National Collaborating Centre for Environmental Health
Centre de collaboration nationale en santé environnementale

A Primer on Evaluating the Quality of Studies for Environmental Health

Background on Epidemiological Studies

Basic Concepts

Epidemiology is the study of causes, distribution, and control of disease in human populations. Epidemiologic methods can be used to identify public health problems, determine the most likely causes, and evaluate efficacy of control measures through intervention studies.[1-3]

Epidemiological studies can provide the most compelling evidence for environmental risks to humans. The strongest study design is experimental, with the investigator assigning either the exposure or the intervention. The weakest study designs are descriptive, because they cannot directly ascertain a relationship between exposure and outcome.

Evaluations of the quality of epidemiological studies have in common the consideration of study design, sources of error, exposure ascertainment and evidence of causality. Further critical appraisal of cohort/intervention, case-control and cross-sectional studies are illustrated in separate documents. Examples of hypothetical studies involving exposure to pesticides are used to illustrate each concept.

Experimental Study Designs	Examples
Experimental study: The investigator manipulates the study factor and randomly assigns subjects to the exposure (or treatment) groups and to the non-exposed groups, example: randomized clinical trial.	Over a season, farm workers were chosen at random to spray either malathion or an inert control substance on berries; researchers were blinded to the exposure group when assessing the incidence and severity of skin rashes.
Intervention trial: An intervention (typically oriented towards education and behavioural change) is applied to the study group and not to the unexposed controls. Random assignment is preferred.	Three out of six similar communities were assigned to an intensive pesticide awareness campaign and three were not. The use of alternative pest-control methods were assessed after one year.

Observational Study Designs	Examples
Cohort study: A group of persons without the disease (but who may or may not be exposed) are selected on the basis of their exposure status and followed over a lengthy period of time, either prospectively or retrospectively. Typically examines multiple health outcomes related to a specific exposure.	Pregnant mothers were enrolled in a study on household pesticides. The children of households using either synthetic or natural indoor insecticides at the time of pregnancy were each compared to those with no pesticide use. The principal outcome was the incidence of asthma in their children.
Case-control: The study group includes people with the outcome or disease status of interest (cases) and people without (controls). The past exposures of cases are compared with those of controls to evaluate which exposures are associated with the outcome or disease status. Typically evaluates multiple exposures in relation to a disease.	Emergency room records were used to identify children treated for severe asthma attacks. The comparison group was children treated for injuries at the same hospital emergency department. The families were then interviewed about household pesticide use, among other exposures.
Cross-sectional: Exposures and outcomes are determined at the same point in time. They are usually quick and inexpensive and have a role for generating hypothesis.	A cross-sectional survey was done on children attending an asthma camp. Included in the survey were questions for parents on asthma severity, household pesticide use and other exposures.

Descriptive Study Designs	Examples
Case report or series in which observations are made on one or a few subjects.	The case of a child with fatal asthma found household use of synthetic pesticides to be one unusual exposure.
Ecologic or correlation studies: Examine rates of disease by population-level factors. Groups (rather than individuals within groups) are the unit of analysis. Ecological fallacy is the inability to apply group-level relationships to individuals.	Residents of urban areas have asthma rates that are increasing over time and sales data on pesticides have also increased over time. However, we have no idea whether asthmatics are more likely to have used pesticides.
Surveillance studies: Focussed, systematic, and routine collection of health or exposure data, including monitoring morbidity and mortality rates or determinants of health.	From a provincial database health surveillance program, reports of asthma fatalities in children were found to be higher since 1990 which coincides with the introduction of pyrethroid pesticides.

Disease clusters: Occurrence of a greater than expected number of cases of a particular disease within a group of people, a geographic area, or a period of time.	A number of employees of a pesticide manufacturer and some members of their families were diagnosed with cancer. The numbers of cases were higher than expected in the general population.

Methods: Exposure Ascertainment	Examples
Exposure definition: What was the primary exposure of interest? Was the exposure clearly defined and accurately measured? Were subjective or objective measurements used and were they validated?	The level of exposure to the pesticide diazinon, determined as low-medium-high based on acetycholinesterase level, was associated with skin cancer in a dose-response relationship.
Route of exposure: Was the route of exposure (inhalation, ingestion or skin contact) defined?	Homeowners applying herbicides often do not wear adequate protective gloves and can be exposed by skin contact.
Exposure duration: Was environmental exposure measured for the short- or long-term? Over the long-term, were previous measurements used? Were all sources and changes considered?	Cumulative exposure to the organophosphate took into account current restricted use as well as past permitted uses.
Exposure measurements: How was personal exposure measured? If biological markers were used, were they appropriate for short- or long-term exposures?	Certain pesticides change the level of a specific enzyme in blood. The magnitude of the change indicates dose absorbed.
Exposure validation: Was there any attempt to verify the exposure metric with another source?	Exposure histories by questionnaire were compared to biological markers of exposure.

Random Sources of Error	Example
Random error: Reflects fluctuation around a true value because of poor precision, sampling error, and variability of measurement. Non-differential misclassification (i.e., no difference in misclassification between those with and without the outcome of interest) of the exposure or outcome will weaken associations, such that the observed effect estimate will be closer to null than the actual relationship.	Exposure to synthetic household pesticides was determined by cumulative exposure (number of years by type of pesticide) based on subject interviews. This imprecise estimate of exposure could result in misclassification of exposure for an individual.

Systematic Sources of Error (Bias)	Examples
Confounding: A distortion of the association between an exposure and outcome due to the relationship of another risk factor with both the exposure and the outcome. A variable cannot be a confounder unless it can affect the outcome and it is distributed differently among the compared populations. "Unknown confounders may have affected the result" is a criticism often applied to any study.	A significant association was found between pesticides exposure in children of farm labourers and their performance in intelligence tests. However, socioeconomic status is related both to farm pesticides exposure and to the outcome of IQ deficits in children.
Selection bias: An error due to systematic differences in characteristics between those selected or not selected for a study. Examples include *self-selection bias* (difference in response rates) and the *healthy worker effect*.	Asthma rates among pesticide applicators were lower than in the general population. (Asthmatic workers would tend to avoid such occupations and the general population includes those ill with asthma).
Information bias: Results from a systematic difference in the way the exposure or outcome is measured between compared groups. • *Recall bias* is a differential level of accuracy in the information provided by compared groups. For example, cases may be more likely than controls to recall and report prior exposures. • *Interviewer bias* is a systematic difference in soliciting, recording or interpreting information.	Interviewers, aware of the health status of the children participating in the case-control study, surveyed the parents about past exposure to household pesticides. A strong association between use of cosmetic herbicides and nervous disorders was found. (Parents with ill children are more likely to recall suspected exposures, particularly if the problem has received media attention. Interviewers may probe the known cases more thoroughly).

Measures of Effect	Examples
P-value: The probability that the study results are determined by chance. A significance level of 5% is typically set as the accepted risk of not identifying a relationship where one actually exists (Type I error).	The relationship between a history of insecticide exposure in children with a diagnosis of leukemia was found to be statistically significant ($P<0.05$).

Multiple comparisons: Although adjustments for making multiple comparisons of data could theoretically avoid rejecting the null hypothesis too readily (Type I error), it may also increase the probability of accepting the null hypothesis for associations that are not null. Generally one should avoid doing any formal corrections for multiple comparisons when the study focuses on only a few planned comparisons, rather than every possible comparison.	The relationship of three types of synthetic pesticides with cancer was evaluated for 20 cancers. By reporting all the individual P values and confidence intervals, one can compare the actual number of "significant" p-values to the expected number of comparisons having uncorrected p-values less than 0.05 if the null hypotheses were true,
Confidence intervals: Margins of error typically given as a 95% probability that the risk estimate will be within the stated interval. Narrow confidence intervals imply greater precision and are affected by sample size and variability.	For children prenatally exposed to insecticides the odds ratio for leukemia was 1.8 (95% CI 1.4, 2.0). However, early childhood exposure to insecticides was not associated with leukemia (OR 1.1, 95% CI 0.8, 1.3).
Power: The measures of a study to demonstrate a true association, which is highly affected by sample size. Type 2 or β (beta) error occurs if a result is non-significant ($p>0.05$) when there is actually a true difference between groups. Power is calculated as ($1-\beta$) and is typically set at 80%.	20 men and women were compared for digital tremor after using synthetic household pesticides and no differences were observed. However, it was later determined that the study was underpowered and to potentially show significant results a sample size of at least 90 subjects was needed per group.

Hill's Guidelines for Assessing Causation[4]	Example
Temporal Relationship: Exposure always precedes the outcome. This is the only absolutely essential criterion.	Residents near to fields sprayed by paraquat, between 1974 and 1999, were later diagnosed with Parkinson's disease (after 2005).
Strength: This is defined by the size of the association as measured by appropriate statistical tests. The stronger the association, such as a larger odds ratio, the more likely it is to be causal.	For the association between the herbicide paraquat and Parkinson's disease, the odds ratio was 2.0 (95% CI1.6, 2.9), indicating that those living close to fields sprayed with paraquat had twice the risk of having Parkinson's disease.
Consistency: The association is consistent when results are replicated in studies in different settings using different methods.	The increased risk for Parkinson's disease, related to use of paraquat and other pesticides, was demonstrated in 3 case-control studies and a cohort study in 4 crop growing areas in North America.

Consideration of Alternate Explanations: Multiple explanations or hypotheses should be considered before making conclusions about the causal relationship between any two items under investigation.	The association between Parkinson's disease and paraquat may be explained by socioeconomic factors; residences close to crop areas are generally of lower socioeconomic status. Exposure measurements were imprecise, relying on historical records and GIS tools, and a possible recall bias of past residence.
Specificity: An association between one causal factor and one outcome (but not other outcomes) can imply causality, but is not sufficient cause.	Chronic effects of paraquat also include respiratory health effects, such as, chronic obstructive lung disease.
Plausibility: The association agrees with currently accepted understanding of pathological processes; however, new research findings must be considered.	Paraquat was found to interfere with dopaminergic systems, known to be compromised in Parkinson's disease.
Experiment: The condition can be altered (prevented or ameliorated) by an appropriate experimental regimen, but may not be feasible or ethical to do so.	Ethical concerns would prevent an experimental study. However, since it is not yet completely banned in Canada, an intervention program which provides a safe substitute for paraquat may be possible.
Coherence: The association should be compatible with existing theory and knowledge. However, research that disagrees with established theory and knowledge are not automatically false.	The association of orchard work and pesticides with Parkinson's disease has been observed in many epidemiological studies, but it was problematic to identify the specific pesticide(s) involved.

Acknowledgements

Helen Ward would like to thank Mona Shum, Sarah Henderson, and Tom Kosatsky for their valuable input and review of the draft document.

References

1. Rothman KJ. Epidemiology, an Introduction. Oxford, UK: Oxford University Press; 2002.
2. Aschengrau A, Seage III GR. Essentials of epidemiology in public health. Sudbury, MA: Jones & Bartlett Learning; 2003.
3. Friis RH, Sellers T. Epidemiology for public health practice. 4th ed. Sudbury, MA: Jones & Bartlett Learning; 2009. Available from: http://www.jblearning.com/catalog/9781449613617/.
4. Elwood M. Critical appraisal of epidemiological studies and clinical trials. 3rd ed. Oxford, UK: Oxford University Press; 2007.

Production of this document was made possible through a financial contribution from the Public Health Agency of Canada.