Evaluating Causation for Occupational Cancer Among Firefighters:

Report to the Workers' Compensation Board of Manitoba

Tee L. Guidotti, MD, MPH, FRCPC, CCBOM, FFOM
Division of Occupational Medicine and Toxicology
Department of Medicine
School of Medicine and Health Sciences
The George Washington University Medical Center
2100 M St., NW, Ste. 203
Washington DC 20052
Tel.: 202 994-1765
Fax: 202 994-0011
eohtlg@gwumc.edu

Version 1, 2 January 2005
Version 2 19 January 2005
Version 2.1 (final) 7 March 2005

Table of Contents

Executive Summary	3
Introduction	5
History	5
General Methodological Issues	7
Criteria of Causality and Causation	10
The Evidentiary Base for Firefighters	20
The Conventional Case: Genitourinary Cancers	24
Cases in which One Pathological Entity Predominates: Brain	27
Cases in which No One Pathological Entity Predominates: Leukemia, Lymphoma, Myeloma	28
Lung Cancer: Removing Smoking as a Confounder	32
Other Tumour Sites	40
Critique of the CCO Team Report	41
Conclusions	52
Bibliography	54

Executive Summary

The occupational health problems of firefighters have been extensively studied, to the point that the world epidemiological literature on this topic is among the most complete and detailed available for any occupation. Even so, many unresolved issues remain, especially whether firefighters are at increased risk for certain cancers. These issues are unlikely to be definitively resolved any time soon with new studies, because they primarily deal with rare outcomes and few studies are likely to have sufficient statistical power. Meta-analysis, while useful as an analytical tool, depends on the underlying data and cannot overcome the power limitations of individual studies if most of the studies involved in the analysis have already incurred a Type II error (missing a true association) or are subject to uncorrectable bias that obscures the association.

There is currently a movement across Canada, led by Manitoba, to adopt legislation establishing rebuttable presumptions for compensation of firefighters who develop certain types of cancer. Such presumptions must meet legal standards of the weight of evidence, in two ways. Assessing the occupational cancer risk of firefighters presents methodological problems common to the interpretation of epidemiological data for other rare outcomes. These problems are common in occupational epidemiology. We discuss criteria for inferring causation in such situations, both in general and by examining of the published policy of the Workers' Compensation Board of British Columbia. The WCB of BC has accepted most claims for cancer of the type under review and has developed a set of criteria that is unusually explicit and therefore worthy of examination as a model.

Epidemiological studies do not distinguish among primary cancers of the brain, leukemias and the lymphomas, because they are individually rare and subject to miscoding and aggregated coding. Environmental risk factors do not necessarily apply to all disease entities in the aggregation. A true excess may be diluted by inclusion in a category of cancers that includes other types not associated with the risk fact, leading to an inherent bias to underestimate the risk. Risk for leukemias is especially difficult to evaluate because studies often aggregate not only types of leukemia but also lymphomas and myeloma.

The presumptions for bladder cancer and kidney cancer are grounded on substantial evidence that already meet scientific standards of certainty. The presumption for testicular cancer

is based on an emerging literature suggesting a high risk but illustrate the "first case" problem: the first case of a previously unrecognized work-related condition to come to adjudication is likely to be denied because the literature does not exist to support it. The presumptions for brain cancer, non-Hodgkin lymphomas and leukemias are based on the inference that within the overall category there are specific disorders for which the evidence suggests an elevated risk but it is not possible to discern which are in excess. The argument for a presumption for nonsmoking firefighters with lung cancer may also apply in some cases to colon cancer.

A team from Cancer Care of Ontario conducted a study for the Workers' Compensation Board of Ontario, which is reviewed in this report, together with peer reviewer's comments and response. Multiple deficiencies are noted, the most significant of which is that the study, notwithstanding its stated intention and the title of the report, only looked at strength of association, did so selectively and rather arbitrarily, and did not in fact comprehensively evaluate the literature.

A brief conclusion reiterates that the problem of cancer in firefighters raises broader issues and represents a class of problems that should be approached by the logic of the problem and the application, not by stereotyped statistical algorithms when they may not apply.

Key words: firefighters, epidemiology, adjudication, study design, weight of evidence, bladder cancer, kidney cancer, testicular cancer, brain cancer, lymphoma, leukemia, myeloma, lung cancer, colon cancer, colorectal cancer.

Introduction

Firefighters may be at risk for a number of exposure-related cancers because of their occupational exposure to a variety of toxic agents.[1] A major unproven hypothesis is that risk increased following the introduction in the 1950's of combustible plastic furnishing and building materials known to generate toxic combustion products which may be carcinogenic.[2] However, the evidence for excess of certain cancers has been equivocal. This has led to great controversy and inconsistency in the adjudication of claims for occupational disease.

Although recent findings have led to an appreciation of the risk for one new outcome, testicular cancer, as will be discussed below, most of these issues are unlikely to be resolved by more data.[3] Most large studies on firefighters are similar in design (and therefore share similar biases arising form design), have similar problems in attaining complete and accurate ascertainment of deaths (the Canadian Mortality Data Base, however, conferring advantages in this country) and face similar limitations on power for rare outcomes. Virtually all inherent biases that predictably affect cohort studies of firefighters tend to underestimate risk and obscure associations through dilution and misclassification. Power is also limited for rare outcomes, even in studies with large subject populations. It does not necessarily follow that this problem can be overcome by meta-analysis, since meta-analysis merely aggregates and pools the risk, on a weighted basis, of studies that may have missed a true association in the first place because of low power, although this cannot be known for certain.

History

The present report was requested by the Workers' Compensation Board of Manitoba in October 2003, for delivery in January 2005.

In 2002, a movement arose in Canada to re-examine compensation for firefighters who develop certain cancers suspected to be work-related. The movement began with the provincial government of Manitoba. In 2002 we were asked to prepare a report on the health risks to firefighters for the Government of Manitoba. By April 2003, bills had been passed by the provincial legislatures in Alberta, Saskatchewan, Manitoba and Nova Scotia. Ontario had already adopted a set of similar provisions as policy in 1994.

The province of British Columbia has been a laboratory for assessment of these issues due to that province's explicit procedures for evaluating occupational disease cases and the activism of the provincial union for firefighters, which has commissioned a number of reports and analyses. In 2004, the Workers' Compensation Board of British Columbia commissioned a study by a team from the Division of Preventive Oncology of Cancer Care Ontario, a government-sponsored cancer treatment and control agency. In September 2004, we delivered a detailed report to the British Columbia Fire Fighters Association on presumption for occupational disease among firefighters. Much of this report, especially the analysis of the epidemiological evidence, closely follows the source document used as the basis for that report, although there are many changes.

The rationale for establishing a presumption for firefighters, and thereby expediting their claims, is a policy issue, not a scientific issue. It is based on the idea that firefighters, like police and a few other public safety occupations, are expected to take risks that would be unacceptable in any other work environment. They may be trained to manage these risks and to protect themselves, but the working environment cannot be made safe because they deal with situations that are inherently dangerous and may lose control. In the interests of society and as safety professionals, however, they essentially waive the right to refuse dangerous work and routinely accept the risk, like a soldier sent into battle to defend the country. It is, by this logic, ultimately in society's interest to compensate for this risk because the work has to be done. Because this formulation is based on values rather than differential risk, it is not subject to scientific debate. However, once the decision is made on policy grounds to expedite compensation for firefighters, establishing a presumption is an administratively convenient way to achieve the goal.

Systems that follow this approach for claimants deemed to be serving the national or community interest include the US Department of Veterans' Affairs Vietnam Veterans Agent Orange compensation program and the US Department of Energy's compensation program for workers exposed to chemicals in the nuclear weapons industry during the Cold War. They do not require balance of probability, but evidence of shared risk.[6] The burden of proof is to demonstrate sufficient expose to have incurred a substantial risk of the outcome in question, even if the probability falls short of 50% or even odds. These programs are intentionally biased in favour of the applicant for reasons of public policy and to alleviate the burden of illness after public service. Both programs require the totality of evidence to be taken into account, not just cohort mortality or cancer incidence studies, and freely admit toxicological evidence.

General Methodological Issues

Meta-analysis, while useful as an analytical tool, depends on the underlying data and cannot overcome the power limitations of individual studies if most of the studies involved in the analysis have already incurred a Type II error (missing a true association) or are subject to uncorrectable biases that obscure the association. Meta-analysis has been useful but has not successfully identified some cancers for which later cohort studies provided strong evidence for a probable increased risk, such as kidney and bladder.[4] Pooled studies with large populations have not fully resolved these issues, either.[5]

The role of meta-analysis in this application is in part a question of the philosophy of study design in epidemiology.[6] Meta-analysis takes a prospective approach to a post-hoc problem. The event (the disease) has already occurred. Meta-analysis is a retrospective look using a prospective statistical model.

The issue is whether it was more likely to be the result of a past exposure or condition existing in the past, not a future likelihood. For such problems, a Bayesian meta-analysis could be valuable but the approach is technically complex and subject to manipulation because the assumptions on prior probabilities may be somewhat arbitrary.

If a study looked for a real association and missed it (because of power considerations or bias), no amount of further analysis or lumping together with other populations is going to make up for the miss. The more the patchwork of studies is both very large (relative to the universe of firefighters) and very representative of all firefighters, the more likely meta-analysis may replicate something like an original idealized cohort of firefighters but it is like sewing a quilt from patches with holes in them, a quilt that does not adequately cover the bed. Small wonder that meta-analysis has to date not demonstrated previously unsuspected associations for firefighters, failed to identify the adverse effects of beta-blockers post-MI (despite what looked like firm evidence), and generally tends to miss as much as it reveals. The real information inherent in the differences among studies tends to get smoothed over in meta-analysis, although the findings are largely revealed in those very differences.

Meta-analysis is hypothesis-generating but not valid for hypothesis testing. It is, in this investigator's opinion, a blunt instrument masquerading as a sharp scalpel, because the

statistics look good. Meta-analysis has a role when there is no other source of information but it is not a substitute for digging deeply into the individual studies to figure out what is going on.

Complicating the issue of power and bias inherent in design is the common problem of misclassification and case definition. The current situation in assessing occupational cancer among firefighters is not unlike the following non-occupational analogy.

Imagine being asked to assess causation in a woman who took an experimental drug 20 years before and has developed cancer. However, imagine further that epidemiology is at a rudimentary state of development. The only epidemiological studies available indicate that the frequency of "female cancer" in Canada is approximately 74.5 per 100,000 per year. First, one must consider the denominator: 100,000 Canadians or 100,000 Canadian women? (Yet even so, occasionally males get breast cancer.) Secondly, what would be the definition of "female cancer"? Would it be logical to mix breast, ovarian, uterine corpus and uterine cervix into one category? Not hardly, but suppose that a medically-defensible breakdown was unavailable. Thirdly, one would then want to know at least which cancer types were more common: it would be noted that the more common types were, in order, breast (the actual rate is 107/100,000), uterus (19), ovary (15) and cervix (8). Suppose the woman in question had a cervical carcinoma. Would the knowledge that, for women taking this drug, the risk of "female cancers" was increased by one half help the consultant in determining whether this excess risk was shared by cervical carcinoma? One would correctly conclude that breast cancer was driving risk in Canada but one cannot know if breast or another type drove the excess. It could as easily be driven entirely by a small elevation in breast or a very large elevation in uterine or ovarian carcinoma. However, a doubling of the risk for cervical carcinoma would probably go unnoticed, a mere blip against the much greater background risk (A doubling of the rate of cervical cancer would increase the relative risk to only 105, statistically insignificant in most studies). If the rate of uterine cancer were falling, while the rate of cervical cancer increased, the overall rate of "female cancer" might still fall overall. Thus, a markedly increased risk in a subtype may be hidden by dilution and easily buried in opposing trends.

In this analogy, the confusion over the denominator is an imperfect analogue to smoking as a confounder, the aggregation of disparate cancers reflects the problem of tissue type when cancer is reported by site or general type and the problem of dilution and opposing trends is a close analogy, especially when cancers have different risk factors.

A third problem, which is inherent in reviews such as the one that will be subject to critique below, is that multiple studies are not just replications; they have an epistemological structure. Some study types, primarily those early studies examining multiple outcomes, are best suited to be hypothesis generating. However, others, regardless of their design, may have been organized to test hypotheses. A good example, described below, is the case of renal cell carcinoma. In that case, one study showed an elevation, a second study confirmed it, and a third study, independent and designed in ignorance of the others, also confirmed it. For a rare outcome, this is impressive validation. Another example is testicular carcinoma, in which one observation was followed byanother, which, although the elevation was not statistically significant, observed the same effect. These examples are provided in detail in the appropriate sections below.

In this context, the decision by the province of Manitoba to evaluate its own experience is noteworthy. It is logical and to be commended that the province would follow adoption of a new policy with a systematic effort to validate the assumptions of the policy. The present study of firefighters in Manitoba will be a useful addition to the literature. That study will be definitive in describing the Manitoba experience during the period it covers, i.e. it will be a (presumably, as we predict with confidence) accurate description of what actually transpired and the historic experience. It will not, however, be definitive as a study of all firefighters. Although a province-wide study will be relatively large for this literature, it will not overcome issues of power for the detection of increases in rare events (and may not even observe outcomes that are rare in the total population of Manitoba) even if there is a true increase in risk for the universe of firefighters. Especially for rare outcomes, it may not predict whether a meritorious case will appear in the near future or will take a much longer period to be statistically likely. It may uncover "clusters" that have other explanations and may not represent a true increase in cancer risk. All of these interpretive issues will be attached to the forthcoming Manitoba study, as they are to any study of this type. These issues can be partially overcome by larger populations and they can be clarified by more advanced statistics. Still, not all issues can be resolved in this way.

We suggest that issues of rare outcomes among firefighters represent a class of problem in occupational epidemiology that is best approached outcome by outcome, using principles of logic, rather than advanced statistical techniques or redundant studies using similar methodoogy.[6] Key to interpretation of these studies and the totality of the literature is the weight of evidence. This, rather

than scientific certainty, determines the outcome of the case or claim in a legal setting, such as tort litigation and adjudication for compensation benefits.

Criteria of Causality and Causation

Scientific standards of certainty do not apply to the assessment of the individual case in workers' compensation, because it imposes an unachievable burden of proof. The legal requirement is a determination on the basis of the weight of evidence, giving the benefit of the doubt to the claimant when the weight of evidence is balanced. (This provision is written into the Workers' Compensation Act in every province.) First, the association with work must be supported by the literature on the balance of probabilities ("more likely than not", or >50% certainty), but not necessarily to a level of scientific certainty (conventionally assumed to be >95% certainty, because the standard for statistical significance is $p < 0.05$). Second, it must be more likely than not in the unselected individual case, coming to adjudication, that the condition arose out of work, which corresponds to an approximate doubling of the relative risk. Factors specific to the individual case are grounds to rebuttal, or challenge to the presumption.[6]

The assessment of causation and the adoption of a presumption rest on the criteria employed for evaluating the scientific evidence for group risk and demonstrating that the general conclusions fit the particular case by individualizing analysis to the features of the case. The WCB of British Columbia has made these criteria very explicit ("Criteria of Causality", outlined in the WCB Reporter, pp. 431 – 467). In this section, the BC criteria will be taken as a model and explored for their implications in assessing causation. References are available in our recent book on this topic.[6]

British Columbia has generally accepted claims for cancer submitted by firefighters under discussion in this report (Table 1, data drawn from a Discussion Paper of the WCB of BC, dated 19 November 2004), except for colon cancer and non-Hodgkin lymphoma; there has been only one case of testicular cancer in the last twenty years, which was not accepted. (One claim for lung cancer was denied but this report does not recommend acceptance in the case of smokers.) In British Columbia, there is no statutory provision or special regulation governing the treatment of claims submitted by firefighters under the Workers' Compensation Act. Cases of occupational disease among firefighters must therefore be managed under the general criteria that the condition

is a recognized occupational disease, arising out of work and disabling such that it prevents the claimant from earning full compensation for work. Occupational diseases may be recognized in four ways: recognition as a rebuttable presumption (in Schedule B of the Act), recognition under section 6(4.2) of the Act as a disorder arising peculiarly from a particular activity specific to an occupation (there is only one example, not relevant to firefighters), recognition by regulation, and recognition "by order" in a specific case, which is not a binding precedent. Currently, adjudication in BC is based on the latter approach and rests on the application of general procedures embodied in the Criteria for Causality.

Table 1. Acceptance of Claims for Cancer by Firefighters, WCB of British Columbia, 1985 through October 2004.

Cancer Type	**Accepted**	**Denied**
Brain*	6	1
Bladder	1	0
Kidney	4	1
Lung	0	1
Non-Hodgkin Lymphoma	0	5
Multiple Myeloma	3	0
Testicular	0	1

* One case was suspended.

Causation

Causality is properly the study of cause and effect. Causation implies assessment of the cause of a particular outcome, such as an occupational disease. Causation is the more common usage in North America but the WCB of BC has used "causality" correctly in the policy.)

The BC Criteria as outlined in their policy handbook takes a rather mechanistic, unitary cause → single effect view of causation. The idea of causation is actually highly nuanced and rather complicated, however. Causation refers, in this context, to the risk factors or exposures that initiate the process leading to the health outcome. The concept is akin to that of etiology in clinical

medicine but without the implication that there must only be a single cause. The concept of causation in epidemiology assumes that the risk factors bear a "causal" relationship in that they either establish necessary condition or set into motion a mechanism that results in the outcome. There is no presumption, as there is in common language and in the work of some epistemologists, that a cause must be "sufficient" in itself to produce an effect.

A useful distinction can be made between causes that are risk factors in an epidemiologic sense, in that they increase the probability of an outcome that is not certain, and those few that are invisible precisely because they are intrinsic components of the mechanism that produces the effect, such as oncogenes or pathways of the metabolism of procarcinogens. The latter "component causes" (as they are called by Kenneth Rothman, who first articulated the concept in the context of epidemiology) are more profitably considered to be a means to the end rather than initiating events. Exposure to these component causes cannot be controlled because they are intrinsic; they may, however, be modified in such a way as to slow or prevent the action of the mechanism in producing the outcome. The distinction has obvious relevance to hereditary risk and family history, especially subjects with described syndromes of familial cancer risk.

A cause, in its sense as used here, is a factor that contributes to the likelihood that an outcome will occur. This is a stochastic, or probabilistic definition, not strictly a mechanistic definition. There is a certain probability, or odds, that a step will occur but no certainty. In daily life, one speaks of "cause and effect" relationships as if there is one cause for every effect and as if an effect necessarily follows the presence of a cause. This is too rigid to be useful in epidemiology and in cancer toxicology, where the mechanisms are complicated and influenced by numerous external and internal factors. It is not even useful, in this context, to speak of a cause as being either necessary or sufficient because causes may be interchangeable in the mechanism or may interact.

For example, exposure to either cigarette smoking or asbestos individually is known to result in lung cancer in a roughly predictable probability. Exposure to both vastly increases the risk beyond that of the summed probabilities of either alone, suggesting a substantial interaction. However, most workers who have been exposed to either or both do not develop lung cancer, although they might if they lived long enough and were free of other risks to their life. A few unlucky people who neither smoke nor are exposed to environmental carcinogens such as asbestos develop lung cancer regardless, although this is uncommon. Neither asbestos exposure nor

cigarette smoking is necessary, sufficient, or predictable in individual cases as a cause of lung cancer, but the association is clear and these factors are truly "causes".

This example also illustrates the fallacy in trying to apportion the contribution in individual cases of multiple causes. Is the interaction in the case of asbestos and cigarette smoking one of asbestos enhancing the effect of cigarette smoking or vice versa? In several exercises, authoritative investigators have attempted to estimate the proportion of cancer "caused" by various classes of external influences and have almost invariably concluded that smoking and diet are major causes of cancer in the population and that occupational and environmental exposures contribute much less. It may well be true that control of smoking is the single most effective approach to the reduction of cancer incidence now available. However, the apportionment of the relative contribution of causes to cancer incidence overall or to a single case assumes that their effects on the underlying mechanism are separable and individually discreet when they clearly are not. It is a useless exercise to apportion causation on the basis of risk estimates for factors in (relative) isolation because the mechanism is intrinsically always interactive. As Rothman (1986) has put it "...it is easy to show that 100 percent of any disease is environmentally caused, and 100 percent is inherited as well. Any other view is based on a naive understanding of causation."

Causation is an elusive concept applied to epidemiology. Human populations are vastly more complex than experimental systems, subject to numerous influences on health, behaviour, and social adjustment. each individual in the population may be subject to numerous other exposures that influence the outcome of interest. Although these are called "confounding" factors, they are often every bit as important in determining the outcome as the risk factor under study; it is a mistake to dismiss them as merely sources of bias. In the presence of numerous "confounding" factors, clear associations demonstrated in epidemiological studies are remarkable observations, with three likely explanations: an effect, a statistical chance event, or a bias in the study method. If the latter two "false" outcomes are excluded, the effect that remains is not necessarily causal in nature. (This insight was articulated in the original article by Bradford Hill, 1965.)

Multifactoral Outcomes

Cancer is the result of a multistep process, reflecting numerous influences and risk factors. Our task in assessing causation in adjudication is to identify whether the risk factors associated

with work increased the risk to the individual by as much or more as the individual would have experienced in daily life as a member of the general population (given the genetic inheritance they have). This issue is always looked at as a problem of discerning the signal from the noise but it is perhaps better conceived as an issue of identifying the feather that tipped the balance or the bet in roulette placed on a colour or a number.

When we study health outcomes from environmental exposures, we are several steps removed from the direct toxic lesions induced by these exposures. They are multifactoral health outcomes. There is no simple one-to-one relationship between a unitary cause and mechanistic, deterministic progression to effect, as there is in, say, physics or biochemistry (cell biology, of course, is another level of integration more complicated than biochemistry). There is often not even a relatively well-defined relationship among agent, host defenses, and environment, as there is for most communicable diseases. Rather, the problems we study are more like cardiovascular disease, in which numerous partial determinants ("agents" and "host factors") converge in a complex structure. This is most obvious for studies of environmental determinants of cancer, but holds true as well for other occupational and environmental diseases.

Biology and the health sciences are less predictive than physics because the object of study is more complicated and more highly variable from individual to individual in important respects. The study of behaviour is notoriously more complicated still because it involves many internal and external factors working interactively. The study of human populations is difficult because there are a multiplicity of external as well as internal determinants affecting each outcome or response.

To make progress in a field as complex as environmental health sciences, we construct hypotheses regarding the relationships between various x and y and look for empirical measures of the strength and consistency of the association between x_0 and y. We also argue endlessly over the relationships between measured variables and determinants and outcomes. Some critics virtually imply that failure to control variables is a character flaw of investigators and indicate inferior science (most infamously Alvin Feinstein, in 1988). This ridiculous. A critique of any study in the environmental health sciences should be based on an acceptance that this is the structure of the discipline, not a failure on the part of investigators to control all variables. The issue is whether these issues have been addressed and how carefully the methods have been applied.

A major limitation on the design and interpretation of epidemiologic studies of toxic exposures has been the difficulty in accommodating the implicit biological assumptions that form the basis of population studies. Instead, we have adopted a simplistic working hypothesis that concentrates on relating clinical outcomes to exposure to the hazard, without rigorously examining the susceptibility of the subject exposed and other important determinants of the response.

The Hill Criteria

As reflected in the WCB Reporter (pp. 444 – 448), the most widely accepted set of criteria for assessing the likelihood of causation reflected in epidemiological data is that proposed by the late Sir Austin Bradford Hill (1965). The criteria are to be applied rigorously and as a group; the more that appear to be satisfied, the more likely it is presumed that the association observed is truly causal. It is important to understand that the criteria can only be applied to the literature, not the individual case, and that they are predicated on a sufficiently well-developed scientific literature to answer each criterion. This is actually rarely available for any but a few very heavily studied hazards and occupations.

These criteria are as follows:

1. Strength of the association
2. Consistency among studies, esp. by different techniques
3. Specificity of outcome
4. Exposure precedes disease outcome
5. Dose-response relationship (epidemiologic)
6. Plausibility of a biological mechanism
7. Coherence of chain of evidence
8. Experimental association, esp. dose-response (toxicologic)
9. Analogy to similar effect produced by a similar agent

The Hill criteria for accepting an association as causal are not absolute. Some criteria are stronger than others. Many authors have commented on the Hill criteria pointing out their

limitations, not least Bradford Hill himself in the original paper. The criteria need to be understood for their limitations as well as their strengths:

1) The strength of an association is a strong criterion; SMRs less than 150, for example, are usually considered unlikely to be strongly associated with a single work-related exposure. A strong association may be a statistically uncertain one, with a wide confidence interval around the estimate, particularly if the number of cases is small. This is a common problem with rare disorders, such as the purported association between phenoxyherbicide exposure and non-Hodgkin lymphoma or soft-tissue sarcomas.

2) The criterion for consistency presents a paradox. Contradictory results from a study using different methods cannot directly refute the original observation because the circumstances of the investigation may not be easily comparable. Discrepancies among studies often provide greater insights than consistencies. However, results from an identical study (possible only in theory) conducted on a different population sample contributes no new insights and serves only as a replication for purposes of statistical inference testing. Few investigators spend much time precisely duplicating the experiments of others. A related issue is that of generaliseability: the results of a specific study may accurately describe the experience of that population but the findings may not be valid as a generalization to other populations.

3) Specificity is the weakest criterion of all. It is now well established that one exposure (e.g. asbestos) may lead to a number of outcomes (asbestosis, bronchogenic carcinoma, mesothelioma, and probably carcinomas of the larynx, colon and ovary). Specificity is an elusive criterion and depends importantly on exposure levels and host characteristics.

4) The temporal relationship (cause must precede effect) is an essential criterion. It is problematical only in so far as it is often difficult to sort out the time frame, particularly when there is a long latency until the effect is observed.

5) The biological gradient of exposure and response is very useful and compelling when it appears. The absence of an exposure-response relationship does not necessarily rule out an association, however. In many toxicological systems, response changes with exposure level and greater levels of toxicity may obscure the expression of more subtle effects, usually by increased lethality. In epidemiological studies, this is less often a plausible explanation for failure to observe an exposure-response relationship.

6) Biological plausibility is often questionable as a criterion. Many strong associations prove elusive in the laboratory, such as the association between arsenic exposure and lung cancer. As well, science remains in doubt about many of the host defenses and adaptive mechanisms that alter outcome.

7) Coherence of evidence is a strong criterion but it assumes a relatively thorough knowledge of the problem. Hence, this criterion (and those for consistency and biological plausibility) remove from consideration virtually all suspected associations in the early stages of investigation, regardless of how compelling the evidence collected initially. These criteria make the Hill guidelines a "scientific" rather than a "public interest" test, driven by the $\alpha = 0.05$ "rule" rather than a threshold of suspicion justifying action to protect the public. Control of a new or newly recognized hazard often cannot wait until the scientific evidence is complete, however.

8) Experimental or collateral validation may involve toxicological demonstration of a similar or comparable effect but may also extend to experimental epidemiology (involving a controlled intervention among human populations) or, less convincingly, quasiexperimental studies (evaluations following interventions without strict control). In effect, this criterion is an attempt to sidestep what is usually seen as the fundamental limitation of epidemiology as a science, that it is inferential (demonstrating associations that suggest hypotheses of causation) as opposed to mechanistic (demonstrating the certain mechanisms by which an effect occurs). To some degree, this represents a misunderstanding of the realities of experimental science; experimental studies are themselves inferential but usually with a greater degree of face validity. Even "hard sciences" such as chemistry and physics are inferential but with a much greater degree of inherent confidence in individual studies and reproducibility because the systems under study are far simpler, even if the measurement technology is usually more elaborate. (Particle physics can be described mathematically; human behaviour cannot.)

9) Reasoning by analogy is one of the weaker criteria. It is actually closely akin to the criterion for coherence of evidence. Analogy is very useful in generating hypotheses and theoretical constructs but is invalid as a means of proof. Certainly the analogy between causes and mechanisms of pneumoconioses would not have predicted the carcinogenicity of asbestos. For proof of causation, empirical evidence is required; analogy is only a step along the way.

Confounding

Confounding is a major issue in epidemiology and is only cursorily and conventionally addressed as in the Criteria for Causality.

A fundamental concept of epidemiology is that main effects will demonstrate consistency in association while other complicating associations will be variable in strength. Two corollary problems are that a relatively weak association may be lost among stronger confounding associations and that bias may mimic this consistency in association. Susser (1991) proposed a set of definitions for discussing these aspects of causation.

Confounding factors are not the same as background noise, or the numerous competing influences that affect opportunity to be exposed to hazard or the health outcome that a person may experience. These competing influences are distributed more or less randomly and are adequately dealt with by conventional inferential statistics. The problem that then proves most vexing is accounting for confounding factors, those characteristics that are associated with both opportunity for exposure and outcome. For example, acceptance of cigarette smoking has become increasingly a characteristic linked to social class in recent years. Smoking rates are much higher among blue-collar workers than white collar. Smoking is also closely associated with risk of lung cancer. Thus, a study of risk factors for lung cancer among foundry workers, who are a blue-collar group, must account for the contribution of smoking before concluding that any workplace exposure plays a role. However, smoking may interact with workplace exposures, magnifying the cancer risk further. Certain subjects who might have died eventually from smoking-related lung cancer may have died earlier from workplace exposure-related lung cancer, and vice versa. It is therefore just as great a mistake to dismiss associations summarily, on the assumption that they are due to smoking, as it is to assume that smoking plays no role.

Further complicating the issue is that exposure and outcome are rarely, if ever, studied directly. Exposure is estimated from representative measurements, at best, and categorized, at worst, by the use of "surrogates", or proxy categories, such as occupational title or years of employment. Outcome is not always easily determined in individual cases or accurately enumerated because of limitations on clinical testing and diagnosis. The best that can be done is to examine indices that reflect the magnitude of the exposure, and indices that reflect the magnitude of the effect.

Additional risk factors can be associated with the effect but only a few bear some relationship to the exposure of interest and to the outcome. These are the confounding factors. Each of these must also be estimated by an index that variably reflects the magnitude of each. The result is that epidemiology measures "reflected" causes and effects, not the direct chain of events. While this is also true of the experimental sciences, even physics, the relationship between the actual event and the indices that represent it are more distant in epidemiology than in other disciplines. This is a fundamental, irreducible problem inherent in the science, just as the scale of observation poses practical limitations on particle physics and the lack of a meaningful experimental model places limitations on astronomy.

Validity

The Criteria do not really speak to the all-important issue of validity. My personal opinion is that a positive test is much more significant than a negative test because every conceivable epidemiological study in the real world is subject to error, introduced by:

1. Exposure assessment: identity, level, and duration
2. Outcome data
3. Latency period
4. Statistical power (the probability of finding an effect when it truly exists)
5. Access to populations of interest
6. Confounding exposures
7. Bias

Exposure assessment is a common problem in epidemiologic studies because accurate data on exposure levels in the workplace are usually lacking, especially for individual subjects. Outcome data, for example, in non-Hodkin lymphoma (NHL), is frequently confused, misclassified or aggregated. This makes it difficult to establish refined exposure-response relationships and forces the investigator to make a number of assumptions. The long latency period of many diseases, particularly many cancers, makes identification of an association difficult and greatly complicates demonstration of an exposure-response relationship. Epidemiologic

studies of many occupational disorders are also difficult to perform because of small numbers of subjects often result in a low statistical power. It is often difficult to do studies because employers refuse access so the worker groups we study are often not representative of their class. Especially for cancer, but not so much for NHL, there are numerous confounding exposures, discussed above.

The misclassification biases introduced by these problems almost all, in almost every case, operate to reduce magnitude of an association and the probability of finding an effect. Given that study implementation in the real world is intrinsically biased against finding an effect, therefore, the actual demonstration of a positive finding is, all other things being equal, more persuasive than a negative finding.

When a well-designed study does show a positive effect, particularly if it has low power to begin with, this to me is far more compelling than similar studies that show no effect, particularly those with low power. This opinion is often challenged by other experts who believe that meta-analysis can by aggregation reproduce an artificial, representative global population from a working population that has already been fragmented by individual studies, many of which may well have missed the association being sought because fo power and bias. However, Sir Richard Doll and Julian Peto, the world's greatest living epidemiologists, took precisely the same position in their seminal book *The Causes of Cancer*.

The Evidentiary Base for Firefighters

The literature up to 1994 has been extensively reviewed in two authoritative sources. In 1994, the Industrial Disease Standards Panel of Ontario produced a widely-quoted report designed to identify candidate conditions that merit occupational disease presumptions in that provinces' system of workers' compensation.[7] We published a similar analysis in 1995, differing in detail but reaching similar conclusions, which are summarized below.[8] Table 2 provides a comparison of these two reports. (Much but not all of the text in this section and following sections on outcomes parallels the text of a report to the British Columbia Professional Fire Fighters Association. Both were adapted from the text of a source document which was also used for a manuscript in progress.)

- Lung cancer: There was evidence by 1995 for an association but not of sufficient magnitude for a general presumption of risk. We suggested that a presumption be considered for non-smoking firefighters.

- Cancers of the genitourinary tract, including kidney, ureter, and bladder: The evidence by 1995 was strong for both an association and for a general presumption of risk.
- Cancer of brain: Incomplete evidence by 1995 strongly suggested a possible association at a magnitude consistent with a general presumption of risk.
- Cancer of lymphatic and haematopoietic tissue: As a group, there was some evidence for both an association and a general presumption or risk. However, the aggregation is medically meaningless. We therefore recommended a case-by-case approach.
- Cancer of the colon and rectum: There was sufficient evidence to conclude that there is an association but not that there is a general presumption of risk.

Table 2. Conclusions of recent reports on firefighters and cancer risk.

Outcome	**Guidotti, 1995[8]**	**IDSP, 1994[7]**
Criteria:	*Totality of evidence, weight of evidence*	*Weight of evidence*
Lung	Association Presumption (for nonsmokers, rebuttable)	No association
Colorectal	Association	Association
Bladder	Presumption	Association
Kidney	Presumption	Association
Testes	N/A	N/A
Leukemia	Association?	Presumption
NHL	Association?	Presumption
Myeloma	Association?	Presumption?
Brain	Association	Presumption

The evidence supporting these recommendations is presented in these reports and will not, with the exception of leukaemia, be revisited here. Since 1995, there have been several additional major studies that have contributed to the world literature on firefighters. They are summarized below without reference to the earlier work.

Burnett et al.[5] conducted a very large proportionate mortality study on firefighters in 27 American states from 1984 through 1990, using data from the National Occupational Mortality Surveillance system. Limitations of these data are partially overcome by the sheer size of the database, 5744 deaths among white male firefighters, which is far beyond what may be achieved in

any one cohort study. This study is an example of population surveillance for occupational disease, which we have advocated elsewhere.[3] Burnett et al. was specifically designed to follow up on findings in the Alberta study. It should therefore be considered to he hypothesis testing, not a typical PMR study valid only for hypothesis generation. This issue is elaborated on in the Conclusion.

Deschamp et al.[9] studied the recent experience of relatively small number of fire fighters in Paris from 1977, as a prelude to a longer-term cohort study. An elevated SMR was found for respiratory cancers (1.12), gastrointestinal cancers (1.14) and genitourinary cancers (3.29) among other findings. However, the study is anomalous in several ways, uniquely demonstrating an elevated mortality from stroke (1.19) and a very low overall mortality (0.52), the lowest reported to date among firefighters. Further experience with this cohort is required to interpret the findings.

Ma et al.[10] conducted a large study using the same database as Burnett et al.[5] to explore race-specific disparities in cancer mortality. For this study, the NOMS database was extended by three years to 1993 but lost data from three states that were removed. Race as coded on the death certificates yielded 1817 deaths of white firefighters and 66 deaths of black firefighters. As expected, the overall results were similar. Of greater interest is the pattern of race-specific elevations. If an environmental or occupational factor is the major risk factor for a certain type of cancer, one would expect elevations in both white and black firefighters.

Bates et al.[11] reported a study on firefighters in New Zealand from 1977 to 1996, conducted to investigate the observation of a cluster of testicular cancer. This elevation was confirmed independent of the cluster. This study is unusual in reporting both cancer incidence and mortality. It also reports one of the lowest mortality ratios reported for firefighters (0.58), suggesting a uniquely strong healthy worker effect. The healthy worker effect is much stronger for cardiovascular disease than for cancer but a mortality level this low suggests that underlying risk factors such as smoking may be hugely different. The authors caution that matching to mortality data and cancer registration data may be incomplete prior to 1990 and suggest that they have greater confidence for findings after this date. Bates et al.[11] observed no significant elevations except for testicular cancer. They found a marked increase in testicular cancer and nonsignificant elevations in incidence in the 1977 – 1996 cohort of cancers of interest: lung (1.14, 95% 0.7 – 1.8), which showed a modest increase with duration of service, bladder (1.14, 95% CI 0.4 – 2.7), brain (1.27, 95% CI 0.4 – 3.0), and “myeloleukemia” (1.81, 95% 0.5 – 4.6), but not kidney (0.57, 95% CI 0.1 – 2.1). Limiting the analysis to the 1990 – 1996 subcohort, however, they found the increase in

testicular cancer but a deficit in most of these same cancers, except for brain (1.59, 95% CI 0.3 – 4.6), and no kidney or "myeloleukemia" cases. A strikingly different picture is observed in the pattern of deaths, however. Mortality among firefighters in the 1977 – 1996 cohort is elevated for bladder cancer (2.73, 95% CI 0.3 – 9.8) but less than expected for lung (0.86, 95% CI 0.4 – 1.6), brain (0.68, 95% CI 0.1 – 2.4) and "hematopoietic cancer" (0.72, 95% CI 0.2 – 1.8), and these use no deaths from testicular cancer. The discrepancy between incidence and morality in cancers with a high case mortality, such as lung, is an anomaly. However, all numbers are small and the authors are candid in describing limitations of the database.

Baris et al.[12] conducted an exemplary cohort mortality study. This study should be accorded great weight because it has exceptional power, spans most of the 20th century, and has the most complete follow-up. The study therefore merits description in detail.

The cohort consisted of 7789 Philadelphia firefighters employed from 1925 to 1986 compared to US white male rates, comprising 204,821 person years of follow-up. The men were hired in their late 20s (on average) and worked on average for approximately 18 years, with an average of 26 years follow up. Baris et al.[12] examined their cohort by age, duration of employment, job assignment and by number of runs to fight fires (enumeration of responses from the firehall) in three broad ordinal categories. There were 2220 deaths among the members of the cohort.

All causes of death and all cancers were approximately equal to the expected rates for all U.S. white males. The authors did observe statistically significant excesses for colon cancer (SMR 1.51; 95% CI =1.18-1.93). Nonsignificant excesses were reported for cancers of the buccal cavity and pharynx (1.36; 95% CI=0.97, 2.14); for non-Hodgkin lymphoma (1.41; 95% CI=0.91,2.19); for multiple myeloma (1.68; 95% CI=0.90-3.11) and for lung cancer (1.13; 95% CI 0.97-1.32). With >20 years of firefighting, the following cancer sites showed elevated risks: colon cancer (1.68; 95% CI 1.17-240); kidney cancer (2.20; 95% CI 1.18-4.08); non-Hodgkin lymphoma (1.72; 95% CI 0.90-3.31); multiple myeloma (2.31; 95% CI 1.04-5.16); and benign neoplasms (2.54; 95% CI 1.06-6.11).

Baris et al.[12] developed a direct index of exposure by assessing risk by volume of firefighting runs over the career of firefighters. Cancer of the pancreas showed a possible exposure-response relationship which rose from 1.02 for low to 1.17 for medium to 1.61 for high exposure. Although there were no other tumour sites with an exposure-response gradient,

comparing low exposure (1.00) to high exposure reveals that, several other cancer sites demonstrated increasing risk: stomach, 1.20; pancreas, 1.42; leukemia, 1.22; and benign neoplasms, 2.06. The authors also evaluated career exposure to diesel fumes, including unexposed. Although there were no continuous exposure-response gradients, several sites demonstrated increasing risks in the medium and high categories compared to unexposed: buccal cavity and pharynx, prostate, brain, multiple myeloma, and leukemia.

The Conventional Situation: Genitourinary Cancers

Genitourinary cancers represent the conventional situation, in which individual diseases are more or less satisfactorily classified and risk estimates probably do reflect the experience of the group for the individual cancers. Perhaps for this reason there seems to have been more widespread acceptance of claims by firefighters in this class than for other outcomes. The data are reasonably easy to interpret by tumor site. One source of uncertainty, however, is ureter cancer. Although linked with kidney cancer in most epidemiological studies, because of their rarity, cancer of the renal pelvis and ureter are diseases of transitional epithelium, like bladder cancer. In the absence of specific information, one may infer some characteristics of ureteral cancer by analogy to bladder cancer.

Bladder cancer

Burnett et al.[5] found no elevation for bladder cancer. The PMR was 101 for firefighters dying under the age of 65 and 99 for those dying at or over the age of 65. With 9 and 37 deaths, respectively, this is a large collection of deaths by bladder cancer. Using the same database, Ma et al.[10] reported that a not-quite statistically significant elevation of 1.2 was observed for bladder cancer among white firefighters and an elevation (but based on a single case) for black firefighters.

For bladder cancers, latencies tend to be shorter and more variable than for other solid tumours. Aniline dye workers in the 1940's and 1950's showed a latency as short as seven years, presumably due to high, constant exposure which may have compressed the latency period to its absolute minimum. This is not plausible for firefighters. The exposure of firefighters to potential

bladder carcinogens is much less than for chemical workers in the 1920's. In our data from Alberta[13] bladder cancer did not appear before age 60 or before 20 years of service and showed a very long peak latency of 40 years. Baris et al.[12] reported a slightly elevated SMR of 1.25 for bladder cancer, with greatest risk being among those hired before 1935 (SMR=1.71 95% CI=0.94,3.08), and among those with greater number of runs during their first 5 years employed (SMR=2.59, 95% CI=0.64,9.84). It would be difficult to accept a latency under 10 years for bladder cancer in a firefighter but the literature from other occupations does not rule out latencies under twenty years. One might expect that the duration of service associated with risk among firefighters to be on the order of 15 years.

Kidney cancer

Following up on the observation by Guidotti in Alberta[13], Burnett et al.[5] found a marked elevation for cancer of the kidney among firefighters in 24 US states. The PMR was 141 for firefighters dying under the age of 65 and 144 for those dying at or over the age of 65. With 24 and 53 deaths, respectively, this is a large collection of deaths by kidney cancer. Using the same database, Ma et al.[10] reported a borderline statistically significant elevation of 1.3 for cancer of the kidney among white firefighters. No cases were observed for black firefighters.

Delahunt and colleagues[21] in a study based on the New Zealand Cancer Registry, identified a threefold, statistically significant, excess of cancer of the kidney among firefighters (3.51, 2.09 – 5.92). The magnitude of the risk was consistent with that obtained by Guidotti.[13]

The standard cancer epidemiology text Schottenfeld and Fraumeni[20] cites several studies in which a near doubling of risk is associated with duration of employment less than ten years, among the aluminum workers exposed to polycyclic aromatic hydrocarbons. These are likely to be the responsible carcinogens in firefighting. In data from Alberta[13] a marked elevation in risk for kidney cancer was visible in the category 10 – 19 years of employment. Baris and co-workers[12] reported a doubling of risk with an SMR=2.20, 95% CI=1.18, 4.08 among those employed for 20 or more years.

It is not clear that kidney cancer follows the same pattern as bladder cancer and latency has not been as intensively studied for kidney cancer. On the basis of current understanding and the literature on firefighters, it might be difficult to accept a latency under 15 years, just on the

basis of the time required for a solid tumour to proliferate, but latency periods less than 20 but greater than 15 would not be unreasonable.

Cancer of the ureter, as noted, is a malignancy of transitional epithelial tissue, like that of the bladder. Cancer of the ureter might therefore be expected to feature a variable latency and a similar exposure duration but a tendency in most cases to appear late in life and after 20 years of employment as a firefighter. It would therefore be reasonable, on analogy to the conclusion for bladder cancer, to accept a latency period no shorter than ten years and a minimum duration of exposure of 15 years.

Testicular Cancer

Bates et al.[11] found a standardized incidence ratio of 3.0 (1.3 – 5.90) for testicular cancer among firefighters in the New Zealand city of Wellington. Stang et al.[22] reported similar findings from northern Germany, although their odds ratio of 4.3 (0.7 – 30.5), while large, was not statistically significant. Such high risks are unlikely to be confounded by differences in the prevalence of cryptorchism (the major known risk factor), smoking (not known to be associated with testicular carcinoma) or other plausible alternative risk factors. In their community-based study of testicular carcinoma, only four firefighters and three controls were firefighters out of 269 and 797, respectively, making the power of their study very limited. Stang et al.[21] also reported on duration of employment. Of the four cases, two had been employed as firefighters more than 20 years and two for less than 4.

There are five basic tissue types of testicular cancer, the most common by far being seminoma (about 95%). Bates et al.[13] does not specify the histology of the tumours. Stang et al.[22] reports that of the four in their study, two were embryomas, an unusually high frequency, which suggests, but does not prove, that this type (which is also found in mixed germ cell types) may be uniquely associated with occupational risk. The evidence is too weak to rely on, however without replication.

Given the totality of the evidence, it is reasonable to establish a presumption for testicular carcinoma on the basis of current evidence. However, given the methodological limitations of Bates et al.[11] and the lack of available evidence on exposure, tissue type of the tumours and latency, no further guidance can be recommended. Testicular cancer was not considered in

earlier studies and an excess may have been hidden in aggregate figures for genitourinary cancers.

Testicular cancer represents a good example of the "first case" problem. When the first case of a previously unrecognized association is asserted in a legal action or a claim is filed, the literature is undeveloped and therefore does not support it. The first case is almost always denied. If the case is decided in litigation, this closes off access to any future remedy because the case has already been decided. Unless there is a provision in a workers' compensation system that requires the system to reopen claims, the rejected claimant goes without compensation regardless of the subsequent accumulation of evidence. A publicized first case often stimulates further research but this usually comes too late for the initial claimant.

Cases in which One Pathological Entity Predominates: Brain

Cancers of the brain arising from brain tissue are relatively rare and may include twenty or more individual types. Each type may or may not be a different disease, with its own risk factors. Epidemiological studies do not distinguish among them because they are individually rare, subject to miscoding and are aggregated into a more general ICD code when they are reported. The most common type of "brain" cancer is glioma but this type only constitutes about half of the total. Meningiomas, which are not usually malignant, do not arise from brain tissue and are not obviously associated with environmental or occupational exposures, are often counted as brain cancer. Gliomas (astrocytomas) are more likely to be associated with environmental and occupational exposures. The risk of brain cancer as an aggregated category is increased in many studies but this risk is probably diluted by inclusion of cancers (and meningioma) that are not associated with environmental or occupational factors.[20] This leads to an inherent bias to underestimate the risk for that subset of cancers that may have a true association with firefighting. Analysis by specific tumour type might identify which, if any, is associated with the risk but these cancers are uncommon and such a study would be very difficult; require large populations and will not be done anytime soon if ever. A different approach is required, inferring risk for the predominant type from the combined risk for the group. Ma et al.[12] reported that no elevation was observed for brain cancer among white firefighters but a very large elevation, with a mortality odds ratio (MOR) of 6.9 (95% CI 3.0 –

16.0) was observed for black firefighters. Burnett[5] did not observe an elevation for cancer of the brain.

Baris[12] observed a relative deficit of brain cancer, with an SMR of 0.61 (95% CI 0.31-1.22). Risk did not appear to be concentrated in any subset of firefighters by assignment, number of runs or duration, although the highest SMR (1.18) was observed among firefighters with more than 729 runs in the first five years of duty. Because brain is an uncommon tumour site, statistical power is usually limited, even in large cohort studies. This study therefore does not contradict the findings of other studies that suggest an elevation in risk (upper 95% CI was 1.22), but it does not support them either. The weight of evidence to date, predominantly from earlier studies, suggests that the elevation in risk for brain cancer reflects a true risk which may be concentrated in certain subgroups, as demonstrated among black firefighters.

Demers et al.[14] documented a doubling or risk (SMR 257) at less than ten years of employment peaking at over a tripling (353) up to 19 years. Heyer et al.[15] also showed a near-doubling of risk (184) at less than 15 years duration of exposure. It is not clear what the minimum latency for a brain cancer might be, especially for rapidly-growing astrocytoma. It would be reasonable to assume that for aggressive brain cancers, exposure periods plus latencies may be under ten years in some cases.

Cases in which No One Pathological Entity Predominates: Leukemia, Lymphoma, Myeloma

This disease aggregation represents the most difficult case and remains refractory to efforts to tease out which individual diseases that are driving the elevated risk.

"Leukemia, Lymphoma, Myeloma" is a common aggregation in epidemiological studies. However, it is not a medically defensible aggregation of disease outcomes. The disease categories are distinct, although there is some overlap, and each category consists of individual disease with very different characteristics. Most epidemiological studies aggregate deaths or incident cases in these three broad categories, and even more commonly together, in order to achieve sufficient numbers for statistical analysis. However, the legitimate purpose for doing so must be to make a provisional assessment, to determine if there is an anomaly. When these

aggregations are taken at face value, as if they were a single disease outcome, elevations in one disease or a deficit in another can easily distort the aggregate risk estimate.

Non-Hodgkin Lymphomas

Lymphomas are uncommon. They tend to contribute a small number of deaths in most studies and are difficult for epidemiologists to analyze. Because they are individually rare and many tend to manifest themselves at older ages, their relationships to environmental factors are more difficult to determine.

Epidemiological studies generally do not separate the various types, or if they do, divide lymphomas into simply Hodgkin's disease and non-Hodgkin lymphomas. Hodgkin's disease is actually a class of apparently closely related lymphomas that tend to peak in young adulthood and again at older age and have not been associated with occupational or environmental exposures or occupational risks. Non-Hodgkin lymphomas are a larger, more heterogeneous category and have been known to be associated, for a long time, with many environmental exposures and occupations.[16] Non-Hodgkin lymphoma is further divided, especially in older epidemiological studies, into the obsolete catagories "lymphosarcomas" and "reticulum cell sarcomas".

This crude system obscures the level of risk that may exist for certain critical types of lymphoma. There are over 30 types of lymphoma recognized in the current classification system (the WHO and R.E.A.L. system). New types will be identified as immunological and genomic methods become more sophisticated. Different types of lymphoma are known to be associated with different occupational risk factors, including follicular cell lymphoma with the meatpacking industry and small cell lymphoma with solvent exposure.[17] Chronic lymphocytic leukemia, which is more accurately considered a lymphoma appearing in blood, was recently identified as a risk of Vietnam veterans exposed to herbicides on this basis, although leukemias in general are not so recognized.[18]

If, as seems plausible, different environmental exposures are associated with different cell types of non-Hodgkin lymphoma, a truly elevated risk that arises, for example, from exposure to some constituent of combustion products, may be diluted by inclusion with all the other types of lymphoma, that have no association with the exposure. Analysis by specific

tumour type might identify which, if any, is associated with the risk but these cancers are uncommon and such a study is probably not feasible for a single occupation.

Lymphatic cancers were separately addressed in Burnett et al.[5], which revealed an elevation for non-Hodgkin lymphoma. The PMR was 161 for firefighters dying under the age of 65 and 130 for those dying at or over the age of 65. With 35 and 66 deaths, respectively, this is a large collection of deaths by lymphoma. These cancers were also separately identified by Ma et al.[10] who found a statistically significant elevation of lymphatic cancer was observed among white firefighters, with a MOR of 1.4. Ma found no elevation was observed among black firefighters, based on a single case.

Baris et al.[12] observed a not-quite significant overall elevation for non-Hodgkin lymphoma, with an SMR of 1.41. While not achieving statistical significance, this rose to 1.72 for firefighters with 20 years or more experience and 2.65 for those assigned to ladder companies. The subset hired between 1935 and 1944 did show a statistically significant elevation of SMR 2.19 (95% CI 1.18-4.07). A reverse dose-response relationship was observed by number of runs, with the group experiencing the lowest number showing a significant elevation, with an SMR of 2.36 (95% CI 1.31-4.26), but no relationship was found with runs during the first five years. Baris et al.[12] found that among those employed more than 20 years, the SMR was 2.20, 95% CI=0.90,3.31). This suggests the possibility that these are true elevations in these subgroups.

Leukemias

Haematopoietic cancers (which affect the blood-forming organs, most particularly bone marrow) are generally known as leukemias. There are about a dozen well-recognized forms of leukemia, of which five or six predominate. Different environmental exposures may be associated with different cell types. Acute myelogenous leukemia is known to be associated with benzene exposure. AML is the most common leukemia in adults and this leukemia has been the subject of many studies. Individually, leukemias are relatively uncommon. A truly elevated risk of AML, which may arise from exposure to benzene in combustion gases, may well be diluted by inclusion with all the other types of leukemia, many of which may have no environmental

association. Unless studies are conducted on specific leukemias among firefighters, this problem cannot be resolved and the risk within the class must be inferred from the available data.

Haematopoietic cancers were separately addressed in Burnett et al.[5], who reported a PMR of 171 for firefighters dying under the age of 65 and 119 for those dying at or over the age of 65. With 33 and 61 deaths, respectively, this is a large collection of deaths by leukemia. Ma et al.[10] observed no apparent elevation for haematopoietic cancers, with an MOR of 1.1. There were no cases among black firefighters. This is unusual but probably reflects the smaller numbers of black firefighters in the American population.

Baris found no overall elevation for the leukemias (SMR 83, 95% CI 0.50-1.37), not specified as acute or chronic or by type. A statistically significant elevation in SMR of 275 (95% CI 1.03-7.33) was observed for firefighters assigned to ladder companies only, but not to those assigned to both ladder and engine companies. A non-significant elevation was observed for those with a high level of runs in the first five years, with an SMR of 2.44 (95% CI 0.70-8.54) and with medium (but not high) levels of runs over a lifetime, with SMR of 2.50 (95% CI 0.56-11.10). These data are not compelling evidence for a true association in this population but do not rule it out. Because of power considerations, the study by Baris et al.[12] does not really clarify this issue.

There is also an important anomaly in the older literature. L'Abbé and Tomlinson,[19] in a study of firefighters in Toronto, uniquely reported risk for types of leukemia. They observed an excess of "lymphatic" [lymphocytic] leukemia at 190 (42 – 485). This finding was highly influential in the IDSP report, but is anomalous. Acute myelogenous leukemia (AML) would be expected to be elevated in circumstances in which benzene is a hazard, not lymphocytic. These findings suggests that it is premature to limit the presumption to AML.

Although Ontario now recognizes lymphocyctic leukemia, the evidence presented by L'Abbé and Tomlinson[19] cannot be used to rule out the possibility of an association with AML. The evidence suggests (again, at the level of "more likely than not") that it cannot, be convincingly argued that only one form of acute leukemia, either myelogenous or lymphocytic, should be recognized. Lymphocytic leukemia is suggested by the empirical data, AML by the known toxicological profile of exposures experienced by firefighters. Thus, it is not possible to recommend a selective criterion that only recognizes AML, lymphocytic or, for that matter, only acute and not chronic leukemias.

Leukemias tend to have short latencies, on the order of five years or so. Short latencies and therefore duration of employment for leukemia are reasonable, on the order of four years to ensure that no errors of exclusion are likely.

Myeloma

Myelomas are B-cell lymphomas and malignant plasma cell dyscrasias. Baris et al.[12] found that increased with duration of employment, with 20+ years having a statistically significant SMR of 2.31, and a statistically significant SMR of 2.54 for engine company employment only, with some suggestion of correlation with medium and high diesel exposures (latter based on small numbers of deaths).

Interpretation

The weight of evidence for lymphatic cancer of the non-Hodgkin type and haematopoietic cancer suggests that the elevation in risk reflects a true risk in certain subgroups but these subgroups cannot be readily identified by usable criteria in adjudication. Thus, the earlier recommendations from IDSP[7] for a presumption, and by Guidotti,[8] for an implied presumption but with individual evaluation of each case, are not contradicted by the new evidence. Because the individual disease risks cannot be separated, they must be taken as a group until more information is available.

Lung Cancer: Removing Smoking as a Confounder

Lung cancer presents a different problem. In this case, the risk associated with occupation is overwhelmed by the effect of cigarette smoking. A different approach must be used.

Lung Cancer

Lung cancer has been among the most difficult cancer sites to evaluate. Despite the obvious exposure to carcinogens inhaled in smoke[2], it has been difficult to document an excess in mortality

from lung cancer of a magnitude and consistency compatible with occupational exposure. Respiratory protection has reduced individual exposure levels to combustion products since the 1970's, and this may be the reason that studies rich in recent person-years of observation, such as Baris et al[12] , do not observe elevations. Studies we conducted in Alberta on fire fighters entering the fire service from 1927 to 1987 do show evidence for an increase in risk .[13]

Without question, cigarette smoking complicates the analysis, but the prevalence of smoking among fire fighters does not appear to be excessive compared to other "blue collar" occupations. A comparison that takes into account the prevalence of cigarette smoking is illuminating.

Many studies have shown an excess of lung cancer on the order of 20 to 80% (i.e. SMRs around 120 or 180), a magnitude not uncommon in studies of other blue collar occupations with less plausible exposure levels.[24] However, the empirical findings on lung cancer from recent, well-designed epidemiological studies have been inconsistent.[8] One study from Denmark reported a standardized mortality ratio of 317 for older fire fighters but the comparison population was unusual and difficult to interpret[25]. Studies on cohorts from San Francisco [26] and Buffalo[27] showed no excess and even suggest a deficit. This might be expected if firefighters, on average, smoke less than the general population and there is some evidence for this.[28]

In 1995, we proposed that the true risk for lung cancer associated with fire fighting was probably on the order of 150.[8] This figure has been disputed. We suggested then that the true risk has been underestimated in career fire fighters and both diluted and confounded by the effect of cigarette smoking, which is a much greater risk factor.

Virtually all extant studies that are positive, relevant, close to the primary data, large and well done seem to cluster in a band from an excess of 30% to 68%.[8] The principal exceptions are Baris et al. [12] and Vena and Fiedler [27] Baris et al.[12] , despite a low overall risk (1.13, 95% 0.97 – 1.32) does report suggestive elevations in certain subgroups, notably fire fighters with less than 9 years of service (1.52, 95% CI 1.16 – 2.01), those assigned to engine companies (1.18, 95% CI 0.93 – 1.51), and those hired before 1935 (1.30, 95% CI 0.97 – 1.73). Vena and Fiedler [27] present one of the lower overall risks in the fire fighting literature (0.94, 95% CI 0.62 – 1.36) but their data show a possible exposure-response relationship with duration of employment (a near-monotonic increase of 0.14 relative risk for each of five decade of fire service, nonparametric $p < 0.07$) and a statistically significant excess (at $p < 0.01$) for fire fighters with more than 40 years of fire service (1.29). Heyer et al. [15] reported an overall risk of only 97 (95% CI 65-139) but observed

an elevated risk among fire fighters aged 65 years or more, when the incidence of lung cancer tends to peak. Thus, even in so-called "negative" studies there are hints of a possible association.

Among those studies that appear to be unequivocally "negative", Beaumont et al.[26] reports the lowest risk (0.84, 95% CI 0.64 – 1.08). This same study is unusual among the major studies because it also shows the largest healthy worker effect, the lowest overall mortality from all causes 0.90 and the lowest mortality rate cancer (0.95) an atypical age distribution and a high rate of cirrhosis.

At the other extreme is one study by Hansen et al. in which an overall risk of 163 (95% CI 75 – 310) was accompanied by a tripling of risk (317) for firefighters aged 60 to 74.[25] This is an imaginative Danish study that aggregated other occupational groups into a synthetic reference group. The artificiality of this construct makes the study difficult to interpret, however.

In our study of urban fire fighters in Alberta,[13] we found trends that we believe suggest a true SMR on the order of 150 in that population. Individually, these trends are not definitive but together they are highly suggestive. The overall SMR for lung cancer was 142 (95% confidence interval 91, 211), statistically not significant, and statistically indistinguishable from 150. However, lung cancer was elevated to an SMR of 167 among fire fighters entering the fire service in the 1960's, the most recent cohort at the time of the study for which the expected latency period had elapsed. This is not strong evidence, because it is based on only two cases, but the following cohort of firefighters entering in the 1970's showed an even greater risk, 261 (although based on a single case). The risk of lung cancer also showed an exposure-response relationship in our data, with groups of fire fighters who had higher exposure opportunities and duration showing elevations on the order of 200. By duration of employment, an initially high risk for those with less exposure declined with duration of employment but achieved a doubling for those working 40 or more years (although only two firefighters were in that group). More persuasively, when duration of employment was corrected for exposure opportunity in job classification, the exposure-response relationship changed to suggest, following an initially high risk among probationary fire fighters or those unfit for duty, a more or less consistent but low elevation for the middling exposed varying around 150 (range 32 to 258), and a significantly elevated risk (408, $p < 0.05$) for those with more than 35 exposure opportunity-weighted years of employment.[13] Baris et al.[12] although negative overall, appears to show the same effect in the first 9 years. Unfortunately, the data from other studies cannot be disaggregated on the same basis as the Alberta cohort.

An important factor in the Alberta study, which was not appreciated at the time of initial publication, is that cigarette smoking is historically less of a confounding factor in Alberta than it has been in other populations. Subsequent studies of smoking-related lung disease outcomes suggest that smoking rates have been historically low in the province compared to the rest of the country and this is reflected in lower mortality from chronic obstructive pulmonary disease. In recent years mortality rates for smoking-related disorders appear to have converged with the rest of Canada as smoking rates in the rest of the country have gone down and those in Alberta have changed less dramatically. Again, this suggests, but does not prove, that the Alberta experience is less confounded by cigarette smoking than elsewhere.[29]

An anomaly of the Alberta data is that the excess was seen in one city (Edmonton) and not another (Calgary). In Edmonton alone, the risk was 201, the highest overall risk for lung cancer reported.[13] This represents an internal replication because the same study team collected data from both cities, matched against death certificates concurrently and analyzed both datasets simultaneously.

Taken together, and supported by the methodologically stronger studies in the literature, 150 seems to be a reasonable estimate of the true (unconfounded) risk for lung cancer among firefighters. The attributable risk fraction would therefore be on the order of 50% for firefighting as an occupation. For the average firefighter, therefore, the most likely estimate of the risk associated with working as a firefighter would be about half that of the risk associated with living in the community.

The findings of epidemiological studies are not necessarily applicable to the circumstances of an individual case. Claims under workers' compensation and other adjudication systems are generally required to be based on individual circumstances, not on broad generalizations, unless there is a relevant presumption and no unusual circumstances to rebut it. One of the individual factors of greatest practical importance is smoking.[6]

When lung cancer occurs in a firefighter who does not smoke, the relevant comparison is to the risk of other nonsmokers, not the population as a whole, which includes many smokers. For a non-smoking firefighter, the a priori risk for lung cancer is low. Is the additional risk attributable to fire fighting sufficient to achieve a doubling, the threshold for presumption? There is evidence that it is but some reasonable assumptions are required.

There is no study available that describes the experience of non-smoking firefighters. This is not unusual: it is difficult to identify or to partition out the risk of non-smokers in most epidemiological studies of occupational risk factors. Although lung cancer is rare in people who do not smoke, when it occurs it is usually adenocarcinoma. However, adenocarcinoma is also increased among smokers, so tissue type does not help as an indicator in the individual case. The association between firefighting and lung cancer is not simple. There is no question that firefighters are potentially exposed to numerous carcinogenic substances in the course of their work. SCBA is highly effective in protecting against these exposures, but it is not necessarily used for low-level fires and is often dispensed with just after the fire is put out, during the overhaul phase of ember suppression, when airborne carcinogen levels are actually highest.

Lung cancer in a non-smoker is rare. Although lung cancer is rare in people who do not smoke, when it occurs it is usually adenocarcinoma. We do not know much about how the risk factors for non-smokers compare to the risk factors for smokers, apart from smoking. The reason is that the smoking effect is so strong that it makes it difficult to isolate other factors in epidemiological studies. It is also difficult to identify non-smokers in most epidemiological studies of occupational risk factors. A known factor is heredity. Lung cancer tends to run in families, but the predisposition tends to express itself only when there is an environmental exposure, such as smoking.

An elevated risk is most clearly demonstrated in our own data among firefighters in Alberta with intensive exposure or long duration of service. For firefighters overall there is a consistently elevated risk but the magnitude of the increase is small and statistically uncertain. The problem is that the group contains both smokers and non-smokers and there is no study available of which I am aware that describes the experience of non-smoking firefighters.

The findings of epidemiological studies are best estimates for the individual, but they are not necessarily applicable to the circumstances of an individual case. A firefighter who does not smoke and is relatively young has an a priori risk for lung cancer that is low compared to the population of firefighters as a whole, an occupational group that includes smokers. (In the 1980's, perhaps 30 to 40% of firefighters smoked; the data available are sketchy but seem to be more or less in line with the general population. These smokers would contribute the majority of cases of lung cancer, as they do in the general population.

In calculating the SMR or relative risk, both the numerator and the denominator typically include smokers. Smokers among the firefighters contribute the great majority of cases of lung cancer, as they do in the general population. Although their risk may be increased compared to similar smokers who do not fight fires, the increase is probably small in absolute terms, given their already increased risk from smoking, which is in the range of 5 to 10 times that of nonsmokers.[20] In the 1980's, perhaps 30 to 40% of firefighters smoked; the data available are sketchy but seem to be more or less in line with the general population.[28] The question therefore is how to estimate the relative risk of nonsmoking firefighters when most of the cases are already coming from smokers.

One may assume that, within a reasonable range of exposure, the magnitude of an increase in risk for lung cancer that is associated with a given exposure to combustion products from fighting fires would be the same for smoking and non-smoking fire fighters. This exposure is added to the greatly increased risk sustained by smoking firefighters who receive much more intense exposure to similar and probably more potent carcinogens in cigarette smoke. [2,6] We may therefore assume a model in which the risk of exposure to combustion products from fires and the risk from smoking are roughly additive. For smoking firefighters, the risk arising from work is added onto the existing risk derived from cigarette smoking, which is about ten times the risk of lung cancer experienced by nonsmokers, overall. Thus, if the risk of lung cancer is increased by 50% for smoking firefighters, the proportionate increase in risk for non-smokers would be much greater, by as much as tenfold, because the same attributable risk is added to a much smaller baseline risk. Seen another way, the relative risk will be hugely increased if nonsmoking firefighters are compared to nonsmokers in the general population, because the risk attributable to occupation would be compared to a much smaller baseline risk for the reference population.

One approach to quantifying the risk of nonsmoking firefighters is to estimate that 40% ($f = 0.4$) of fire fighters smoke and that 60% do not ($1 - f$), that the relative risk (RR) of lung cancer for smokers is 10 times that of nonsmokers ($R = 10.0$), and that the relative risk (r) of lung cancer for fire fighters overall is 1.5. If x represents the attributable risk fraction,

$$RR = 0.4(10+x) + 0.6(1+x)/0.4(10) + 0.6(1.0) = 1.5.$$

Solving for x yields an attributable risk fraction of 2.3. The relative risk is that number plus 1, representing the baseline risk of the unexposed, general population. This translates to a relative

risk for nonsmoking fire fighters of 3.3, comfortably above a doubling. The exact value is unimportant because of the compounded uncertainties; that it exceeds a doubling is what matters.

Another way to approach the problem is to determine, based on the same assumptions, what the minimum relative risk for the firefighters as a whole would have to be to reflect a true doubling of risk for nonsmoking firefighters. The calculations are similar and yield $r = 1.22$, which is comfortably supported by the world literature (whether or not the true risk is 1.5, as has been argued above). How sensitive is this model to underlying assumptions? Reducing the estimate of the proportion of the firefighting population that smokes to 30% barely changes the overall relative risk required to support the presumption, to 1.27. Reducing the estimate of the relative risk associated with smoking from 10 to 5, which is a low estimate and which intentionally biases the model against nonsmokers, increases the overall relative risk required to support the presumption to 1.38, still in line with the world literature and below the 1.5 level that probably represents the "true" risk. Again, the exact number is unimportant; what matters is that the overall risks that would be associated with a doubling in the subgroup of nonsmoking fire-fighters falls into an area entirely consistent with the literature and therefore best evidence.

However, the most relevant comparison of all is a simple ratio to the nonsmoking population using the attributable risk function defined above. If a nonsmoking firefighter were compared to a similar population of people who also do not smoke, the expression would be $0.6(1 + x)/0.6(1.0) = 1 + x = 3.3$, which is the relative risk given above. (This is not coincidence, just the result of a mathematical identity. The group risk of people who do not smoke is defined as unity.) In other words, compared to nonsmokers as a group, smoking firefighters have much more than a doubling of risk. Again, the exact value is unimportant because of the compounded uncertainties; that it clearly exceeds a doubling is what matters most. Thus, it seems apparent that the available evidence supports the conclusion that the risk for lung cancer among nonsmoking firefighters is at least doubled compared to the general nonsmoking population.

Extant data suggests that any recorded history of smoking is associated epidemiologically with an increased risk of lung cancer. However, applying these findings to workers' compensation require operational definitions of what it means to be a non-smoker. The calculations above assume that a nonsmoker has never smoked, an unrealistic assumption for most claimants. Epidemiological data suggest that the square of the number of cigarettes consumed per day and the fourth- or fifth-power of the length of smoking in years are terms that

contribute independently to lung cancer risk. There is a therefore need to define nonsmoker operationally with respect to both intensity in terms of consumption and duration of smoking habit for purposes of identifying cases in which the smoking habit can be considered negligible for the purposes of adjudication.[20]

Never-smokers are generally assumed to have had no experience with tobacco and to be true nonsmokers. However, most nominally nonsmokers have at least experimented with tobacco and many have smoked and quit. Without spelling it out, some studies appear to consider up to 10 cigarettes per day as the threshold for counting a subject as a true smoker and nomograms for calculating risk (such as that put on-line by Sloan-Kettering Hospital at http://www.mskcc.org/mskcc/html/12463.cfm) seldom go below this number, although the risk of lung cancer is still increased compared to never smokers. Likewise, a duration of smoking of 20 or 25 years is often taken as a rule of thumb for a substantial increase in lung cancer risk, reflecting the fifth-power exponential function relating exposure duration to outcome. Data from one relevant study[20] (of Chinese men) suggests that smoking up to 19 cigarettes is not associated with an increased risk of lung cancer until 29 years of duration and that more than 19 cigarettes per day must be consumed for an increase in risk before 29 years duration; however, these data report consumption categories, not precise exposure-response relationships, and are crude associations. It is probably going too far, given the state of the evidence, to claim that either consumption of ten cigarettes a day or less for any duration or duration of less than 29 years at any intensity qualifies a firefighter to be considered equivalent to a nonsmoker. However, it is reasonable to assume that smoking less than ten cigarettes a day and for less than 20 years (as indicated in extant studies) constitutes an exposure that probably does not contribute an excess risk of the same magnitude as that of firefighting among never-smokers.

It is also well established that a person who quits smoking has approximately half the risk of lung cancer ten years later that would be the case if that person continued smoking. The most optimistic major study cited in the reference literature[20] (the American Cancer Society study) excluded smokers who smoked less than one pack per day and suggests that the risk of lung cancer in these smokers subsequently dropped to near-baseline after 10 years. Other studies, particularly the study of British physicians, suggest that twenty years was required to approach a relative risk approximately equal to the threefold risk calculated above for firefighting among nonsmokers. In the absence of specific data to apply to the problem, ten to twenty years serves as

a reasonable indicator of the time since cessation that a previously smoking firefighter would probably return to a baseline risk sufficiently reduced compared to lifelong nonsmokers that firefighting could be deemed the predominant risk factor.

In summary, in the absence of more specific data that would yield a more precise definition, for the purposes of evaluating causation in an individual case a "nonsmoker" may be defined operationally as a firefighter who never smoked more than ten cigarettes a day and has not smoked at that rate for more than 20 years or as a previously smoking firefighter who had quit for twenty years or more.

Other Tumour Sites

There may now be sufficient evidence to consider colon cancer for a presumption. Overall, Baris and co-workers found an SMR of 1.51 (95% CI 1.18,1.93), based on 64 deaths; there was no consistent dose-response for duration of employment nor for cumulative number of runs. However the risks were greater than 1.00 for all three levels, 1.93 for low; 2.22 for medium and 1.22 for high number of runs. It is worth noting that excess colon cancer risk was also reported by Guidotti[13]; Howe and Burch[4]; Schwartz and Grady[23]; and Vena and Fiedler [24], who reported a significant SMR of 1.83. Thus, two studies, one in two out of three subgroups and the other in the population as a whole, have demonstrated relative risks close or equal to a doubling. Although not as easily demonstrable, it may be argued that claims for colon cancer may be justified in the same way as for lung cancer in an individual with a low a priori risk for the disease. We have argued this in the case of a young vegan with no family history of the disease or of polyps, but the claim was not accepted. In the opposing opinion in the case, it was pointed out that there are differences in risk factors between colon and rectal cancers. This is true only insofar as there exists a set of distinct risk factors for rectal carcinoma that are primarily associated with lifestyle. However, both colon and rectal cancer share all risk factors, including diet or genetic predisposition, which are known risk factors for colon cancer. Thus, unless one posited a major difference in these lifestyle factors between firefighters and the rest of the population, there is no a priori reason not to combine colon and rectal cancers into one category of colorectal cancer. Occupational exposures that affect cancer risk in the colon may reasonably be expected to affect cancer risk in the rectum.

Critique of the Cancer Care Ontario Report

We have received the final report of the Cancer Care Ontario (CCO) team, in the form of a document entitled "The Occupation of Firefighting and Cancer Risk: Assessment of the Literature: Report to the Workers['] Compensation Board of British Columbia" [sic – the report omitted the apostrophe], dated 15 June 2004. We also received the "Addendum to the Final Working Draft", dated March 2004 and the Peer Review and the response to the peer review. All five documents have been reviewed.

We have also received the penultimate draft, dated February 2004, which was the version circulated to reviewers and which was labeled "Working Final Draft: Subject to Revision Following Review". Ordinarily, we would not comment on a working draft. However, in this case we shall because of the unexplained change between drafts in describing the objectives of the study.

We begin with the peer review submitted by Drs. Band and Parker.

Peer review of the report

The comments by Drs. Band and Parker are compatible with most of the comments that will follow. Their general impression of the report can be summed up

Specific comments included:

- The title is misleading because this is not a comprehensive assessment of the literature or of the association of cancers and firefighting.
- Literature search may not have been complete.
- Terminology was not standard.
- Exclusion criteria were unclear.
- No consideration was given to exposure assessment.
- The omission of surveillance studies (meaning PMR studies) jettisons useful information.
- Reviewers disagreed with decisions to exclude or include certain studies.
- Reviewers essentially conducted their own meta-analyses and came to different conclusions in about half the cancer types on the issue of sufficiency of evidence.

- Reviewers state, and discuss at length, that the CCO report addresses only sufficiency of evidence, almost entirely on the basis of strength of association, and does not address other important criteria for association and indicators of causation.
- Reviewers state that the team members call for further research but in doing so do not appear to appreciate that the exigencies of workers' compensation require that the evidence that is currently available must be weighed at the time.

Among comments submitted by the reviewers are the observation that "the final report … does not go far enough" and that the team members "large ignore other Bradford-Hill criteria in their final conclusions." The reviewers also observed that "the authors attribute causality on what is largely a statistical number and not the sum of the relationships…"

In general, we agree with the reviewers' comments, and offer additional comments below. (Note: Most, but not all, of these comments were incorporated in a report to the British Columbia Professional Fire Fighters Association and are adapted from that report.)

Overall impression of the report

The CCO team report is misrepresented in scope and application, based on a faulty premise and imperfectly executed.

It has yielded results at variance with other thoughtful analyses, such as IDSP report. It is instructive to compare the CCO report with the IDSP report. Sufficiency of evidence was one, among several criteria used by IDSP in 1994, and the absence of evidence would have precluded conclusions about association or presumption. As well, the literature on firefighters has grown in the last ten years and one might therefore expect there to be more sufficient evidence today for certain outcomes. Remarkably, however, a comparison of the CCO team report and the IDSP report shows marked variations in conclusions. One is left to wonder how the CCO can judge the extant evidence to be insufficient for one condition but strong enough for the IDSP team to conclude that a presumption was warranted. (The IDSP team was larger and experienced in occupational epidemiology.) See Table 3.

Table 3. Comparison between the CCO and IDSP team reports on firefighters and cancer risk.

Outcome	**CCO Team Report, 2004**	**IDSP, 1994[7]**
Criteria:	*Sufficiency of evidence, only*	*Weight of evidence*
Lung	No evidence	No association
Colorectal	Limited	Association
Bladder	Insufficient	Association
Kidney	No evidence	Association
Testes	Insufficient	N/A
Leukemia	Insufficient	Presumption
NHL	Insufficient	Presumption
Myeloma	Insufficient	Presumption?
Brain	Limited	Presumption

Body of the report

The Background section of the report lacks a review of the literature, which is customary in such reports. The authors fail to address the many specific problems and limitations of occupational epidemiology, including the healthy worker effect (mentioned in passing in the Discussion), misclassification bias (a more likely cause of underestimation), the proper exposure metric (a central problem in firefighters studies) and confounding exposures.

There is no mention of the central problems in firefighter research: the issues of changing combustion products and of personal protective equipment. The introduction does not convey the complexity or of knowledge of common problems and issues associated with assessing risk arising from occupational exposures. Many of the word choices are peculiar and suggest a lack of familiarity with firefighter terminology (e.g. “after the fire is exhausted”).

There is a useful (but very brief and incomplete) subsection on “firefighter characteristics”, by which the authors apparently mean lifestyle characteristics. There are other characteristics of this working population they do not mention, such as the low turnover, the frequent occurrence of sideline employment (“moonlighting”), the fitness requirement and the ethnic and demographic profile of the occupation.

On page 2, the discussion of occupational epidemiology is rather naïve. The authors state “studies have suggested an underestimation of occupational cancers because occupational hazards are not the subject of systematic study.” On the contrary, some occupational hazards have been investigated exhaustively (e.g. asbestos). It is also logically fallacious, because lack of systematic investigation does not imply underestimation (it could as easily lead to selection bias that would overestimate the risk). It would be more accurate to say that the risk of occupational

cancer in firefighters is underestimated because exposure assessment is imperfect, potential biases almost all lead to underestimation, and the risk estimates of many outcomes of interest are diluted by aggregation.

The Background section also makes no mention of the standard of certainty applied to the evidence. The fundamental principle of workers' compensation, which is the balance of evidence, is nowhere mentioned. Apparently by default, the team applies a scientific standard of certainty (95%) to a problem requiring balance of probabilities (50% + 1). It is unclear whether they were aware of this fundamental assumption, as it is nowhere mentioned.

The section on Objectives was originally stated, in the penultimate draft of the report, as a single sentence that reads "To determine if, based on the published literature, there is evidence to support a relationship between the occupation of firefighting and risk of specific cancers among firefighters." In the final version of the report this has changed. The "goal" is now stated as "to synthesize and analyze the scientific/medical literature regarding the causal association between the occupation of firefighting and specific cancers". It is disquieting that something as fundamental as the objectives of the study could change from one draft to another.

The two sets of objectives also have different interpretations. The first implies a standard of evidentiary strength, which this study report has attempted. The second implies a broad mission to interpret and integrate material across disciplines, which this report has not done. Neither objective, in fact, is correct for such an exercise. Because the standard applied in workers compensation is the weight of evidence, not its sufficiency, the investigators are evaluating the wrong aspect of the question. They are determining whether the current body of evidence is sufficient or lacking to draw a conclusion with certainty. This is not helpful, for several reasons. First, the workers' compensation system requires resolution of cases in a timely manner on the basis of the existing evidence, not a judgment on whether the existing evidence is sufficient – cases are not going to wait for the evidence to improve. Besides, over the last five decades the weight of evidence for disease outcomes associated with firefighting has generally grown stronger, not weaker. The passage of time, in other words, is demonstrating new likely associations as evidence accumulates. Second, the evidentiary base for occupational risks among firefighters is as complete as it is for any other common profession: it is therefore rather ingenuous to fault it for being incomplete. Third, the weight of evidence is not the standard used in this report, although the adjudication criteria for workers' compensation are already very

specific in requiring a balance of probabilities. Thus, the report unavoidably and inherently biases its findings against the claimant, in violation of the Workers' Compensation Act of BC, which requires that the benefit of the doubt be given to the claimant.

The section on Methods presents problems for the careful reader. In practice, the methodology is a simple statistical algorithm that is applied much like a cookie cutter to the overall findings in a set of selected studies, ignoring relative strengths and features within the studies chosen and discarding the more detailed information found in the more extensive studies.

The subsection on exclusion criteria labels three studies on firefighters in the Northwest US as "duplicate" but this is incorrect. The three represent attempts to estimate risk more accurately by different means and with different local firefighter groups and are therefore better characterized as overlapping; they are not duplicate studies in the usual sense.

The description of quality assessment is wholly unsatisfactory. The Quality Assessment procedure is not specified, only that a three-point scale was used "for each dimension", but what is meant by "dimension" is unclear. (In the prior section there are two lists of characteristics and parameters for which data were abstracted but no "dimensions" of technical quality.) The Methods description states that each article was reviewed by one "scientist" with expertise in the area of occupational and environmental epidemiology (which is limited in this team) and one "research associate", who can be assumed to be inexperienced in the field (see above). Given the overall lack of experience in occupational epidemiology of the team, this is not reassuring.

The report lists a "number of issues raised" with respect to cohort studies but does not make it clear that most of the studies omitted are old and that Deschamps et al. [8] was only a preliminary study conducted to support a future cohort study. On another occasion, in response to a question put to them for clarification by the WCB of BC, the team failed to recognize that Eliopoulos et al.# was a study on a special population of firefighters engaged in suppressing brush fires (in Western Australia); it is rather basic, when examining an epidemiological study, to determine what population is being studied.

The study team omitted all PMR studies indiscriminately, on the grounds that "the authors intend them to be a surveillance tool". This omission on purported grounds that the team can read the mind of the authors misrepresents the intent of Burnett et al.[5], which is spelled out in their paper. Burnett et al. is a special case of a PMR study that was conducted by NIOSH to validate associations that were reported by other studies of firefighters, in particular the Alberta

study.[13] It was intended to be a test of a hypothesis, not a surveillance tool.[9] and confirmed the result (elevated risk of cancer of the kidney) in a different and much larger population. Burnett et al. is also a special case because it overcomes, through sheer numbers, many of the limitations of more conventional PMR studies. Thus, in omitting Burnett et al., the team threw out the closest thing to hypothesis testing in the firefighting literature, at least before the recent studies on testicular carcinoma. They may not have realized this at the time. The error was compounded by the team's failure to recognize that a third study, that of Delahunt et al., also constituted a *de facto* test of a hypothesis because they were actively searching for evidence of excess risk of one particular outcome, kidney cancer, not multiple outcomes subject to multiple comparisons. [21]

The team also fails to identify the considerable differences among the studies and the patterns that they show. For example, Beaumont et al. [10] is a consistent outlier among studies of firefighters, recognized as such by workers experienced in the field, but this is nowhere mentioned and appears not to have been noticed by the team. Because the team does not recognize or appreciate these differences, they appear not to have noticed an important characteristic of the better studies (including Hansen [11], L'Abbé/Aronson [12], Guidotti [9] and Baris [13]), which is that they provide evidence for an underlying effect that is misclassified or confounded and which becomes stronger with better exposure assessment. .

One of they key tests used by epidemiologists to determine the magnitude and effect of confounding and misclassification bias is to compare the trend in the association with more refined approaches to exposure assessment. If one observes higher risk estimates for plausible subgroups at risk as the exposure assessment improves and misclassification becomes less of a problem, the effect is more likely to be a primary effect associated with the exposure of interest. The team does not appear to have applied or understood this basic principle or to have noticed the trend in the studies (such as Guidotti et al.[9] and Baris et al.[13]) that consciously applied the method and provide detail on exposure assessment.

Of greatest concern, however, is the basic methodology. A simple algorithm for meta-analysis was applied to the surviving studies in the database, without regard for the nuances of study design, exposure assessment or validity of the diagnostic rubric. There is no effort to understand the biology of the disease or the reasonableness of the outcomes that were "grouped into similar site groups using ICD codes." This appears to be a case of having only a hammer

and all problems therefore looking like a nail. It merely perpetuates the problem of aggregated outcomes, dilution of effects and misclassification bias, without resolving anything.

The last subsection (3.6) in the Methods section features "Criteria for Conclusions". These are rather general guidelines, some of which resemble the Hill criteria and some of which deal with issues of interpretation (such as, the "likelihood that study results are due to uncontrolled confounding or other methodological problems"). Although it is quite valid to provide room for interpretation, the authors are silent on how they came to these conclusions and who did the interpreting and by what consistent framework. It is also of great concern that these criteria were not inserted until the last draft: if they were truly operational criteria for establishing conclusions of the study, they should have been established at the outset and not added later. By the Discussion section in the final report, the team essentially quits the pretense and ignores the Hill criteria altogether, except for strength of association. (This deficiency is discussed further below.)

The Table of Criteria (Table 5, p. 11) outlines the algorithm for each conclusion on evidentiary strength. These criteria (with the sole exception of evidence for a dose-response relationship) are based on summary risk estimates for the entire working population studied. This is not satisfactory. Too much information is lost by not considering job classifications of firefighters, qualified by job assignment or duration of exposure or intensity of exposure, that may provide additional evidence, corroboration or identification of subgroups at elevated risk.

The Criteria also emphasizes "substances" (an example of nonstandard terminology) to which firefighters may be exposed. In assessing an association with firefighting, relevant risk factors are not necessarily toxic exposures. For example, colon cancer may have little to do with combustion-related exposures but may still be work related if they arise out of work. Since firefighters live in the fire hall for the duration of their shift and cook there, dietary factors and stress-related effects on the gastrointestinal tract (well-documented to relate to intestinal motility) may mediate such effects and may well be work-related. Thus, the definition of a plausible causal exposure is much too narrow.

The Findings are presented in stereotyped subsections using essentially identical terminology. The risk point estimates and standard errors derived from the meta-analyses are presented in a log scale, which is not standard and which visually minimizes the appearance of variation, making it appear that point estimates "above" (to the right of) unitary risk are closer

than they would normally appear on a linear scale. Oddly, the standard error bars are of equal length on both sides of the point estimate, although the scale is logarithmic. Is this an error?

From page 31 on, without caption, informative heading or discussion, there are a series of tables (starting with Table 6) that appear to summarize the studies reviewed and that treat exposure-response relationships based on the duration of employment, when available. This is not accompanied by text, which is peculiar for such a report.

The tables omit one major study in which detailed exposure-response data are available for many outcomes (Guidotti, 1993[9]) for no obvious reason and without explanation. The information on exposure in Guidotti, 1993, is provided in the report for kidney cancer and lung cancer only in the form of duration of employment, although a refined exposure index, which takes into account actual opportunity for exposure by job classification for each year, was published in the study and used to classify firefighters by relative exposure. The team apparently either overlooked this aspect of the study or intentionally ignored it, although it is a refinement of considerable importance. Likewise, the study team ignores the considerable detailed exposure-response data beyond duration of employment available in Baris et al.[13] This is an elementary error, the equivalent, if this were a study of nutrition instead of occupational epidemiology, of measuring dietary intake by how long a person has eaten a food rather than estimating actual consumption.

The refusal to use all the information available is another fundamental failure of this study. Inexplicably, the study team confined itself to duration of employment in those studies they chose to examine. Estimating relative exposure by different methods is important not only to assess exposure-response more accurately but, by determining whether the association is strengthened by improvements in exposure measurement, is a critical check on confounding. The study team does not apply this elemental principle of interpretation in occupational epidemiology and gives no reason why they fail to do so. On page 49, the authors mention the importance of exposure opportunity, yet they failed to use this information when it is available.

The Discussion section is superficial, abbreviated and highly unsatisfactory. The Discussion section does not draw conclusions on causation, which is generally the point of adjudicating claims in the first place. Rather it, evaluates the strength of the evidence. While this is clearly an important aspect of interpretation for purposes of adjudication, it is not the only or even the most important step. That is to weigh the extant evidence and to determine the weight or

balance of evidence. The study report does not do this, confining itself entirely to the assessment of strength of summary evidence, only, from comparable studies, rejecting those studies and subanalyses that do not fit the predetermined analysis plan. Thus, the job is less than half done, from the standpoint of proper adjudication.

The discussion of the healthy worker effect (p. 50) is naïve and ignores the literature on firefighters. The authors do not seem to appreciate that there are two healthy worker effects, constituting recruitment and retention bias. Firefighters are subject to both but the retention bias is attenuated by a historical pattern of job assignments that tends to protect firefighters who are less fit. Both patterns are strongest for cardiovascular disease and diminish rapidly after initial hire for cancer, so that the two healthy worker effects are not as great an interpretive problem for cancer incidence and standardized mortality at later ages as they are for other outcomes and for proportionate mortality. Empirically, studies of firefighters have not shown the expected magnitude of the overall (combined) healthy worker for mortality, but there is evidence (discussed in Guidotti, 1993[9]) of an effect in younger age groups. The authors mention none of this and the discussion is highly superficial.

The Discussion section does mention the survivor effect. This is one bright spot in an otherwise highly disappointing Discussion section, because it acknowledges the related issues of competing mortality and "exhaustion of susceptibles" (a concept more often recognized in infectious disease epidemiology), which have previously been disputed in claims before the WCB of BC.

The Discussion section also has significant omissions. Crucially, it fails to address the issue from the standpoint of adjudication. The penultimate draft criticized the sufficiency of the literature, which is even so among the most complete for any occupation. This has been toned down in the final draft and the authors have reoriented the Discussion section toward methodology. It makes no effort to synthesize the literature, to assess major problems, to determine generalisability of conclusions or to summarize the findings.

The report ends with the formulaic comment that further research is needed. However, this apparently benign comment demonstrates that they have missed the point. The problem is not accumulation of detail or assembling an every-larger synthetic population of firefighters to meta-analyze. The fundamental problem is one of disaggregating outcomes and studying each

according to the intrinsic limitations of measurement, bias and confounding found in the study design.

Findings for cancer at specific sites

Despite the manipulations of statistical presentation, the findings are remarkably inconclusive. Given that workers' compensation is rather explicit in its requirement that the claim be based on the weight of available evidence, not its sufficiency for scientific certainty, this inconclusiveness cannot be very helpful. (This subsection is identical to the text of the BCPFFA report.)

Bladder. Bladder cancer is the first outcome considered. The team concludes that insufficient evidence is available, despite noting a consistent (but not statistically significant) excess risk. (The two studies that do not show an increase show consistently lower risk estimates for mortality overall and for other outcomes than other comparable studies in the literature, suggesting that they are different in some systematic attribute.) Certainly the association does not achieve conventional standards of statistical significance but the weight of evidence clearly does suggest an increased risk. Bladder cancer is therefore an example of the most fundamental problem of the study report, which is that the study team has defined a criterion for "limited evidence" that is inappropriate for the adjudication process.

Brain. Remarkably, the authors come to the conclusion that there is limited evidence (in the penultimate draft the conclusion was that the evidence was "sufficient") to conclude that there an increased risk of cancer of the brain among firefighters. This conclusion is reached on the basis of evidence that, on closer inspection, is no stronger than that for bladder (with almost identical pooled risk estimates around 1.4) simply because there are a larger number of studies. (Recall that adjudication under workers' compensation requires consideration of the weight of evidence, not its sufficiency for scientific certainty.)

Colorectal. The final study report concludes that there is limited evidence of an increased risk of colorectal cancer among firefighters. This conclusion appears to have been modified during the course of the study, raising questions about the application of the criteria for a conclusion. Insufficient evidence to determine whether an association exists was originally reported for colorectal cancer in the penultimate draft. Inconclusive evidence to determine

whether an association exists was originally reported for colon cancer and rectal cancer individually and together. Later, the addendum revisited these same outcomes and concluded that there is indeed evidence to support an association. The findings of the addendum carried over into the final report.

Kidney. This section was added to the final report, suggesting that it was omitted from the original analysis. The team concludes that there is no evidence of an increased risk associated with firefighting, despite two relatively strong studies with quite large risk estimates and confirmatory evidence from Burnett et al.[5] (which they do not consider), which is also reflected in Ma et al.[20] and a strong suggestion in a subgroup of Baris et al.[13]. To the findings of these studies using one methodology are added the "positive" findings of another, Delahunt et al.[21], using another method applied to a different population. Given that an excess of kidney cancer is now apparent in several studies although few studies would have the power to detect a true increase in risk, the team members have put forward a heroic conclusion by stating that there is "no evidence" of an association.

Leukemia. The team concluded that there was insufficient evidence to determine if there is an association. The team is correct in observing that leukemia is a heterogeneous group of diseases. (This observation was added to the final version: it does not appear in the penultimate draft.) They do not, however, use this information to interpret their findings or to point out the methodological implications. The implications of this omission are described in detail later in this report.

Lung cancer. Despite elevations in three mortality studies, the authors conclude on the basis of their meta-analysis that there is "no evidence" of an increased risk of lung cancer among firefighters. They come to this remarkable conclusion on the basis of summary statistics alone, without considering the exposure-response relationships or the known differences in population characteristics. The implications of this superficial examination are described later in this report.

Multiple myeloma. The team concluded that there was insufficient evidence to determine if there is an association with multiple myeloma. The report fails to acknowledge that the same heterogeneity as previously noted for leukemia is also a feature of myeloma (and, for that matter, brain). No reason is given for this highly selective discussion. The penultimate draft placed a heavy emphasis on the association with benzene exposure (more so than for leukemia, inappropriately) but this has been removed in the final draft.

Non-Hodgkin lymphoma. The team concluded that there was insufficient evidence to determine if there is an association with non-Hodkgin lymphoma. The discussion of non-Hodgkin lymphoma is very abbreviated for such an important and complicated group of outcomes. Oddly, the penultimate draft acknowledged that this group of disorders "represents a heterogeneous group of cancers of the lymphatic system, which may have different etiologies" but this observation, which is true, was omitted from the final report. The final report therefore treats the rubric as a single entity, which is inappropriate. This issue is discussed further later in this report.

Testicular cancer. The authors conclude that there is insufficient evidence to evaluate the risk of testicular cancer when all four studies show an elevated point estimate and three out of the four studies have demonstrated elevations over 1.5, one as high as 4.30. (Similarly, they make the same conclusion for bladder cancer when seven out of nine studies show elevations.)

Conclusion

Firefighers are an important occupational group and their work-related health problems are issues of importance for themselves, for compensation, and for public safety. The occupational health problems of firefighters are also examples of deeper issues in occupational epidemiology and compensation policy. Studying firefighters brings out these deeper issues.

The application of epidemiology to adjudication and litigation is based on a different set of rules than for scientific investigation. The role of the expert is to give guidance as to the weight of the evidence, not to produce more data or to determine revealed truth through the scientific method. While it is not the role of the expert or the investigator to make such determinations, an understanding of the system and an interpretation based on the principles of the system is more likely to be useful than a scientific interpretation irrelevant to the application.

This issue is a prime example of a class of problem in occupational epidemiology that is best approached by examining the structure of the problem, outcome by outcome, using principles of logic rather than indiscriminately applying advanced statistical techniques.[6] Adjudication under workers' compensation requires an examination of the weight of evidence, not scientific certainty. It is incumbent on investigators in this field to provide analysis, studies and data that will assist in

determining the weight of evidence. However, it is also incumbent on those who apply the evidence to examine it in full and not just parts of it.

Bibliography

1. Guidotti TL, ed. Emergency and Security Services. Chapter 95 in: ILO Encyclopaedia of Occupational Health and Safety. Geneva, International Labour Organization, 1998.

2. Guidotti TL, Clough V. Occupational health concerns of firefighting. Ann Rev Pub Health 1992; 13:151-171.

3. Guidotti TL Occupational epidemiology. Occupational Medicine 2000; 50(2):141-145.

4. Howe GR, Burch JD. Fire fighters and risk of cancer : An assessment and overview of the epidemiological evidence. Am J Epidemiol 1990; 132:1039-1050.

5. Burnett CA, Halperin WE, Lalich NR, Sestito JP. Mortality among firefighters: a 27 state survey. Am J Ind Med 1994;26:831-833.

6. Guidotti TL, Rose SG. *Science on the Witness Stand: Evaluating Scientific Evidence in Law, Adjudication and Policy.* Beverley Farms MA, OEM Press, 2001.

7. Industrial Disease Standards Panel. Report to the Workers' Compensation Board on Cardiovascular Disease and Cancer Among Firefighters. Toronto, WCB of Ontario, September 1994.

8. Guidotti TL. Occupational mortality among firefighters: assessing the association. J Occup Envir Med 1995;37(12):1348-1356.

9. Deschamps S, Momas I, Festy B. Mortality amongst Paris fire-fighters. Eur J Epidemiol 1995 Dec;11(6):643-6.

10. Ma F, Lee DJ, Fleming LE, Dosemeci M. Race-specific cancer mortality in US firefighters: 1984 – 1993. J Occup Environ Med 1998; 40(12):1135-1138.

11. Bates MN, Fawcett J, Garrett N, Arnold R, Pearce N, Woodward A. Is testicular cancer an occupational disease of fire fighters? Am J Ind Med 2001;40:263-270.

12. Baris D, Garrity TJ, Telles JL, Heineman EF, Olshan A, Hoar Zahm S. Cohort mortality of Philadelphia firefighters. Am J Ind Med 2001;39:463-476.

13. Guidotti TL, Mortality of urban firefighters in Alberta, 1927 – 1987. Am J Ind Med 1993;23:921-940.

14. Demers PA, Heyer N, Rosenstock L. Mortality among firefighters from three Northwestern United States Cities. Br J Ind Med 1992;49:664-670.

15. Heyer N, Weiss NS, Demers P, Rosenstock L. Cohort mortality study of Seattle fire fighters: 1945 – 1983. Am J Ind Med. 1990;17:493-504.

16. Goldsmith JR, Guidotti TL. Environmental factors in the epidemiology of lymphosarcoma. Pathol Annu. (Part II) 1977;12:411-425.

17. Tatham L, Tolbert P, Kjeldsberg C. Occupational risk factors for subgroups of non-Hodgkin lymphoma. Epidemiology 1997 ;8(5):551-8

18. Committee to Review the Health Effects in Vietnam Veterans of Exposure to Herbicides (Third Biennial Update), Institute of Medicine. *Veterans and Agent Orange: Update 2002*. Washington DC, National Academy Press, 2003.

19. L'Abbé KA, Tomlinson GA. Special Report: *Mortality Study of Fire Fighters in Metropolitan Toronto*. Toronto, Ontario Industrial Disease Standards Panel, 1992.

20. Schottenfeld D, Fraumeni JF. *Cancer Epidemiology and Prevention*. New York, Oxford University Press, 1996.

21. Delahunt B, Bethwaite PB, Nacey JN. Occupational risk for renal cell carcinoma. A case-control study based on the New Zealand Cancer Registry.Br J Urol. 1995;75(5):578-82.

22. Stang A, Jökel K-H, Baumgardt-Elms C, Ahrens W. Firefighting and risk of testicular cancer: results from a German population-based case-control study. Am J Ind Med 2003;43:291-294.

23. Schwartz E, Grady K. Patterns of occupational in New Hampshire (1975-1985). New Hampshire Division of Public Health Services, Bureau of Disease Control, 1986.

24. Guidotti TL, Baser M, Goldsmith JR. Comparing risk estimates from occupational disease monitoring data. Public Health Review 1987; 15:1-27.

25. Hansen ES. A cohort study on the mortality of firefighters. Brit J Ind Med. 1990;47:805-809.

26. Beaumont JJ, Chu GS, Jones JR, Schenker MB, Singleton JA, Piantanida LG, Reiterman M. An Epidemiologic study of cancer and other causes of mortality in San Francisco firefighters. Am J Ind Med. 1991;19:357-372.

27. Vena JE, Fiedler RC. Mortality of a municipal-worker cohort: IV. Firefighters. Am J Ind Med. 1987 (11):671-684.

28. Gerace T.A. Road to a smoke-free fire service for Florida: policies and progress. J Public Health Pol. 1990;11(2):206-217.

29. Guidotti TL. Trends in mortality from COPD in Alberta: back to the future? Can Resp J 1995;2:97-103.