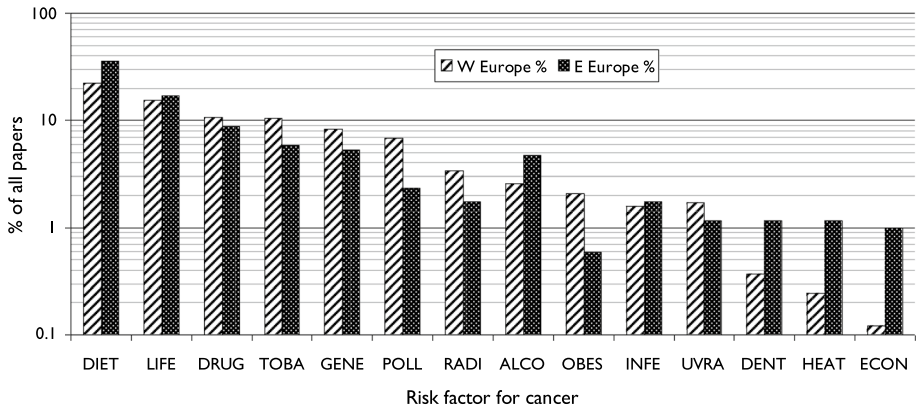
**Newspaper reportage on epidemiology of CVD and cancer**

The database of these newspaper stories was compiled for an earlier study (Pallari & Lewison, 2017) as part of the evaluation of European research in these disease areas. The database was a product of a consistent and systematic search enquiry that was set up in English and translated into 19 languages and applied to 32 newspapers from 21 European countries. There were no language restrictions on the news stories, although most of the identified research studies were in English. For more information on the methodology, please see Pallari et al. (2019). Using the aforementioned database, we examined the subset of those news stories that described epidemiological studies for these two diseases, and classified them by inspection of their titles and synopses into one (or occasionally more) of the 14 risk factors listed in Table 2. [A few stories were about other ones]. Figures 5 and 6 show the distributions for CVD and cancer, respectively, and for the two geographical groups of newspapers.

Lifestyle changes, particularly exercise and being sedentary, together with dietary choices, are the main subjects. They are comparable in prominence for CVD, but diet has a bigger role in stories about cancer. The effects of drugs, both positive and negative, are important for CVD in Western Europe, but not in Eastern Europe, where the influence of infections occupy more space. Perhaps surprisingly, pollution gets more attention in Western than in Eastern Europe; for CVD it is noise from aircraft and motor traffic that dominates. Many of the stories give contradictory messages, for example on the effects



**Fig. 5** Newspaper stories in 2001–2013 on the risk factors in the epidemiology of CVD



**Fig. 6** Newspaper stories in 2001–2013 on the risk factors in the epidemiology of cancer

of alcohol and of hormone replacement therapy (HRT), which may be either positive or negative. Under the heading of genetics, comparisons were sometimes made between the relative risks suffered by men and women. Obesity as a risk for both diseases was the subject of relatively more stories in Western Europe than in Eastern Europe. The risks from socioeconomic deprivation were featured for CVD in Western Europe, but for cancer in Eastern Europe.

### IHME data on the risk factors in the nine world regions

The IHME data on the causes of cancer in different continents are available, and some results for 2010 are shown in Table 7. They are based on the selection of Advanced Settings > DALYs > 2010 > All ages > Both sexes > Level 2 (CVD and Neoplasms). For each selected geographical area (Location), and risk factor, the percentages of Risk factor attribution (RFA) were read from the diagram. For example, for Location = Eastern Europe, and Risk = Tobacco, RFA for CVD = 25.6% and for Neoplasms = 27.2%. These data show that there are some big differences between continents on which causes are the most important with regard to the development (or the reduction) of CVD and cancer.

The individual causes (from tobacco use to pollution and occupational causes) add to more than 100% because of double counting with some causes combining to increase the risk of the two diseases. Tobacco as a cause of cancer varies greatly, from 31% in North America to only 8% in sub-Saharan Africa. Alcohol has the biggest effects in Australia & New Zealand and Europe (CEE and WEU), but much less in the Middle East & North Africa (MEM), all of whose countries are Muslim-majority where alcohol is not freely available. Obesity has the largest effect in North America, and for CVDs in the Middle East & North Africa, and the least in Asia (EAS and SAS). Within EAS, China suffers mainly from its diet, and from pollution. So these data tend to confirm some subjective views of the world.

This table shows that genetics plays a much larger part in cancer than in CVD, where its role is quite small apart from in Sub-Saharan Africa. For CVD, the main causes are diet (especially in South Asia and Eastern Europe), followed by pollution (mainly in the same two regions). Tobacco is the main causative factor for cancer, greatest in North America,

**Table 7** Data from IHME on the percentage contributions to the risk for CVDs and cancer in the different continental regions in 2010, DALYs, all ages and both sexes

Region	Genetics GENE	Alcohol ALCO	Dietary risks DIET	High BMI OBES	Envir/occup POLL	Tobacco TOBA
Risk factors for CVDs						
ANZ	19.8	3.2	34.7	22.9	10.3	17.4
CEE	9.5	8.0	42.0	27.7	20.7	25.6
EAS	13.1	3.6	41.4	15.1	35.9	25.5
LAC	17.3	2.4	33.9	28.0	20.9	20.2
MEM	15.0	0.1	33.5	31.3	33.7	22.3
NAM	17.7	2.6	37.4	30.1	12.3	24.3
SAS	16.7	0.8	43.0	14.1	40.5	19.4
SSA	23.0	2.8	28.9	18.6	32.8	9.7
WEU	17.8	4.2	32.5	19.5	15.6	20.3
World	14.7	3.3	39.4	19.8	31.4	22.4
Risk factors for cancer						
ANZ	58.7	7.5	6.3	6.1	7.8	20.8
CEE	52.5	7.2	6.2	6.0	6.0	29.5
EAS	56.5	4.9	6.5	3.2	7.8	26.5
LAC	64.6	4.1	4.7	4.6	3.8	16.5
MEM	67.7	1.1	3.4	4.7	5.7	19.5
NAM	51.4	4.9	5.3	6.8	7.2	31.3
SAS	68.3	4.4	4.3	1.6	3.7	16.6
SSA	70.3	3.8	4.2	2.5	3.2	7.8
WEU	52.6	7.3	5.8	5.2	9.2	28.9
World	58.5	5.1	5.7	3.9	6.6	24.3

and least in Sub-Saharan Africa. It also has a noticeable effect on CVD, especially in Eastern Europe (notably in Montenegro, 34.6%) and East Asia.

It is possible to make a comparison of these risk factors in the nine continental regions with research data (taken from Table 4) for four of the subject areas listed in Table 1, tobacco, alcohol, diet and obesity. This can show whether the researchers are focusing their attention on the areas deemed to be of importance to the control of cancer in the eight continents. The correlation is positive for all four; it is highest for obesity and lowest for diet, probably because dietary data are not so readily available, or indeed so reliable, as data for the other three risk factors.

## Discussion

In this work, we report (what we think is for the first time) a comparison between the causes of CVD and cancer, the amount of research on each of the main risk factors, and the extent of their news coverage in Europe. News coverage is important because several of the causes of these diseases are to some extent under readers’ control, and therefore may be influenced by the stories that they read. It was clear from these data that the journalists had selected research stories, not as an unbiased sample but rather ones that they thought would

interest their readers. Thus, of the 74 stories about genetics and cancer, the largest number ( $n=27$ , or 36%) were about breast cancer, which is always a popular subject (Lewison et al., 2008).

The 14 risk factors are somewhat different from each other. Genetics (GENE) is the dominant subject area in most of them, but infection (INFE) is top in both Africa and Latin America because of the papilloma virus and its effect on cervical cancer, which is less well screened and treated there compared with Western Europe and North America (Vaccarella et al., 2017). Pollution (POLL) is evidently a subject for a lot of research in Eastern Europe, as is smoking (TOBA). There is relatively little research on this in China; this may reflect the policy of the Chinese government, which derives very substantial tax revenues from cigarettes (Li & Lewison, 2020). The interest of the countries of Eastern Europe in radiation (RADI) is probably attributable to the long-term health effects from the Chernobyl nuclear accident in April 1986 (Ivanov et al., 2012; Rivkind et al., 2020). It is also notable that obesity (OBES) is of particular interest in Australia & New Zealand and the rich countries of North America and Western Europe. Australia is also the most concerned with research on ultraviolet radiation (UVRA) because of its high incidence of solar-induced skin cancer (McLoone et al., 2014; McMeniman et al., 2020). We added the subject areas of dental issues and hot drinks because of the belief that dental health has wider effects on bodily health and the risk of various manifestations of cancer (Chen et al., 2020; Kawasaki et al., 2020; Wu et al., 2021), and the problems, particularly in China, with oesophageal cancer which can result from the drinking of very hot tea (Li & Lewison, 2020; Yu et al., 2018).

The impact of such modifiable risk factors and the disease burden of CVD and cancer, however, may not be entirely understood at the population level until there is a clear communication and awareness strategy from governments. Such a distinction between policy and action was made in a recent study that showed that government strategic plans focused on the treatment of diabetes rather than its prevention in Cyprus, Iceland, Luxembourg, Malta and Montenegro (Cuschieri et al., 2021). Furthermore, the importance of translating burden of disease studies for use in policy-making was recently raised with efforts concentrated in producing a knowledge translation framework (Pallari et al., 2020). To support these efforts the European COST Action burden of disease network was set up. It is intended to integrate technical initiatives in the development and use of these metrics into findings relevant to stakeholders. It will also strengthen the capacity for cross-country collaboration, as reported by Devleeschauwer (2020). To this end, and to effect such a change in health communication and policy-making, data visualisation tools may enhance the engagement of decision-makers and relevant stakeholders, as a disconnect between research and public health has been reported (Lundkvist et al., 2021).

## Strengths and limitations

The study involved the analysis of voluminous research outputs at the global level, the comparative assessment using burden of disease data from IHME and examination of news-media reports. We were fortunate to have developed such a database of newspaper stories from so many different countries, albeit confined to Europe. Additionally, we have drawn comparisons between the countries concerned from the research outputs and newspapers with data on the causes of CVD and cancers as part of an effort in linking these data and identifying gaps in public health communication and informing policy-makers on how to reduce the burden caused by these diseases in these countries.

The study has several limitations. Because of time, the two files of epidemiology papers were “virtual” ones, and the analysis was based only on the WoS software. [Our analysis was confined to this database as it provided a good coverage of research on the two diseases, with careful curation of journals, and our own analytical software has been developed for it over many decades]. The details of the papers were not actually downloaded to file, which would have allowed us to examine the fractional counts of countries, which provide a fairer assessment of their individual contributions to the research. It would also have allowed us to link the various causes, *e.g.*, diet, alcohol use, to the different manifestations of cancer that are particularly affected, which might help researchers to focus their work better. We have done this only for two examples, but there must be many others besides the link between smoking and lung cancer, which is one of the best-established. The allocation of papers to causes that we performed has similarities with the topic modelling process in which documents are classified in multiple subject areas based on text mining. However, instead of using natural language processing techniques for this paper, we visually inspected samples of the papers and used our judgment to classify them through a complex iterative process.

### Implications for further research

There is some evidence that CVD and cancer researchers in different continental regions are studying the subjects that are the most important for their influence on the risks. This is only an association, and it would be necessary to carry out further studies such as a survey of some of the leading researchers to establish whether they are aware of the relative importance of the different causes, and whether this knowledge had influenced their choice of research topics. It would also be useful to ask funders of such research whether they try to support the research that was most likely to reduce the disease burden in their countries, or whether their allocation of funds was based only on their perceptions of the potential quality of the work. Additionally, mapping out the factors based on a country analysis may provide useful results that are of value to local policy-makers.

### Conclusions

Epidemiological research into the causes and prevention of CVD and cancer only represent small percentages of the total, despite this being more cost-effective than treatment once disease is established in a patient. However, such research appears to be targeted at the most important causes in different continental regions. Although the prevention of these two diseases depends heavily on governmental public health campaigns (notably to control smoking), the lifestyle choices of members of the public also make a difference. These can be influenced by information about new research findings that are reported in the news-media. Within Europe, it seems that these are chosen to reflect the topics that are likely to have the most interest for their readers, which is helpful.

**Acknowledgements** The authors are grateful to the many research assistants and volunteers who worked to analyse the European newspaper stories at King’s College London. This project was led and managed by EP and was supported by the European Commission’s Seventh Framework Programme (Grant No. EC/FP7/602536). The work of the current paper was funded by the UK Research and Innovation GCRF ESRC Grant No ES/P010962/1. This paper is an extended version of the one by the same authors that was presented at the ISSI conference in Leuven in July 2021 (Pallari & Lewison, 2021).

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Begum, M., Lewison, G., Lawler, M., & Sullivan, R. (2018). Mapping the European cancer research landscape: An evidence base for national and Pan-European research and funding. *European Journal of Cancer*, *100*, 75–84.
- Boffetta, P. (2007). Molecular cancer epidemiology: A tale of >3842 publications. *Carcinogenesis*, *28*(8), 1621.
- Cancer Research UK. (2021). *Does obesity cause cancer?* Retrieved October 29, 2021, from <https://www.cancerresearchuk.org/about-cancer/causes-of-cancer/obesity-weight-and-cancer/does-obesity-cause-cancer>
- Chen, Y., Zhu, B. L., Wu, C. C., Lin, R. F., & Zhang, X. (2020). Periodontal disease and tooth loss are associated with lung cancer risk. *BioMed Research International*. <https://doi.org/10.1155/2020/5107696>
- Christensen, A. S. P., Meyer, M. K. H., Dalum, P., & Krarup, A. F. (2019). Can a mass media campaign raise awareness of alcohol as a risk factor for cancer and public support for alcohol-related policies? *Preventive Medicine*, *126*, 125722.
- Cuschieri, S., Pallari, E., Terzic, N., et al. (2021). Mapping the burden of diabetes in five small countries in Europe and setting the agenda for health policy and strategic action. *Health Research Policy and Systems*, *19*, 43. <https://doi.org/10.1186/s12961-020-00665-y>
- De Vito, C., Angeloni, C., De Feo, E., et al. (2014). A large cross-sectional survey investigating the knowledge of cervical cancer risk aetiology and the predictors of the adherence to cervical cancer screening related to mass media campaign. *Health Education Research*, *30*(2), 336–346.
- Devleeschauwer, B. (2020). European burden of disease network: Strengthening the collaboration. *European Journal of Public Health*, *30*(1), 2–3.
- Donell, L., Hoffman-Goetz, L., & Clarke, J. N. (2004). Portrayal of genetic risk for breast cancer in ethnic and non-ethnic newspapers. *Women & Health*, *40*(4), 93–111.
- Flor, L. S., Reitsma, M. B., Gupta, V., et al. (2021). The effects of tobacco control policies on global smoking prevalence. *Nature Medicine*. <https://doi.org/10.1038/s41591-020-01210-8>
- Ginsberg, G. M., Edejer, T. T. T., Lauer, J. A., & Sepulveda, C. (2009). Screening, prevention and treatment of cervical cancer—A global and regional generalized cost-effectiveness analysis. *Vaccine*, *27*(43), 6060–6079.
- Huffman, M. D., Baldrige, A., Bloomfield, G. S., et al. (2013). Global cardiovascular research output, citations, and collaborations: A bibliometric analysis. *PLoS ONE*, *8*(12), e83440.
- Hurry, M., Eccleston, A., Dyer, M., & Hoskins, P. (2020). Canadian cost-effectiveness model of BRCA-driven surgical prevention of breast/ovarian cancers compared to treatment if cancer develops. *International Journal of Technology Assessment in Health Care*, *36*(2), 104–112.
- IHME. (2020). *GBD compare* | Viz Hub. Retrieved October 29, 2021, from <https://vizhub.healthdata.org/gbd-compare/>
- Irvin, M. R., Zhi, D. G., Sitlani, C., et al. (2016). Genome-wide meta-analysis of SNP-by-antihypertensive treatment class effect on serum potassium in cohorts of European and African Ancestry from cohorts for heart and aging research in Genomic Epidemiology Consortium. *Circulation*, *133*, 80.
- Ivanov, V. K., Kashcheev, V. V., Chekin, S. Y., et al. (2012). Radiation-epidemiological studies of thyroid cancer incidence in Russia after the Chernobyl accident (estimation of radiation risks, 1991–2008 follow-up period). *Radiation Protection Dosimetry*, *151*(3), 489–499.
- Kawasaki, M., Ikeda, Y., Ikeda, E., et al. (2020). Oral infectious bacteria in dental plaque and saliva as risk factors in patients with esophageal cancer. *Cancer*. <https://doi.org/10.1002/cncr.33316>
- Klingelhofer, D., Braun, M., Schöffel, N., et al. (2021). Gastric cancer: Bibliometric analysis of epidemiological, geographical and socio-economic parameters of the global research landscape. *International Journal of Health Policy and Management*, *10*(3), 118–128.

- Kowall, B., & Stand, A. (2019). Mistakes and drawbacks in German newspaper articles on epidemiologic studies: “Meat consumption and colorectal cancer” as an example. *Gesundheitswesen*, *81*(5), 438–443.
- Lee, C. W., Weit, L., & Gelmon, K. A. (2001). Effect of news media coverage of cancer risk on clinical oncology practice. *Journal of the National Cancer Institute*, *93*(21), 1656–1657.
- Levois, M. E., Layard, M. W. (1995). Publication bias in the environmental tobacco-smoke coronary heart-disease epidemiologic literature. *Regulatory Toxicology and Pharmacology*, *21*(1), 184–191. <https://doi.org/10.1006/rtp.1995.1023>
- Lewison, G. (1999). The definition and calibration of biomedical subfields. *Scientometrics*, *46*(3), 529–537.
- Lewison, G., Tootell, S., Roe, P., & Sullivan, R. (2008). How do the media report cancer research? A study of the UK’s BBC website. *British Journal of Cancer*, *99*, 569–576.
- Li, A., & Lewison, G. (2020). Chinese cancer research in 2009–18 and the disease burden. *Cancer Management and Research*, *12*, 5031–5040.
- Lin, M. Q., Lian, C. L., Zhou, P., et al. (2020). Analysis of the trends in publications on clinical cancer research in mainland China from the surveillance, epidemiology, and end results (SEER) database: Bibliometric study. *JMIR Medical Informatics*, *8*(11), e21931.
- Lundkvist, A., El-Khatib, Z., Kalra, N., et al. (2021). Policy-makers’ views on translating burden of disease estimates in health policies: Bridging the gap through data visualization. *Archives of Public Health*, *79*(1), 1–11.
- Ma, X. Y., Zhang, B., & Zheng, W. (2014). Genetic variants associated with colorectal cancer risk: Comprehensive research synopsis, meta-analysis, and epidemiological evidence. *Gut*, *63*(2), 326–336.
- Macdonald, S., Cunningham, Y., Patterson, C., et al. (2018). Mass media and risk factors for cancer: The under-representation of age. *BMC Public Health*, *18*, 490. <https://doi.org/10.1186/s12889-018-5341-9>
- McLoone, J. K., Meiser, B., Karatas, J., et al. (2014). Perceptions of melanoma risk among Australian adolescents: Barriers to sun protection and recommendations for improvement. *Australian and New Zealand Journal of Public Health*, *38*(4), 321–325.
- McMeniman, K. D. D. L., Jagirdar, K., et al. (2020). The interplay of sun damage and genetic risk in Australian multiple and single primary melanoma cases and controls. *British Journal of Dermatology*, *183*(2), 357–366.
- Merino, J., Dashti, H. S., Li, S. X., et al. (2019). Genome-wide meta-analysis of macronutrient intake of 91,114 European ancestry participants from the cohorts for heart and aging research in genomic epidemiology consortium. *Molecular Psychiatry*, *24*(12), 1920–1932.
- Mirza, F. N., Yumeen, S., & Walter, F. M. (2021). The epidemiology of malignant melanoma, squamous cell carcinoma and basal cell carcinoma in the UK from 2004 to 2014: A population-based cohort analysis using the Clinical Practice Research Datalink. *British Journal of Dermatology*, *184*(2), 365–367.
- Morrison, A. C., Bis, J. C., Hwang, S. J., et al. (2014). Sequence analysis of six blood pressure candidate regions in 4,178 individuals: The cohorts for heart and aging research in genomic epidemiology (CHARGE) targeted sequencing study. *PLoS ONE*, *9*(10), e109155.
- Pallari, E., & Lewison, G. (2017). European newspaper reports of non-communicable disease research, 2002–2013. In *ISSI-International conference on scientometrics & informetrics* (pp. 587–598). University of Wuhan.
- Pallari, E., & Lewison, G. (2019). How biomedical research can inform both clinicians and the general public. In W. Glänzel (Ed.), *Springer handbook of science and technology indicators* (pp. 581–607). Springer.
- Pallari E & Lewison G. (2021) World research and European news-media reports of cancer risk factors: Implications for public health communication. In W. Glänzel, S. Heeffer, P.-S. Chi, & R. Rousseau (Eds.) *Proceedings of the 18th International Conference on Scientometrics and Informetrics* (pp. 871–882).
- Pallari, E., Thomsen, S. T., & Hilderink, H. B. M. (2020). Knowledge translation and health policy for burden of disease. *European Journal of Public Health*, *30*, 165–1380.
- Pratt, S., Dixon, H., Scully, M., et al. (2014). “Alcohol causes cancer in more places than you think”: Using a mass media campaign to raise awareness of the link between alcohol and cancer. *Asia-Pacific Journal of Clinical Oncology*, *10*(S1), 147.
- Psaty, B. M., O’Donnell, C. J., Gudnason, V., et al. (2009). Cohorts for heart and aging research in genomic epidemiology (CHARGE) consortium design of prospective meta-analyses of genome-wide association studies from 5 cohorts. *Circulation-Cardiovascular Genetics*, *2*(1), 73-U128.
- Rivkind, N., Stepanenko, V., Belukha, I., et al. (2020). Female breast cancer risk in Bryansk Oblast, Russia, following prolonged low dose rate exposure to radiation from the Chernobyl power station accident. *International Journal of Epidemiology*, *49*(2), 448–456.

- Smith, N. L., Felix, J. F., Morrison, A. C., et al. (2010). Association of genome-wide variation with the risk of incident heart failure in adults of European and African Ancestry a prospective meta-analysis from the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium. *Circulation-Cardiovascular Genetics*, 3(3), 256–U79.
- Stryker, J. E., Fishman, J., Emmons, K. M., & Viswanath, K. (2009). Cancer risk communication in mainstream and ethnic newspapers. *Preventing Chronic Disease*, 6(1), A23.
- Stryker, J. E., Moriarty, C. M., & Jensen, J. D. (2008). Effects of newspaper coverage on public knowledge about modifiable cancer risks. *Health Communication*, 23(4), 380–390.
- Ugolini, D., Puntoni, R., et al. (2007). A bibliometric analysis of scientific production in cancer molecular epidemiology. *Carcinogenesis*, 28(8), 1774–1779.
- Vaccarella, S., Laversanne, M., Ferlay, J., & Bray, F. (2017). Cervical cancer in Africa, Latin America and the Caribbean and Asia: Regional inequalities and changing trends. *International Journal of Cancer*, 141(10), 1997–2001.
- Wu, H. D., Zhang, J. J., & Zhou, B. J. (2021). Toothbrushing frequency and gastric and upper aerodigestive tract cancers risk: A meta-analysis. *European Journal of Clinical Investigation*, 51, e13478. <https://doi.org/10.1111/eci.13478>
- Yu, C. Q., Tang, H. J., Guo, Y., et al. (2018). Hot tea consumption and its interactions with alcohol and tobacco use on the risk for esophageal cancer: A population-based cohort study. *Annals of Internal Medicine*, 168(7), 489.
- Yu, N., Su, X. M., Wang, Z. F., et al. (2015). Association of dietary vitamin A and beta-carotene intake with the risk of lung cancer: A meta-analysis of 19 publications. *Nutrients*, 7(11), 9309–9324.
- Zhang, B., Beeghly-Fadiel, A., Long, J. R., & Zheng, W. (2011). Genetic variants associated with breast-cancer risk: Comprehensive research synopsis, meta-analysis, and epidemiological evidence. *Lancet Oncology*, 12(5), 477–488.
- Zhang, T. C., Yin, X. L., Yang, X. R., et al. (2020). Research trends on the relationship between microbiota and gastric cancer: A bibliometric analysis from 2000 to 2019. *Journal of Cancer*, 11(16), 4823–4831.
- Zimmermann, M. R., Vodicka, E., Babigumira, J. B., et al. (2017). Cost-effectiveness of cervical cancer screening and preventative cryotherapy at an HIV treatment clinic in Kenya. *Cost-Effectiveness and Resource Allocation*, 15, 13. <https://doi.org/10.1186/s12962-017-0075-6>

## Authors and Affiliations

Elena Pallari<sup>1</sup> · Grant Lewison<sup>2</sup> 

Elena Pallari  
elena.pallari@outlook.com

<sup>1</sup> Health Innovation Network, Minerva House, 5 Montague Close, London SE1 9BB, UK

<sup>2</sup> Department of Cancer Policy, King's College London, Guy's Hospital, Great Maze Pond, London SE1 9RT, UK