

Chapter 2 — Methods

This project began in February 2003 with meetings between Dr. Cole and Dr. Sanborn to hire research assistants and approach physicians with experience and training in environmental health to act as reviewers. The project team was completed in mid-March, and since then, project group meetings have been held twice monthly by teleconference, with project decisions made by e-mail between teleconferences. In March 2003, discussions began regarding the general framework and direction for proceeding with the literature review. We agreed to begin by examining review articles to gauge the state of the current literature and identify any gaps that we would like to address in the review project. Details of the inclusion and exclusion criteria were also decided on at this time.

We decided to include all peer-reviewed studies published between 1992 and 2003 that investigated the human health effects of pesticides. These studies had to be systematic in their approach, and be written in English, French, Spanish, or Portuguese. The searches were done using the PreMedline, Medline, CancerLit, and LILACS (Spanish-language) databases. This review excluded the organochlorine literature, as most of these chemicals are no longer in use as pesticides in Canada and have been reclassified as persistent organic pollutants.

Phase 1 — Review of review papers

Selection of health effects for study

In the search for review papers, the inclusion criteria were expanded to include reviews conducted between 1990 and 2003. The initial search, using the term “pesticides,” yielded 12,061 papers. For the second selection stage, the term “pesticides” was combined with “systematic review,” “meta-analysis,” and “review” to select those studies that apply a systematic approach. When the papers were limited to review articles only, 1684 articles were found. However, these papers had a number of limitations, and although many were categorized as reviews, most did not describe a systematic approach. In addition, a number of the papers were primarily reviews dealing with organochlorines that had been picked up in the search because they were classified broadly as dealing with pesticides. As a result, many of the papers did not meet our inclusion criteria. From this list of 1684 articles, 49 relevant review papers were selected. Two members of the pesticide review group evaluated these review paper abstracts to determine their relevance to the current project. Of the 49 identified reviews, 30 were selected for further assessment using the quality and relevance criteria.

The review articles were collected and categorized according to health effect. We decided to organize the papers according to health effect rather than specific pesticide exposure, since most of the human health effect literature considers people exposed to cumulative or aggregate pesticide mixtures. To ensure that all possible health effects were included, a comprehensive search of medical subject headings (MeSH) in Medline was done early in the review process. This ensured that all relevant terms were included in the search process, and minimized the exclusion of relevant studies.

The health effect categories included in Phase 1 included: cancer, genotoxic, immunologic, neurotoxic, reproductive, and other. Four members of the group reviewed these review articles: two focused on the cancer papers and two reviewed the non-cancer papers. To standardize this process all reviewers used a common assessment tool in an effort to ensure consistent application of the evaluation criteria. After this step, a full project team meeting was held in Toronto on May

5, 2003 to introduce project members, conduct in-depth assessments of the review papers, and formulate the next phase of the project. This process allowed the group to determine gaps in the literature, and decide on the focus and limits of the review. The group decided to apply the following guidelines in choosing the health effects for review:

1. For the health effect, absence of a current, methodologically excellent review paper;
2. For cancers, data on incidence, premature loss of life, and increasing incidence were used to select specific tumour types for review; as well, a chosen cancer had to account for a substantial portion of total cancer incidence.

Using these guidelines, the following health effects were chosen for review, and categorized broadly as “cancer” or “non-cancer”:

A) Cancer	B) Non-Cancer
1. Lung	1. Reproductive effects
2. Breast	2. Genotoxic/immunotoxic
3. Colorectal	3. Dermatologic
4. Pancreas	4. Neurotoxic
5. Non-Hodgkin’s lymphoma	
6. Leukemia	
7. Brain	
8. Prostate	
9. Stomach	
10. Ovary	
11. Kidney	
12. Testicular	

Over the next month, work proceeded on the quality assessment tool for reviewing the primary studies. This tool, referred to as the data extraction form, was designed through extensive consultation with Andrea Furlan, the Evidence Based Practice Coordinator at the Institute for Work and Health. In addition, all members of the pesticide group took part in a pilot exercise, using the data extraction tool to assess the same article. All ratings of the article used in the pilot exercise were within one point of each other on the global rating scale. This exercise also led to some minor changes that made the form easier to use for the primary studies.

Phase 2 — Assessment of Primary Studies

During June and July 2003, relevant primary studies for each of the health effects were retrieved using a number of search strategies that incorporated the above criteria. The main strategy used in the review article search was repeated with modifications for each health effect using relevant MeSH terms. All searches began by using the term “pesticides,” limiting the search to human studies published between 1992 and 2003. For the search concerning non-Hodgkin’s lymphoma, the following terms were also included to capture all relevant articles: “lymphoma, non-Hodgkin,” “non Hodgkin lymphoma,” “NHL,” and “lymphoma.” A list of abstracts was produced from each search and distributed to the appropriate reviewers.

Of the Spanish papers, 723 abstracts were identified from the LILACS database and 79 of these selected on the basis of relevance. Based on the abstracts, 21 papers were selected for assessment, eight of which met the quality criteria and were included in the review. Two more Spanish papers were selected via Medline for assessment and were included.

The teams read the abstracts and selected those articles that met the inclusion criteria. Case reports and small case series were excluded. When articles lacked abstracts or contained too little information on which to make a selection, the original primary studies were obtained for evaluation. The same reviewer teams who read the review articles for a health effect also assessed the primary studies, as they were already familiar with the background literature. In addition, three new reviewers joined the group at this point: two assisted in reviewing the neurotoxicology papers and one reviewed the reproductive papers. Disagreements between reviewers concerning selection of articles for inclusion were resolved by discussion and input from a third reviewer. After the abstract selections were agreed on, the primary studies were retrieved from hospital and university libraries, the Internet, and, in some cases, ordered from other external sources through the library document delivery service. A total of 109 cancer papers and 156 non-cancer papers met the criteria for review and were distributed to the reviewing teams.

Over a three-month period, the primary articles were assessed using the data extraction form (Appendix 1). This form contains detailed questions regarding the study design, exposure assessment measures, outcome assessment measures, analysis, and methodological quality assessment. These categories were reviewed and discussed at group meetings in July, August, and September to ensure consistency within the group in the approach to evaluating these study characteristics. After a detailed assessment of methodology, each reviewer ranked each paper according to a global assessment scale on a range from 1–7, where 1 represented a paper with clearly unacceptable methodology, 4 an adequate study, and 7 an excellent one. This global assessment scale covered all aspects of study methodology, and was used as a general guideline to decide which studies would be included in the summary tables and report. All studies ranked below a global score of 4 were excluded, as these were considered of insufficient methodological quality to provide reliable data.

The reviewer pairs met to resolve disagreements about global ratings (defined as a difference of two or more points) by first discussing elements of their assessment of the study. If the difference in ratings could not be resolved, a third reviewer acted as a tiebreaker to decide whether to include the paper.

In August, the project group presented the project methodology and preliminary results concerning two health effects to the University of Toronto community medicine residents, and received helpful critique and suggestions.

Phase 3 — Analysis, external review, and report

As assessments of the primary articles were completed, summary tables were made showing the studies included for each health effect category. These tables include information on the study population, design, pesticide type, exposure and outcome measurements, covariates, analysis, and specific measures of association and values. The full review group met November 6–7 to discuss these tables and the written summaries for each health effect category. At this time, the results of the papers were discussed in detail, questions about methodology resolved, and controversial papers discussed in detail. Some methodological issues were common across all health effects, and some, such as the use of measures of incidence versus mortality, were specific to cancer. At these meetings we also developed guidelines for writing chapters for the final report.

The report was written over the following three months, reviewing and editing being done by peer and expert reviewers.

Sources of bias

The group also discussed potential sources of bias in the selection of articles for the review. One of these involved potential bias in the search strategy. The search strategies look for relevant terms in the title or abstract of the primary study. However, if the selected term is not mentioned in these sections, this type of search will pass over the study. Another possibility is that those papers showing positive associations are more likely to mention these terms in the abstract and thus be collected in the search. This is particularly true in the case of cohort studies that examine a wide range of health outcomes. Also, a bias may exist toward publication of papers with positive results. To expand our search for articles we examined the reference lists of each paper and added relevant articles missed by the initial search. We also consulted experts on specific health effects at several stages of the review to solicit comments on the selection of articles and analysis, and reduce the possibility of missing important studies.

Another concern was that some of the cancer and reproductive studies reported the outcomes for many health effects and levels of exposure. As a result, there was sometimes a lack of analysis specific to each health effect; more information was included when results were statistically significant. The large number of statistically significant results in these papers made it difficult to develop summary tables for them.

Two studies revealed that the authors had paid to have them published. Also, two studies were funded by pesticide-producing corporations, introducing the potential for a bias toward publication of negative findings.

Types of studies and methodological issues

The papers reviewed were cohort, case-control, or ecological studies. There are several methodological issues specific to each of these designs.

Cohort studies

There are several limitations and advantages innate to the design of cohort studies. One advantage of cohort studies is the large number of subjects that can be included, which increases their ability to find elevated rates of rare illnesses. However, when large numbers of subjects are studied, it is usually impossible to attain detailed pesticide exposure histories. For instance, instead of measurements of direct exposure and specific pesticide use, surrogate measurements are often used to describe exposure. These could be indirect measurements such as the job description of an occupation involving high exposure to pesticides, length of time employed in an exposed environment, length of membership in a union of occupationally exposed workers, or location of residence or workplace. For farmers, exposure is often estimated by type of crop grown, number of acres farmed, or expenditure on pesticides. Usually specific pesticides are not identified and sometimes results are confounded because the study dates back to an exposure period when organochlorines or other chemicals now banned were still in use. For example, some older studies of 2,4-D occurred during the period when this chemical was heavily contaminated with dioxins. Some studies use follow-up times that are too short to show up higher incidence rates when there is a long latency period between exposure and illness onset (some latency periods are as long as 30 years). In addition, sometimes subjects are lost to follow-

up because of migration. Often standardized mortality rates (SMR) are used instead of standardized incidence rates (SIR), which underestimates the actual occurrence of curable and treatable cancers. Mortality information is usually derived from death certificates, and these are not always accurate. For instance, death certificates do not always describe subtypes of cancers; when pesticide exposure causes one subtype and not another, the lack of subtyping could mask the true association between an exposure and the cancer. A histological diagnosis, the most accurate way to diagnose cancer, is not always available, and this lack makes a study less reliable. Important covariates such as smoking, family history, and race are usually not included in cohort studies because of the large numbers of subjects involved.

Cohort studies are confounded by the “healthy worker effect.” Ill people are less likely to hold down jobs or are more likely to be absent when the study is conducted. Hence, they will not be included in a group of workers studied, and the group will be pre-selected for healthy people. Also groups that possibly experience higher pesticide exposure, such as farmers, gardeners, and foresters, are known to be healthier than other workers in the population because they smoke less, exercise more, and are exposed to less air pollution. This makes them a pre-selected group within the larger working population. Another major limitation of these cohort studies and other studies is their lack of data on women.

Case-control studies

There are major limitations of case-control studies, the main one being that of recall bias: people who agree to participate may comprise a preselected group that is more likely to remember or exaggerate exposures in an effort to explain their own illness or that of a loved one. Some authors attempt to minimize recall bias by using a control group derived from hospital patients with other illnesses; however, these patients may also have illnesses related to exposures and are therefore not always valid controls.

Response rates vary with case-control studies, and low rates may introduce bias. In addition, random-digit dialing as a means of finding controls, can lead to controls with higher socioeconomic backgrounds that tend to be healthier than the general population, making them less appropriate controls. Another possible source of bias in case-control studies is a lack of interviewee blinding and data quality control. Also sometimes the case involved has died so a proxy respondent (usually a family member) must give information on exposure and other matters. Often the information available from these people is less accurate. However, because case-control studies offer the ability to interview or use questionnaires to obtain detailed exposure histories, details of covariates (such as smoking, family history, and race) can be determined.

Ecological Studies

Ecological studies consider groups rather than individuals as the unit of analysis. For example, such a study in the pesticides literature may compare mortality rates from national cancer registries to sales of pesticides by region. Others link disease cases from other national registries to employment information from national censuses to assess the relationship between pesticide exposure and adverse health effects. Ecological studies can be useful for detecting associations between exposure distributions and cases of disease, and are typically more easily and inexpensively conducted than are other study designs. However, because the data are measurements averaged over many individuals, the degree of association does not reflect individual-level associations. The failure of the effect estimate at the ecological level to reflect

the effect at the individual level is the primary limitation of this study design. Nevertheless, these studies are useful for generating hypotheses and suggesting future directions for observational studies.

The following Chapters (3–9) detail the findings for each health effect, and are accompanied by reference lists and summary tables.