

Attributable fraction: example,
cancers due to occupation in the US

Kyle Steenland

Rollins School Public Health
Emory University, Atlanta, Ga

Definition

The population attributable fraction (AF) estimates the proportion of cases (or deaths) from a disease which could be avoided if exposure were eliminated.

The AF is calculated using the relative risk from exposure (risk of disease for the exposed divided by risk for the non-exposed) and population prevalence of an exposure (proportion exposed).

AFs can then help to inform public health intervention

Here I will present

- 1) methods and issues in calculating the AF
- 2) A specific example, the AF for occupational cancer from US*

*Another good example of AFs for occupational cancer in England (Rushton et al 2010, which estimated 5.2% of cancer deaths attributable to past exposure).

There are two ways to calculate the AF.

One requires access to the raw data from a study (or studies) which generates the RR and the prevalence of exposure. The calculated AF strictly pertains only to that study population, although it is often used for other populations.

The other AF – more common - can be calculated using one RR or many RRs from the literature, without access to the raw data - and also estimates of exposure prevalence from the literature. We will focus on that measure, although it is not as 'clean' as the first measure and can be somewhat inaccurate.

$$AF = (R - R_0)/R \quad (1)$$

where R is the risk of disease, and R_0 is the risk of disease in nonexposed only. This is the proportion of the overall risk of disease due to exposure, which is equivalent to the proportion of cases which occur due to exposure. Formula 1 can be re-written

$$AF = p(RR - 1) / 1 + p (RR - 1) \quad (2)$$

Where p is the proportion of the population exposed, and RR is the relative risk of disease (risk in exposed/risk in nonexposed, or R_1/R_0)*.

*This expression is derived by substituting $R = p(R_1) + (1-p)(R_0)$ into (1), dividing by R_0 to get RR s.

Example

	case	not case	total
exp	50	450	500
Not exp	50	1450	1500
	100	1900	2000

$$RR = 50/500 / 50/1500 = 3$$

$$AF \text{ formula } 2 = [(500/1500)(3-1)] / [(500/1500)(3-1) + 1] = 0.40$$

40% of disease in this study is attributable to exposure. That means that 40 out of 100 cases would not occur if exposure were eliminated.

Using formula 2, we need only to agree on what the RR for a given exposure is, and what the prevalence of exposure is in the population of interest (eg Canada).

However, a number of issues arise which must be decided along the way.

Suppose we want the “all cancers combined” AF

Cancer is not one disease, but many

To get the “all cancers combined AF” in practice we

- 1) calculate the AF for specific cancers
- 2) use the AF and the total number of cancer cases to determine the number of cases which would be avoided if exposure were eliminated, for each cancer
- 3) add the avoidable cases together across all cancers
- 4) divide the avoidable cases by total cancer cases to get all cancer AF

Which cancer-causing exposures (occupational carcinogens) should be included (this in turn decides which cancers)?

In our US work included all occupational exposures that IARC determined were class 1 human carcinogens (definite). We also included one class 2A (probable) carcinogen (diesel fumes)

Rushton et al. in their recent work in England included all class 1 and class 2A IARC carcinogens, more than we did. This is the main reason they had a higher estimate of the occupational cancer AF in England than we did in the US (5.3% vs 3.6% cancers due to occupational exposures). For example, we did not include shift work and breast cancer (class 2A); this IARC classification had not yet been made when we did our estimate for the US

Where do we get the RRs?

We take the RRs from the scientific literature. Sometimes there is one good study giving the RR. Sometimes there are many good studies and we take some kind of average RR from these studies (eg, via a meta-analysis). Or we can use the range of RRs and report a range of AFs

Example, an average RR for lung cancer for workers exposed to asbestos across many studies might be 2.0

Where do we get the proportion exposed?

We may have an estimate of the proportion of the general population **currently** exposed to a given carcinogen eg, asbestos. This proportion, however, does not take into account past exposure. We want to include workers who are alive today and at risk of cancer from a previous occupational exposure.

In our work in the US we had an estimate of the proportion of workers exposed to many workplace toxins in the 1980s from NIOSH, but we wanted to estimate the AF for the US population in 1997.

Where do we get the proportion exposed?

We assumed

- 1) The proportion exposed in the 1980s was similar in the previous years
- 2) A 40 year latency (risk) period, so that workers exposed from 1957-1997 were of interest
- 3) There was a 10% occupational turnover yearly over 40 years, and a 10% loss for mortality over that time
- 4) We multiplied the estimate from the 1980s by 4 to estimate the proportion of the US 1997 population with past exposure

Multiple carcinogenic exposures for the same cancer?

Many agents cause lung cancer, such as asbestos, nickel, cadmium, chromium VI, etc.

Assuming little overlap (exposed to one but not another), the AF for multiple agents can be calculated by calculating the AF for each one and then combining them by

$$AF_{\text{all}} = 1 - (1 - AF_1)(1 - AF_2)(1 - AF_3)$$

If the AFs are all small, simply adding them together will get about the same answer, eg , 5%, 5%, 5%

$$AF_{\text{all}} = 1 - (.95)(.95)(.95) = .143, \text{ about the same as } .05 + .05 + .05 = .15$$

Does exposure level matter?

Many AFs are calculated for exposed vs non-exposed. However, it would be better if one had RRs for high exposure and the proportion highly exposed to derive an AF, and a second RR for low exposure and a proportion exposed to low levels, to calculate a 2nd AF for low exposure.

Typically, the RR will be higher for higher exposure, but the proportion exposed will be lower for high exposures.

Often however we do not have sufficient data to calculate separate AFs by exposure level; often the proportions exposed to high levels (defined how?) vs low levels are not available..

Mortality vs morbidity?

AFs are equally valid for cancer incidence data or cancer mortality data, if you can assume that the RRs for cancer mortality are the same for cancer morbidity. This is generally a reasonable assumption.

Most RRs from the literature are for cancer mortality, but a good number also exist for cancer incidence.

If you have good national cancer incidence data, it will be worthwhile to calculate the number of incident cancers due to current and past occupational exposures, as well as the numbers of cancer deaths due to occupational exposure.

YLLs and DALYs?

It is possible to consider the burden due to occupational cancer not only in the number of avoidable deaths, or the number of avoidable cases, but also in terms of years of life lost (YLLs) for mortality, or disability life years (DALYs) after cancer incidence (ie, morbidity).

YLLs

Assume the average age of death from lung cancer is 70, and the average life expectancy is 80. Assume there are 50,000 lung cancer deaths per year. That means that are 500,000 YLLs for lung cancer. To determine the proportion of YLLs due to occupational carcinogens causing lung cancer, multiply the YLLs by the AF for lung cancer (eg 10%). This would mean that 50,000 years of life are lost each year due to past exposure to occupational carcinogens.

YLLs can account for the different impact on premature mortality of different diseases.

DALYs = YLLs + YLDs, accounting for loss due to both premature mortality and disability after disease incidence

YLDs are years of live lived with disability

DALYs are more controversial, because they are based on a subjective weight given to measure the disability due to living with a given disease. More disability=higher weight.

WHO assigns a weight of 0.81 to each year lived with a late stage cancer, 0.15 for an early stage lung cancer, 0.9 for an early stage bladder cancer. Compared to 0.64 for a year lived with Alzheimer's, or 0.50 for a year with AIDs.

Then $YLD = (\text{annual incident cases} * \text{weight} * \text{mean duration disease})$

Past or future?

AFs and number of avoidable cancer deaths refer to past exposures and current burden.

It may be worth considering future burden, which means considering exposures today and likely exposures in the future. In many cases the classic carcinogens have been greatly reduced in the workplace, but other new exposures may be widespread and carcinogenic. In particular shift work and breast cancer is a potentially important new early stage bladder cancer.

It may also be useful to estimate the cost of future occupational cancers vs. the cost of prevention

US example

We used mortality data only

We did not calculate YLLs or YLDs or DALYs

We used only class 1 IARC carcinogens, except diesel fumes (conservative)

We calculated only one AF per carcinogen, ignoring exposure level

We calculated AFs and number of avoidable deaths for all work-related causes, not just cancer

Exposure data

When using RRs from cohort studies, we used estimates of the worker population exposed taken from a NIOSH surveys in the early 1980s (NOES) of 4500 workplaces, representing about one third of employed civilians.

In some instances, when RRs came from US case-control studies representative of the US, the proportion exposed was taken from the controls in that study (or studies).

RRs

RRs were taken almost exclusively from US studies. We used a range of RRs rather than calculating a midpoint or average RR. If only a single study was available we used the 95% CI as the range.

Calculation of AFs

In many cases we calculated the AFs ourselves, but in other cases we used a range of AFs calculated in the literature, if we thought they were adequate.

Cancer	Exposure	% exposed	RRs
Lung	Asbestos+ chemicals*	from literature	AFs from lit
Lung	ETS at work, nonsmokers	20% of nonsmokers	1.3
Lung	Indoor radon at work	used estimate # ca's due to indoor radon in homes, and ratio of dose home/work (5 to 1)	n.a.
Bladder	Aromatic amine dyes	from literature	AFs from lit
Mesothelioma	Asbestos	n.a., used % of meso deaths (85-90% male, 23- 90% female)	n.a., used % meso deaths
Leukemia	Benzene, ETO, ionizing radiation	0.07%, 0.04%, 0.05%	2-4, 1.1-3.5, 1.3-2.1, 1.1-1.4
laryngeal	Sulfuric acid+mineral oils	4.4%, 6.4% (males)	1.1-5.0, 1.1-2.0
Non-melon. skin cancer	PAHs, arsenic	12%, 0.4% (males)	1.1-1.5, 2
Sinonasal+ Nasopharynx	Wood dust, nickel, chromium VI	6.8%, 6.0%, 5.2% (males)	3.1, 2.2, 5.2-10.8
Kidney cancer	Coke production	0.8%	2
Liver cancer	Vinyl chloride	0.5% (males)	2.5

*Arsenic, beryllium, cadmium, chromium, diesel fumes, nickel, silica

Cancer	Exposure	AF	Number deaths annually	Attributable deaths annually
Lung	Asbestos+ chemicals*	6-17% male, 2% females	91,000 male, 62,000 female	9,700-19,900
Lung	ETS at work, nonsmokers	1%	15,300 (10% all lung deaths)	870
Lung	Indoor radon at work	1%	153,000	2,000
Bladder	Aromatic amine dyes	7%-10% male, 0-19% female	7600 male, 3900 female	650-2,100
Mesothelioma	Asbestos	85-90% male, 23-90% female	2100 male, 500 females	1,900-2,400
Leukemia	Benzene, ETO, ionizing radiation	1-3%	19,000	200-500
laryngeal	Sulfuric acid+mineral oils	1-5%	3000	100-600
Non-melon. skin cancer	PAHs, arsenic	1-2%	1400	<100
Sinonasal+ Nasopharynx	Wood dust, nickel, chromium VI	30%-40%	300	200-300
Kidney cancer	Coke production	1-2%	7200	100
Liver cancer	Vinyl chloride	1%	7200	<100
All combined				12,000-26,000 (mid-point 19,000, 3.6%)

*Arsenic, beryllium, cadmium, chromium, diesel fumes, nickel, silica

Sources of underestimation

Exposures/cancers not included (other IARC class 2A carcinogens)

Contributory causes of death not included (cancer may not be the underlying cause- we considered only underlying cause)

Limited exposure data

- *unknown workplace carcinogens

- * did not include generic exposures (eg, painting/lung ca)

Sources of overestimation

Inadequate adjustment for confounding factors (some RRs overestimated)

Assumed independence of multiple exposures (could be synergistic)

RRs too high because based on historically high exposures (recent exposures lower)

Conclusion

Our best estimate of cancer deaths attributable to occupational carcinogens in 1997 was 19,000. Our best estimate of all deaths attributable to all occupational hazards was 55,200 (49,000 due to disease, 6,200 due to injury). Cancer accounted for 39% of occupational disease fatalities.

Leading US causes of death 1997	Number deaths
<i>Acute myocardial infarction</i>	206,212
<i>lung cancer</i>	151,464
<i>Coronary atherosclerosis</i>	129,649
<i>COPD, not specified</i>	81,893
<i>Pneumonia)</i>	75,815
<i>All deaths due to occupational exposure</i>	55,200
<i>Diabetes</i>	46,328
<i>Colon cancer</i>	45,462
<i>Congestive heart failure</i>	45,419
<i>Breast cancer</i>	41,940
<i>Motor vehicle crash</i>	37,324
<i>Prostate cancer</i>	32,891
<i>Suicide</i>	30,505
<i>Pancreas cancer</i>	27,193
<i>Alzheimer's disease</i>	22,475
<i>lymphomas</i>	22,354
<i>stroke</i>	20,998
<i>homicide</i>	19,491
<i>Occupational cancers</i>	19,000
<i>Emphysema</i>	17,518
<i>Renal failure, unspecified)</i>	13,733

Selected References

Driscoll et al. , The global burden of disease due to occupational carcinogens, *Am J Ind Med* 2005; 48: 419-431

Fritschi and Driscoll, Cancer due to occupation in Australia, *Aust N Z J Public Health* 2006;30:213-9

Nurminen and Karjalainen, Epidemiologic estimate of the proportion of fatalities related to occupational factors in Finland, *Scan J Wk Env Hlth* 2001; 27: 161-213

Rushton et al, Occupation and Cancer in Britain, *Brit J Can* 2010; 102: 1428-1437

Steenland et al. , Dying for Work: the magnitude of US mortality for selected causes of death associated with occupation, *Am J Ind Med* 2003; 43: 461-482

Steenland and Armstrong, An overview of methods for calculating the burden of disease due to specific risk factors, *Epidemiol* 2006; 17: 512-519