ORIGINAL ARTICLE

Environmental risk factors for cancers of the brain and nervous system: the use of ecological data to generate hypotheses

Frank de Vocht,[1] Kimberly Hannam,[2] Iain Buchan[3]

▶ Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/oemed-2012-100954).

[1]Centre for Occupational and Environmental Health, Institute for Population Health, Manchester Academic Health Sciences Centre, The University of Manchester, Manchester, UK
[2]Maternal and Fetal Health Research, Medical and Human Sciences, Manchester Academic Health Sciences Centre, University of Manchester, Manchester, UK
[3]NIBHI & NWeH Science, Manchester Academic Health Sciences Centre, The University of Manchester, Manchester, UK

Correspondence to
Dr Frank de Vocht, COEH, Institute for Population Health, Manchester Academic Health Sciences Centre, The University of Manchester, Oxford Road, Manchester M13 9PL, UK; frank.devocht@manchester.ac.uk

Received 29 May 2012
Revised 27 November 2012
Accepted 14 December 2012
Published Online First 23 January 2013

To cite: de Vocht F, Hannam K, Buchan I. *Occup Environ Med* 2013;**70**: 349–356.

ABSTRACT

Background There is a public health need to balance timely generation of hypotheses with cautious causal inference. For rare cancers this is particularly challenging because standard epidemiological study designs may not be able to elucidate causal factors in an early period of newly emerging risks. Alternative methodologies need to be considered for generating and shaping hypotheses prior to definitive investigation.

Objectives To evaluate whether open-access databases can be used to explore links between potential risk factors and cancers at an ecological level, using the case study of brain and nervous system cancers as an example.

Methods National age-adjusted cancer incidence rates were obtained from the GLOBOCAN 2008 resource and combined with data from the United Nations Development Report and the World Bank list of development indicators. Data were analysed using multivariate regression models.

Results Cancer rates, potential confounders and environmental risk factors were available for 165 of 208 countries. 2008 national incidences of brain and nervous system cancers were associated with continent, gross national income in 2008 and Human Development Index Score. The only exogenous risk factor consistently associated with higher incidence was the penetration rate of mobile/cellular telecommunications subscriptions, although other factors were highlighted. According to these ecological results the latency period is at least 11–12 years, but probably more than 20 years. Missing data on cancer incidence and for other potential risk factors prohibit more detailed investigation of exposure–response associations and/or explore other hypotheses.

Conclusions Readily available ecological data may be underused, particularly for the study of risk factors for rare diseases and those with long latencies. The results of ecological analyses in general should not be overinterpreted in causal inference, but equally they should not be ignored where alternative signals of aetiology are lacking.

What this paper adds

- This study shows how existing open-access online databases can be used to explore potential risk factors for rare diseases at an ecological level, and enables timely generation of hypotheses where standard epidemiological study designs may not be able to elucidate risk factors in an early period of emerging risks.
- We show a clear association between national penetration of cellular telecommunications subscriptions and higher incidence of brain and nervous system cancers, with a latency between exposure and clinical onset of at least 11–12 years, but probably more than 20 years.
- This methodology might be used more widely to test the generalisation of existing hypotheses and to generate new ones, especially for rare diseases, where definitive epidemiological studies are infeasible for addressing public health concerns in a timely manner.

INTRODUCTION

Environmental risk factors for cancers receive substantial public attention, and there is a public health need to balance timely generation of hypotheses with cautious causal inference. Cautious and thorough epidemiological studies are required to confirm a causal link between an exposure and a disease outcome. Such studies usually take a long time before the results are shared with the scientific and policy communities; meanwhile, public debates may ensue over the topic under investigation. In a more proactive approach, however, hypotheses about risk factors could be generated from available data that would enable more informed early debate under a precautionary principle. Moreover, faster generation of hypotheses based on routine data may better direct research resources at emerging risk factors.

For most common cancers the main risk factors are well established. This knowledge is largely based on classical study designs with sufficient statistical power to make confident causal inferences. For lung cancer, for example, the most frequently occurring cancer in the world, it is well established that the main risk factor is tobacco smoking, accounting for 75%–90% of the risk, while other main risk factors include exposure to radon, environmental tobacco smoke, asbestos and other occupational exposures.[1] However, for many cancers, especially the rarer ones, most of the aetiology is unknown and for many may involve a complex interaction of demographic, genetic, socioeconomic and environmental risk factors.

The multi-factorial initiation and development of cancers complicate strategies to prevent or reduce

the number of new cases. Epidemiological studies are generally designed to investigate one, or few hypotheses, rather than being of a wide exploratory or hypotheses generating nature. More hypothesis generating strategies are however emerging with the generation of large dataset from high-throughput *omics* and open linked data sources. Therefore, it is of interest to explore alternative methodologies for generating and shaping hypotheses prior to definitive investigation.

Here we consider the hypothesis generating potential[2] of an ecological combination of: (1) open data on incidence and mortality from cancers at national level for all countries of the world, which are available from the International Agency for Research on Cancer's (IARC) GLOBOCAN 2008 project[3] and (2) open information on putative risk factors and related confounding factors collected at a national level, available from the World Bank list of development indicators[4] and from the United Nations Development Report.[5] The approach may be especially useful in studies of rare cancers where it is difficult to obtain studies of sufficient size to study environmental and occupational risk factors. So, we evaluate this open ecological approach for the specific case study of environmental risk factors for malignant brain tumours.

Malignant brain tumours are considered a rare cancer in that they account for only 1%–2% of all cancers in adults.[6] Furthermore, the incidence has been increasing worldwide over the past 3 decades, especially in industrialised societies.[7–10] Ageing populations[9] and better diagnostic methods[11 12] have been related to the rising incidence. Exogenous risk factors, however, have been underinvestigated. The only established environmental risk factor for gliomas is ionising radiation exposure to the head and neck, while in contrast allergies are consistently inversely associated with glioma risk.[13] However, a large number of environmental risk factors have been proposed that may contribute to increasing brain cancer incidence.[6] These include exposure to pesticides,[14 15] metals,[16] polycyclic aromatic hydrocarbons,[17 18] solvents,[19] glues[20] and electromagnetic fields[21]—with radiofrequency exposure from mobile phones gaining most public attention.[22] Although these exogenous factors are potentially amendable to interventions, the studies linking them to brain cancers have been inconsistent in terms of causality.[16]

This absence of evidence provides an important context for the present study to evaluate the epidemiological worth of open data sources exploring potential risk factors at an ecological level to generate or shape hypotheses about possible causal mechanisms.

MATERIALS AND METHODS

Brain cancer incidence data

Data on incidence rates of cancers of the brain and nervous system for all countries of the world were obtained from the open, online GLOBOCAN 2008 resource.[3] Within the GLOBOCAN project, cancer statistics (including numbers of cases, incidence and mortality rates, and cumulative risk) are available at national and further aggregated levels. From here we extracted the age standardised, or age-world-standardised incidence rate (ASR(w)) per 100 000 population, based on weighted averages of the age-specific rates from the 'world standard population',[23] for cancers of the brain and nervous system in the year 2008.

The types of malignant neoplasms of the brain (C71) were coded using the International Statistical Classification of Diseases and Related Health Problems, 10th revision (ICD-10), and included neoplasms of the cerebrum (except lobes and ventricles) (C71.0), frontal lobe (C71.1), temporal lobe (C71.2), parietal lobe (C71.3), occipital lobe (C71.4), cerebral ventricle (excluding the fourth ventricle) (C71.5), cerebellum (C71.6), brain stem (C71.7), and neoplasms in overlapping lesions of the brain (C71.8) or unspecified (C71.9). Within the GLOBOCAN project, these are grouped together with malignant neoplasms of the meninges (C70), including neoplasms of the cerebral meninges (C70.0), spinal meninges (C70.1) and unspecified meninges (C70.9), and with malignant neoplasms of the nervous system (C72), which include neoplasms of the spinal cord (C72.0), cauda equine (C72.1), olfactory nerve (C72.2), optic nerve (C72.3), acoustic nerve (C72.4), other and unspecified cranial nerves (C72.5), and also overlapping lesions of brain and other parts of the central nervous system (C72.8) and unspecified central nervous system malignant neoplasms (C72.9).

Potential risk factor and confounder data

Data on potential risk factors or confounding factors were collected from two open resources: the World Bank list of development indicators[4] and the United Nations Development Report for Human Development Index (HDI)[5] statistics. Potential confounders were defined as available indicators that may (to some extent) account for differences in quality of cancer registration resulting in differences in incidence rates that are not due to differences between populations or risk factors between countries, or proxies thereof. From the list of development indicators, indicators that were included as potential confounders described population demographics, development and quality of healthcare (listed in table 1). Development indicators describing aspects of urbanisation, energy usage, distribution of occupational sectors or environmental/pollution factors were identified as potential environmental risk factors (listed in table 2). Data on each indicator were obtained for the year 2008. To investigate the latency between when biologically relevant exposure to one of the identified risk factors occurred and when the brain cancer was detected we also extracted risk factor data for the year 1998, and where possible additionally for the year 1995 (older data were not available on most risk factors).

Data treatment and statistical analyses

We linked the datasets by country. Of 208 countries, 23 had no data on risk factors, 10 did not have corresponding ASR(w)s and for 10 the ASR(w) was specified as an impossible '0'. These were removed prior to the analyses, resulting in a final sample of 165 countries. Similarly, indicators (risk or confounding factors) for which data were not available for more than 25% of countries were also deleted (completeness of total death reporting (percentage of reported total deaths to estimates total deaths) (missing=79), malnutrition prevalence-weight for age (percentage of children under 5) (missing=149), poverty headcount ratio at national poverty line (percentage of the population) (missing=125), physicians (per 1000) (missing=105), and public spending on education, total (percentage of government expenditure) (missing=104)).

Data were analysed using linear least-squares regression modelling in which the dependent variable ASR(w) being a rate, was $\log_e$-transformed to resemble a Gaussian distribution prior to statistical modelling. A level of $p<0.05$ was taken as evidence of statistical significant influence of an indicator on ASR(w), while $p<0.10$ was treated as 'borderline significance'.

Prior to adding risk factors in the analysis, a 'confounder model' was developed. All potential confounders were analysed separately in a univariate linear regression model and parameter estimates, statistical significance and explained variance (R^2) were registered. Subsequently, potential confounders with

Table 1 Univariate results potential confounding factors

Confounding factors	Corresponding World Bank indicator	1998 (p Value)*	2008 (p Value)*	Countries 1998	Countries 2008
Quality of cancer registration	Completeness of total mortality reporting (% of reported total deaths to estimates total deaths)		0.03 (<0.01)		80
Continent			p<0.01†		165
World (1st, 2nd, 3rd, other)			p<0.01†		165
Human Development Index			4.34 (p<0.01)	NA	154
Gender distribution	Population female (% total)	0.04 (0.30)	0.04 (0.20)	165	165
Age distribution	Population (0–14)	−0.07 (<0.01)	−0.07 (<0.01)	165	165
(% total)	Population (15–64)	0.11 (<0.01)	0.10 (<0.01)	165	165
	Population (>64)	0.14 (<0.01)	0.13 (<0.01)	165	165
General population health	Life expectancy at birth	0.08 (<0.01)	0.08 (<0.01)	162	165
	Mortality rate, under 5 (per 1000)	−0.01 (<0.01) ‡	−0.01 (<0.01)	163	163
	Survival to age 65 (% cohort) (f) §	0.05 (<0.01)	0.05 (<0.01)	165	165
	Survival to age 65 (% cohort) (m)§	0.04 (<0.01)	0.04 (<0.01)	165	165
	Malnutrition prevalence, weight for age (% children under 5)	−0.03 (0.01)¶	−0.09 (0.13)¶	49	5
	Birth rate, crude (per 1000 people)	−0.07 (<0.01)	−0.07 (<0.01)	164	165
	Mortality rate (per 1000) (f)	−0.01 (<0.01)	−0.01 (<0.01)	165	141
	Mortality rate (per 1000) (m)	−0.01 (<0.01)	−0.01 (<0.01)	165	141
Population wealth	Gross national income per capita (PPP ($))	0.00 (<0.01)	0.00 (<0.01)	151	156
	Poverty headcount ratio at national poverty line (% population)	NA	−0.02 (<0.01)	NA	30
Money spent on healthcare	Health expenditure per capita, PPP (constant 2005 international $)	0.00 (<0.01)	0.00 (<0.01)	161	160
	Health expenditure, public (% of government expenditure)	0.08 (<0.01)	0.04 (0.05)	160	161
	Physicians (per 1000)	0.36 (<0.01)	0.62 (<0.01)	76	51
Education	Public spending on education, total (% of government expenditure)	−0.02 (0.40)**	−0.06 (0.02)	80	53

*Dependent variable is log(e) transformed age-adjusted incidence rates per 100 000 population (ASR(w)).
†p Value based on ANOVA.
‡Year 2000 instead of 1998.
§Female (f) or male (m).
¶Years 2000 and 2009 instead of 1998 and 2008.
**Year 1999 instead of 1998.
ASR(w), Age-world-standardised incidence rate; NA, not available.

p values below 0.20 were added to a multiple regression confounder model using a forward selection method. Variables were subsequently kept in the model based on statistical significance ($p<0.05$) and decreased Bayesian Information Criterion. Multicollinearity was evaluated using Variance Inflation Factors (VIF), using a VIF above 10 as a rule of thumb for evidence of significant multicollinearity (realising that this cut-off should be interpreted with caution, so all situations were evaluated individually[24]) in which case the variable with lowest Bayesian Information Criterion was kept in the model. Model-fit was further evaluated by assessing distributions and trends of the residuals, half-normal probability plots and evaluation of influential observations using Cook's distance.

The 'risk factor' model was subsequently generated by adding independent risk factors to the final confounder model, using a similar model-building approach and evaluation criteria as described for the confounder model.

Additional exploratory and sensitivity analyses using the final model were done by stratification by continent, and by evaluation of different latencies (0–28 years). After obtaining the final multivariate regression model, the data were explored for non-linearity in covariate response by sequentially replacing each covariate by a quadratic cubic spline function using 1–3 knots using the *lmer* function in R.

RESULTS

Univariate analyses of all confounder and risk factor variables for the indicative years 1998 and 2008 indicate that many are, at face value, correlated to 2008 age-world standardised brain cancer incidence. These results are summarised in tables 1 and 2 for potential confounders and risk factors, respectively. For the majority, but not all, of the potential risk factors for which data were available from 1998, data were also available from 1995, and the results of these univariate analyses are provided in online supplementary table S1. Development of the multiple regression confounder model indicated many potential confounding factors were highly correlated with one another, but following the model-building strategy described above resulted in a confounder model that best described these data and included continent, 2008 gross national income and the HDI score as covariates (see online supplementary table S2). This confounder model explained about 67% of the variation between countries with the highest VIF of 5.22 for HDI, which indicated low to moderate multicollinearity only.

The univariate results for risk factors where data were available from 1995 and 1998 were largely similar, so to allow for another 3 years of latency the 1995 data of potential risk factors were subsequently added to the confounder model (table 3). These results indicate that ($p<0.10$) the national incidence of brain cancer is correlated to the proportion of people in large (>1 million people) urban agglomerations, but also to the percentage of the working population working in agriculture, electric power consumption, mobile cellular subscriptions (per 100 people), internet connections per 100 people, and the percentage of combustible renewables and waste (ie, solid biomass, liquid biomass, biogas, industrial waste and municipal waste) in

Table 2 Univariate results potential environmental risk factors

Risk factors	Corresponding World Bank indicator	1998 (p value)*	2008 (p value)*	Countries 1998	Countries 2008
Urbanisation (% population)	Urban population	0.02 (<0.01)	0.02 (<0.01)	165	165
	Population in urban agglomerations of more than 1 million	0.01 (0.09)	0.01 (0.20)	106	106
	Rural population	−0.02 (<0.01)	−0.02 (<0.01)	165	165
Energy usage	Energy use (kg of oil equivalent per capita)	0.00 (<0.01)	0.00 (<0.01)	129	129
Electrical usage	Electric power consumption (kWh per capita)	0.00 (<0.01)	0.00 (<0.01)	128	128
Cell phone use	Mobile cellular subscriptions (per 100 people)	0.04 (<0.01)	0.01 (<0.01)	162	163
Internet/broadband	Fixed broadband internet subscribers (per 100 people)	0.02 (0.99)	0.06 (<0.01)	52	161
	Internet users (per 100 people)	0.06 (<0.01)	0.03 (<0.01)	155	163
Employment (% total employment)	Agriculture	−0.01 (<0.01)	−0.02 (0.04)	84	26
	Industry	0.03 (<0.01)	−0.00 (0.95)	84	26
	Services	0.01 (0.07)	0.01 (0.05)	84	26
Environment/pollution					
Pesticides	Fertilizer consumption (kg per hectare of arable land)	NA	0.00 (0.03)	NA	139
Waste	Combustible renewable and waste (% of total energy)	0.00 (0.69)	−0.00 (0.57)	129	129
Organic water pollutant	Organic water pollutant (BOD) emissions (kg/day)	−5.29 (0.01)	NA	66	NA
Other greenhouse gas emissions	PFC/HFC/SF6 gas emissions (thousand metric tons of CO_2 equivalent)	0.00 (0.21)†	0.00 (0.22)†	129	129
	PM10, country level (ug/m^3)	−0.01 (<0.01)	−0.01 (0.01)	159	156
Traffic	Road sector gasoline Fuel consumption per capita (kt of oil equivalent)	0.00 (0.30)	0.00 (0.32)	128	129
	Road sector diesel fuel consumption per capita (kt of oil equivalent)	0.00 (0.15)	0.00 (0.08)	128	129
	Motor vehicles (per 1000)	NA	0.00 (<0.01)	NA	77
Smoking (% adults)	Smoking prevalence (f)		0.05 (<0.01)‡	NA	124
	Smoking prevalence (m)		0.03 (<0.01)‡	NA	127

*Dependent variable is log(e) transformed age-adjusted incidence rates per 100 000 population (ASR(w)).
†Years 2000 and 2005 instead of 1998 and 2008.
‡Year 2006 instead of 2008.
ASR(w), Age-world-standardised incidence rate; NA, not available; BOD, biochemical oxygen demand; PPP, purchasing power parity; ANOVA, analysis of variance; PFC, perfluorocarbons.

1995 but not 2008. Conversely, ASR(w) was further associated with the number of motor vehicles per 1000 people in 2008 but not in 1995. Note, however, that all associations were evaluated with varying numbers of countries where these data were collected. Further multiple regression modelling using the steps outlined above uncovered just one risk factor that was consistently correlated with increased incidence of cancers of the brain and central nervous system in 2008: the number of mobile phone subscriptions per 100 people (table 4), which indicated about a 4% increase in brain cancer incidence in 2008 for each additional per cent of mobile phone subscriptions in 1995 (β~0.04 (SE~0.02), p~0.04). Again, only low to moderate collinearity (VIF≤6) between covariates was present. Additional stratification (table 5) by continent indicates that these trends can primarily be observed in European countries (although only Africa, Asia and Europe had data from enough countries). Crude associations between the one remaining significant risk factor, mobile phone subscriptions per 100 people, and ASR(w) are shown graphically in figure 1A,B.

Additional exploration of the latency period (figure 1D, and see online supplementary table S3), including the number of mobile subscribers per 100 people for each year between 1980 and 2008 (where possible) and brain cancer incidence in 2008, shows that statistical significant correlations could be observed until 1996. This was similar when all countries were used to analyse the effects of latency or only those 17 countries that had data on mobile phone subscriptions for 1995 (see online supplementary material).

Graphical assessment of the residuals (figure 1C) indicates a relatively good model fit of the final model, although some outliers are present. Sensitivity analyses using quadratic cubic splines to evaluate non-linearity in the confounder/exposure–response associations did not indicate improved fit of the models, based on Akaike Information Criterion values, or changes in the association between mobile phone penetration rate and brain cancer incidence. However, this could primarily be attributed to the relatively limited number of data points (data not shown).

DISCUSSION

In this study, we explored the epidemiological value of combining open data sources on cancer incidence from the GLOBOCAN project, development indicators from the World Bank list and United Nations to explore potential environmental risk factors of malignant neoplasms of the brain and central nervous system at an ecological level. This approach widened the generalisation of an existing hypothesis and highlighted new hypotheses for attention, and as such, this general approach may be applicable more widely, particularly to other rare diseases.

In our case study, cancers of the brain and nervous system represent a rare outcome, and relating it to mobile phone use at individual level is impractical due to: (1) the large numbers required to achieve sufficient statistical power; (2) problems with the accurate assessment of exposure;[25] and (3) difficulty in identifying controls given the advancing ubiquity of mobile phones. This work confirms that mobile phone use may be a risk factor, thereby confirming previous ecological findings from the USA.[26] Although our analyses indicate a relatively small risk, explaining only about 1% of the variation in incidence rates between countries, it is supported by data from individual-level studies.[27] A causal association, however, has not been confirmed by other studies[28–30] and remains controversial.[31–34] Overall,

Table 3 Risk factors added to confounder model

Environmental risk factor	Year	p Value	BIC	dF	VIF
Urban population (% total)*	2008	0.29	332.74	138	2.74
	1995	0.43	333.25	138	3.14
Population in urban agglomerations of more than 1 million (% of total population)	2008	0.01	210.36	86	2.39
	1995	0.01	181.72	87	2.50
Employment in agriculture (% of total employment)	2008	0.17	33.87	20	4.34
	1995	0.03	124.53	69	3.47
Employment in industry (% of total employment)	2008	0.90	36.42	20	2.14
	1995	0.06	125.71	69	1.71
Employment in services (% of total employment)	2008	0.36	35.30	20	3.73
	1995	0.26	128.34	69	4.17
Energy use (kg of oil equivalent per capita)	2008	0.98	251.25	108	2.31
	1995	0.86	251.21	108	4.33
Electric power consumption	2008	0.23	249.62	108	2.00
	1995	0.03	281.27	117	3.76
Mobile cellular subscriptions (per 100 people)	2008	0.28	323.86	136	3.91
	1995	0.04	322.89	135	2.67
Internet connections per 100 people	2008	0.32	327.43	136	7.41
	1995	0.04	216.57	97	1.75
Fertilizer consumption (kg per hectare of arable land)	2008	0.61	287.84	116	1.48
	1995	NA			
Combustible renewables and waste (% of total energy)	2008	0.32	250.15	108	1.10
	1995	0.06	247.34	108	3.98
Organic water pollutant (BOD) emissions (kg/day)	2008	NA			
	1995	0.15	93.60	36	1.19
Greenhouse gases (PFC /HFC/SF6 gas emissions (thousand metric tons of CO_2 equivalent))	2000	0.90	251.23	108	1.43
	1995	0.76	251.16	108	1.39
Road sector gasoline fuel consumption per capita (kt of oil equivalent)	2008	0.91	251.23	108	1.62
	1995	0.19	249.34	108	5.59
Road sector diesel fuel consumption per capita (kt of oil equivalent)	2008	0.83	251.20	108	1.41
	1995	0.62	250.97	108	2.72
Motor vehicles (per 1000)	2008	0.07	135.89	64	6.05
	1995	NA			
PM_{10}, country level ($\mu g/m^3$)	2008	0.52	323.26	132	1.33
	1995	0.42	325.31	134	1.51
Smoking prevalence (% adults) (female)	2006	0.66	257.97	103	3.95
	1995	NA			
Smoking prevalence (% adults) (male)	2006	0.42	259.19	105	2.42
	1995	NA			

*Note that urban population (% total) and rural population (% total) add up to 100%, so rural variable not added.
BIC, Bayesian Information Criterion; NA, not available; VIF, Variance Inflation Factors.

Table 4 Final multivariate results

R^2_{adj} ~0.68	BIC ~322.89	dF=135			
ANOVA	**dF**	**SS**	**F value**	**Pr>F**	**R^2**
GNI.08	1	51.508	134.193	<0.01	0.04
Continent	7	55.012	20.475	<0.01	0.09
HDI	1	13.505	35.183	<0.01	0.09
cellphone.95	1	1.606	4.185	0.043	0.01
Residuals	135	51.817			
Variable		**β**	**SE**	**p Value**	**VIF**
Intercept		−1.81	0.28	<0.01	
GNI.08		−0.00	0.00	<0.01	6.02
Continent	Asia	0.70	0.18	<0.01	2.26
	Central Americas	0.75	0.27	0.01	1.47
	Caribbean	0.46	0.30	0.12	1.33
	Europe	1.29	0.23	<0.01	3.86
	North America	0.91	0.44	0.04	1.49
	Oceania	−0.18	0.34	0.60	1.18
	South America	0.48	0.25	0.06	1.80
	Africa	–	–	–	
HDI		3.92	0.64	<0.01	5.29
cellphone.95		0.04	0.02	0.04	2.67

BIC, Bayesian Information Criterion; cellphone.95, mobile cellular subscriptions (per 100 people) in 1995; GNI, gross national income; GNI.08, gross national income per capita (2008); HDI, Human Development Index; VIF, Variance Inflation Factors.

however, the conclusion of the International Agency for Research on Cancer Monograph working group indicated that limited evidence for an association existed.[22] Our analyses further suggest that the latency between relevant exposure (mobile phone use) and clinical manifestation of the disease (brain and nervous system malignancies) is (at population level) at the very least 11–12 years but should ideally be more than 20 years, which is not reflected in most study designs.

Mobile phone use may be a proxy for another risk factor that correlates with mobile phone use but was not included in the available databases, or was available but was of inferior quality. Given that correlations were also noted in intermediate analyses for national electricity consumption, internet usage and population in urban agglomerations of more than 1 million we hypothesise that if this is the case, this is most likely to be some, yet unknown, factor related to urbanisation and development. However, given that mobile phone use both remained a significant factor independent of the inclusion of other potential risk factors in multivariate modelling and is also in broad agreement with some analytical studies to us indicates this may well be the most important exposure for further study; in agreement with conclusions reached by others.

Our approach further highlighted several other potential risk factors that may be associated with increased risk of brain cancer, namely: populations in urban agglomerations of more than 1 million; percentage of the working population employed in agriculture; percentage of the working population employed in industry; national electricity consumption; internet usage; and combustible renewables and waste. A correlation between urbanisation/population density and increased brain cancer risk has been reported before[35 36] but is inconsistent with other data[37 38]

Table 5 Stratification by continent (cell phone use in 1995)

Continent	β*	SE	p Value	dF	Rt_{adj}
Africa	0.20	0.97	0.84	18	0.40
Asia	0.00	0.04	0.98	23	0.31
Europe	0.039	0.010	<0.001	30	0.40
Central America	NA[2]				
Caribbean	NA				
South America	NA				
North America	NA				
Oceania	NA				

*Adjusted for gross national income per capita (2008) and Human Development Index.
†Not enough (<10) countries with available data.
NA, notavailable.

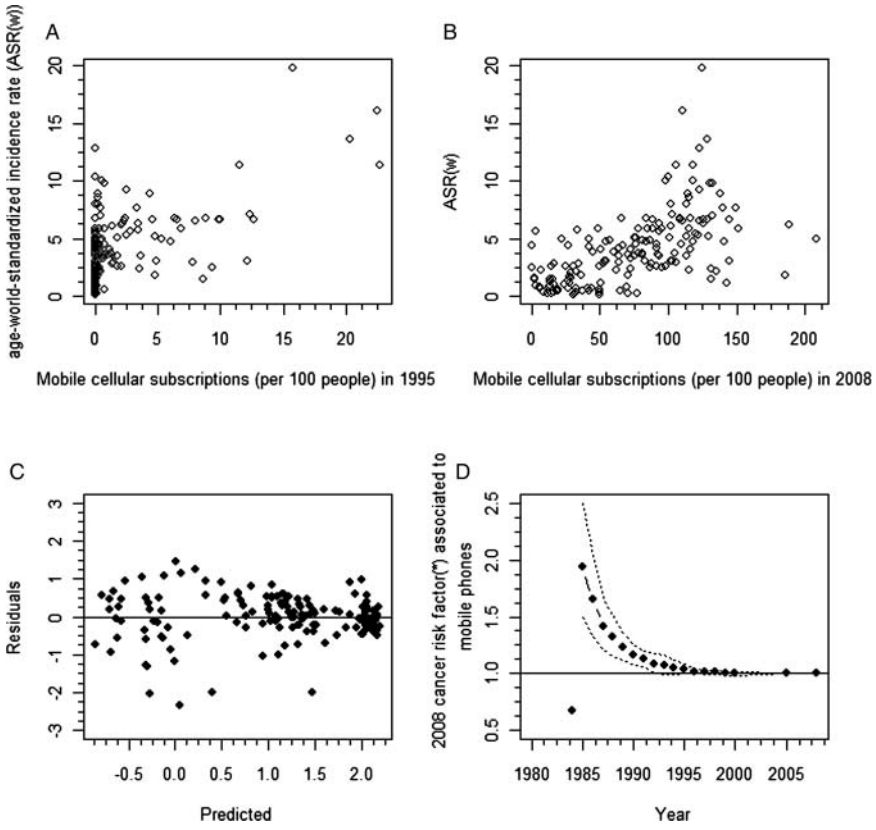

Figure 1 Age-world-standardised incidence rate (ASR(w)) and mobile cellular subscripts (A and B), residuals versus predicted values of the confounder model (C), and risk factors (95% CI) of change in national incidence of cancer of the brain or nervous system in 2008 related to one additional mobile cellular subscription per 100 people in year x (D).

and is likely rather a surrogate for another risk factor.[39] Industry, and more specifically metal,[19] electrical/electronics[40] and textile,[41] has been reported previously. Agriculture has also previously been reported as a risk factor for brain cancer.[42] National electricity consumption and internet usage (broadband and wireless) could be interpreted as indicators of electromagnetic field exposure, although primarily to 50/60 Hz extremely low-frequency.[43] Although speculation, associations with combustible renewables and waste may be associated with biological agents similar to exposures encountered in agriculture.

However, after adjustment for confounding factors these additional potential risk factors that were identified could not be investigated further because of issues of multicollinearity or because not enough data from different countries were available. Lack of data on many risk factors, especially for earlier years, is one of the main limitations of this approach and prohibits more detailed exploration of many risk factors or inclusion of additional confounding factors. This limitation is a general issue for ecological studies using similar data sources. In addition, we highlight the need for consistent collection and collation of such routine data across nations. In the later years of this study, as linkable data became more available, the potential for ecological study increased considerably.

Because of the lack of data from many countries, it was not possible to evaluate non-linearity of the exposure–response associations. This may have contributed to some outliers in the model fit and an underlying heteroscedasticity. If the current trends toward fuller datasets continue then future evaluations of non-linearity and better model fitting may be possible. At present we recommend that linearity is assumed when using such databases to identify potential risk factors, and that properly powered studies with individual-level data be carried out to evaluate non-linearity of exposure–response associations.

Another limitation of the approach described here is the difference in the quality of cancer registration between different countries. For example, incidence data may cover entire national populations but may also, especially in developing countries, cover subnational areas or major cities.[44] It has been reported that in 2006 only about 21% of the world population was actually covered by population-based cancer registries,[45] and only about 8% of the world population by 'good quality' registries (matching CI5 criteria). Sparse registration is most pronounced in Asia (8% of the total population) and Africa (11%),[45] which may in these analyses have resulted in differential impact of residual confounding. Brain cancers (more specifically the incidence of glioma) have indeed been reported to be strongly related to social and economic factors.[46] We aimed to adjust for such differences by adjusting the models for 'Continent', 'Gross National Income in 2008' and 'Human Development Index values', which yielded the best available confounder model explaining about 67% of variation in incidence between countries. However, there may still have been significant residual confounding present. Although we cannot rule this out, we are reassured by the fact that the trend was most notable within Europe where differences in quality of cancer registration are expected to be minimal compared with between-continent differences.

We used weighted age-standardised incidence rates provided by the GLOBOCAN project, and originally based on the methodology outlined by Doll *et al*.[23] As such, we are comparing brain cancer incidence rates as if they were from countries with similar age distributions. Although these require additional assumptions on the age distribution of each country, these have been argued to be most useful for comparing incidence rates over time or between countries since they remove the effects of historical events such as wars and famine.[47] Alternatively, crude

incidence rates were also provided within GLOBOCAN. The differences are relatively small, but nonetheless may have influenced the results.

Ecological studies compared with individual analytic studies are suspect of unavoidable bias.[48] Most notably, 'ecological fallacy' (aggregation bias) indicates that average exposure to a group of people does not, generally, determine their average risk.[49] It has been argued however that in some situations ecological associations can be closer to the true effect than individual-level associations because the latter are themselves also subject to many biases,[50] which may be important in this particular case where mobile phone use was investigated as a risk factor.[51 52] As such, it is important to re-emphasise that ecological studies should not be used to infer causality in a policy context, but the results can, and arguably should, guide further research.

CONCLUSIONS

In conclusion, readily available data may be underused for ecological studies, particularly for exploring risk factors for rare diseases and those with long latency times. These data can, through linkage of different sources, be exploited systematically to explore potential risk factors and to further the exploration of established risk factors for (rare) diseases. We have demonstrated this for cancers of the brain and central nervous system and mobile phone use. In general, the results of ecological studies should not be overinterpreted in causal inference, but equally they should not be ignored where alternative signals of aetiology are lacking.

Acknowledgements The authors would like to thank Dr. Kurt Straif for his insights and discussion on the draft versions of this manuscript.

Contributors FV conceived the study. Study protocol finalised by all authors. FV and KH conducted data collection and statistical analyses. All authors were involved in interpretation of the results. FV wrote the first draft and all authors contributed to subsequent and finals drafts. All authors approved the final version of the manuscript to be published.

Competing interests All authors have completed the Unified Competing Interest form at http://www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous 3 years; and no other relationships or activities that could appear to have influenced the submitted work.

Provenance and peer review Not commissioned; externally peer reviewed.

Funding No external funding was obtained for this study.

Data sharing statements The data used in this work are open-access.

REFERENCES

1 Mc Erlean A, Ginsberg MS. Epidemiology of lung cancer. *Semin Roentgenol* 2011;46:173–7.
2 Piantadosi S. Invited commentary: ecological biases. *Am J Epidemiol* 1994;139:761–3.
3 Ferlay J, Shin HR, Bray F, *et al*. *GLOBOCAN 2008 v1.2, cancer incidence and mortality worldwide: IARC cancerbase No. 10*. [Internet]. Lyon, France: International Agency for Research on Cancer, 2010. http://globocan.iarc.fr (accessed 2 Feb 2012).
4 The World Bank. World Bank Development Indicators. http://data.worldbank.org/about (accessed 2 Feb 2012).
5 UNDP. Human Development Reports data. HDI 2010 Index. http://hdr.undp.org/en/media/Lets-Talk-HD-HDI_2010.pdf (accessed 2 Feb 2012).
6 Ostrom QT, Barnholtz-Sloan JS. Current state of our knowledge on brain tumor epidemiology. *Curr Neurol Neurosci Rep* 2011;11:329–35.
7 Hess KR, Broglio KR, Bondy ML. Adult glioma incidence trends in the United States, 1977–2000. *Cancer* 2004;101:2293–9.
8 Kaneko S, Normura K, Yoshimura T, *et al*. Trend of brain tumor incidence by histological subtypes in Japan: estimation from the Brain Tumor Registry of Japan, 1973–1993. *J Neurooncol* 2002;60:61–9.
9 De Vocht F, Burstyn I, Cherrie JW. Time trends (1998–2007) in brain cancer incidence rates in relation to mobile phone use in England. *Bioelectromagnetics* 2011;32:334–9.
10 Deltour I, Johansen C, Auvinen A, *et al*. Time trends in brain tumor incidence rates in Denmark, Finland, Norway, and Sweden, 1974–2003. *J Natl Cancer Inst* 2009;101:1721–4.
11 Desmeules M, Mikkelsen T, Mao Y. Increasing incidence of primary malignant brain tumors: influence of diagnostic methods. *J Natl Cancer Inst* 1992;84:442–5.
12 Polendak AP. Interpretation of secular increases in incidence rates for primary brain cancer in Connecticut adults, 1965–1988. *Neuroepidemiology* 1996;15:51–6.
13 Bondy ML, Scheurer ME, Malmer B, *et al*. Brain tumor epidemiology: consensus from the brain tumor epidemiology consortium. *Cancer* 2008;113(Suppl 7):1953–68.
14 Bohnen NI, Kurland LT. Brain tumor and exposure to pesticides in humans: a review of the epidemiological data. *J Neurol Sci* 1995;132:110–21.
15 Lee WJ, Colt JS, Heinema EF, *et al*. Agricultural pesticide use and risk of glioma in Nebraska, United States. *Occup Environ Med* 2005;62:786–92.
16 Schlehofer B, Hetttinger I, Ryan P, *et al*. Occupational risk factors for low and high grade glioma: results from an international case control study of adult brain tumours. *Int J Cancer* 2005;113:116–25.
17 Buffler PA, Kelsh M, Chapman P, *et al*. Primary brain tumor mortality at a petroleum exploration and extraction research facility. *J Occup Environ Med* 2004;46:257–70.
18 Cocco P, Heineman EF, Dosemeci M. Occupational risk factors for cancer of the central nervous system among US women. *Am J Ind Med* 1999;36:70–6.
19 Rodvall Y, Ahlbohm A, Spannare B, *et al*. Glioma and occupational exposure in Sweden, a case-control study. *Occup Environ Med* 1996;53:526–32.
20 Spinelli V, Chinot O, Cabaniols C, *et al*. Occupational and environmental risk factors for brain cancer: a pilot case-control study in France. *Presse Med* 2010;39:e34–44.
21 Schüz J, Lagorio S, Bersani F. Electromagnetic fields and epidemiology: an overview inspired by the fourth course at the International School of Bioelectromagnetics. *Bioelectromagnetics* 2009;30:511–24.
22 Baan R, Grosse Y, Lauby-Secretan B, *et al*. Carcinogenicity of radiofrequency electromagnetic fields. *Lancet Oncol* 2011;12:624–6.
23 Doll R, Payne P, Waterhouse JAH, eds. *Cancer incidence in five continents*. Vol. I. Geneva: Union Internationale Contre le Cancer, 1966.
24 O'Brien RM, Robert M. A caution regarding rules of thumb for variance inflation factors. *Qual Quantity* 2007;41:673–90.
25 Auvinen A, Toivo T, Tokola K. Epidemiological risk assessment of mobile phones and cancer: where can we improve? *Eur J Cancer Prev* 2006;15:516–23.
26 Lehrer S, Green S, Stock RG. Association between number of cell phone contracts and brain tumor incidence in nineteen U.S. States. *J Neurooncol* 2011;101:505–7.
27 Hardell L, Carlberg M, Hansson Mild K. Pooled analysis of case-control studies on malignant brain tumours and the use of mobile and cordless phones including living and deceased subjects. *Int J Oncol* 2011;38:1465–74.
28 Schüz J, Steding-Jessen M, Hansen S, *et al*. Long-term mobile phone use and the risk of vestibular schwannoma: a Danish nationwide cohort study. *Am J Epidemiol* 2011;174:416–22.
29 INTERPHONE Study Group. Brain tumour risk in relation to mobile telephone use: results of the INTERPHONE international case–control study. *Int J Epidemiol* 2010;39:675–94.
30 Larjavaara S, Schüz J, Swerdlow A, *et al*. Location of gliomas in relation to mobile telephone use: a case-case and case-specular analysis. *Am J Epidemiol* 2011;174:2–11.
31 Aydin D, Feychting M, Schüz J, *et al*. Impact of random and systematic recall errors and selection bias in case—control studies on mobile phone use and brain tumors in adolescents (CEFALO study). *Bioelectromagnetics* 2006;32:396–407.
32 Peres J. WHO classification sparks debate over cell phone safety. *J Natl Cancer Inst* 2011;103:1146–7.
33 Cardis E, Sadetzki S. Indications of possible brain-tumour risk in mobile-phone studies: should we be concerned? *Occup Environ Med* 2011;68:169–71.
34 Olsen J. The interphone study: brain cancer and beyond. *Bioelectromagnetics* 2011;32:164–7.
35 Nasca PC, Burnett WS, Greenwald P, *et al*. Population density as an indicator of urban-rural differences in cancer incidence, upstate New York, 1968–1972. *Am J Epidemiol* 1980;112:362–75.
36 Bidoli E, Franceschi S, Montella M. Cancer mortality by urbanization and proximity to the sea coast in Campania Region, southern Italy. *Tumori* 1998;84:460–6.
37 Yang CY, Hsieh YL. The relationship between population density and cancer mortality in Taiwan. *Jpn J Cancer Res* 1998;89:355–60.
38 Schouten LJ, Meijer H, Huveneers JA, *et al*. Urban-rural differences in cancer incidence in The Netherlands 1989–1991. *Int J Epidemiol* 1996;25:729–36.
39 Monroe AC, Ricketts TC, Savitz LA. Cancer in rural versus urban populations: a review. *J Rural Health* 1992;8:212–20.
40 Kheifets L. Electric and magnetic field exposure and brain cancer: a review. *Bioelectromagnetics* 2001;22(Suppl 5):S120–31.
41 Heineman EF, Gao YT, Dosemeci M, *et al*. Occupational risk factors for brain tumors among women in Shanghai, China. *J Occup Environ Med* 1995;37:288–93.

42 Khuder SA, Mutgi AB, Schaub EA. Meta-analyses of brain cancer and farming. *Am J Ind Med* 1998;34:252–60.
43 Kheifets L, Ahlbom A, Crespi CM, *et al*. A pooled analysis of extremely low-frequency magnetic fields and childhood brain tumors. *Am J Epidemiol.* 2010;172:752–61.
44 Ferlay J, Shin HR, Bray F, *et al*. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* 2010;127:2893–917.
45 Parkin DM. The evolution of the population-based cancer registry. *Nat Rev Cancer* 2006;6:603–12.
46 Clouston SAP. Social and economic patterning in the Interphone study. *Int J Epidemiol* 2011;40:1122.
47 Ahmad OB, Boschi-Pinto C, Lopex AD, *et al*. Age standardization of rates: A new WHO standard. GPE Discussion paper series: Mo.31. EIP/GPE.EDB World Health Organization. 2001. http://www.who.int/healthinfo/paper31.pdf (accessed 23 Aug 2012).
48 Greenland S, Morgenstern H. Ecological bias, confounding, and effect modification. *Int J Epidemiol* 1989;18:269–74.
49 Robinson WS. Ecological correlations and the behaviour of individuals. *Am Sociol Rev* 1950;15:351–7.
50 Greenland S, Robins J. Accepting the limits of ecological studies: Drs. Greenland and Robins reply to Drs. Piantadosi and Cohen. *Am J Epidemiol* 1994;139:769–71.
51 Vrijheid M, Armstrong BK, Bédard D, *et al*. Recall bias in the assessment of exposure to mobile phones. *J Expo Sci Environ Epidemiol* 2009;19:369–81.
52 Aydin D, Feychting M, Schüz J, *et al*. Mobile phone use and brain tumors in children and adolescents: a multicenter case-control study. *J Natl Cancer Inst* 2011;103:1264–76.

Environmental risk factors for cancers of the brain and nervous system: the use of ecological data to generate hypotheses

Frank de Vocht, Kimberly Hannam and Iain Buchan

Occup Environ Med 2013 70: 349-356 originally published online January 23, 2013
doi: 10.1136/oemed-2012-100954

Notes