RESEARCH TRENDS IN EPIDEMIOLOGY DOCTORAL THESES EXAMINING ILLNESS OUTCOMES, 1982-1992

by

Deborah L. Morin

Department of Epidemiology and Biostatistics

Submitted in partial fulfillment
of the requirements for the degree of
Master of Science

Faculty of Graduate Studies
The University of Western Ontario
London, Ontario
April 1995

ISBN 0-612-03268-X

Canada

Name DEBORAH L. MORIN

Dissertation Abstracts International is arranged by broad, general subject categories. Please select the one subject which most nearly describes the content of your dissertation. Enter the corresponding four-digit code in the spaces provided.

HEALTH SCIENCES, PUBLIC HEAL. I
SUBJECT TERM

0573
SUBJECT CODE

U·M·I

Subject Categories

THE HUMANITIES AND SOCIAL SCIENCES

COMMUNICATIONS AND THE ARTS
Architecture 0729
Art History 0377
Cinema 0900
Dance 0378
Fine Arts 0357
Information Science 0723
Journalism 0391
Library Science 0399
Mass Communications 0708
Music 0413
Speech Communication 0459
Theater 0465

EDUCATION
General 0515
Administration 0514
Adult and Continuing 0516
Agricultural 0517
Art 0273
Bilingual and Multicultural 0282
Business 0688
Community College 0275
Curriculum and Instruction 0727
Early Childhood 0518
Elementary 0524
Finance 0277
Guidance and Counseling 0519
Health 0680
Higher 0745
History of 0520
Home Economics 0278
Industrial 0521
Language and Literature 0279
Mathematics 0280
Music 0522
Philosophy of 0998
Physical 0523
Psychology 0525
Reading 0535
Religious 0527
Sciences 0714
Secondary 0533
Social Sciences 0534
Sociology of 0340
Special 0529
Teacher Training 0530
Technology 0710
Tests and Measurements 0288
Vocational 0747

LANGUAGE, LITERATURE AND LINGUISTICS
Language
- General 0679
- Ancient 0289
- Linguistics 0290
- Modern 0291

Literature
- General 0401
- Classical 0294
- Comparative 0295
- Medieval 0297
- Modern 0298
- African 0316
- American 0591
- Asian 0305
- Canadian (English) 0352
- Canadian (French) 0355
- English 0593
- Germanic 0311
- Latin American 0312
- Middle Eastern 0315
- Romance 0313
- Slavic and East European 0314

PHILOSOPHY, RELIGION AND THEOLOGY
Philosophy 0422
Religion
- General 0318
- Biblical Studies 0321
- Clergy 0319
- History of 0320
- Philosophy of 0322

Theology 0469

SOCIAL SCIENCES
American Studies 0323
Anthropology
- Archaeology 0324
- Cultural 0326
- Physical 0327

Business Administration
- General 0310
- Accounting 0272
- Banking 0770
- Management 0454
- Marketing 0338

Canadian Studies 0385
Economics
- General 0501
- Agricultural 0503
- Commerce-Business 0505
- Finance 0508
- History 0509
- Labor 0510
- Theory 0511

Folklore 0358
Geography 0366
Gerontology 0351
History
- General 0578
- Ancient 0579
- Medieval 0581
- Modern 0582
- Black 0328
- African 0331
- Asia, Australia and Oceania 0332
- Canadian 0334
- European 0335
- Latin American 0336
- Middle Eastern 0333
- United States 0337

History of Science 0585
Law 0398
Political Science
- General 0615
- International Law and Relations 0616
- Public Administration 0617

Recreation 0814
Social Work 0452
Sociology
- General 0626
- Criminology and Penology 0627
- Demography 0938
- Ethnic and Racial Studies 0631
- Individual and Family Studies 0628
- Industrial and Labor Relations 0629
- Public and Social Welfare 0630
- Social Structure and Development 0700
- Theory and Methods 0344

Transportation 0709
Urban and Regional Planning 0999
Women's Studies 0453

THE SCIENCES AND ENGINEERING

BIOLOGICAL SCIENCES
Agriculture
- General 0473
- Agronomy 0285
- Animal Culture and Nutrition 0475
- Animal Pathology 0476
- Food Science and Technology 0359
- Forestry and Wildlife 0478
- Plant Culture 0479
- Plant Pathology 0480
- Plant Physiology 0817
- Range Management 0777
- Wood Technology 0746

Biology
- General 0306
- Anatomy 0287
- Biostatistics 0308
- Botany 0309
- Cell 0379
- Ecology 0329
- Entomology 0353
- Genetics 0369
- Limnology 0793
- Microbiology 0410
- Molecular 0307
- Neuroscience 0317
- Oceanography 0416
- Physiology 0433
- Radiation 0821
- Veterinary Science 0778
- Zoology 0472

Biophysics
- General 0786
- Medical 0760

EARTH SCIENCES
Biogeochemistry 0425
Geochemistry 0996
Geodesy 0370
Geology 0372
Geophysics 0373
Hydrology 0388
Mineralogy 0411
Paleobotany 0345
Paleoecology 0426
Paleontology 0418
Paleozoology 0985
Palynology 0427
Physical Geography 0368
Physical Oceanography 0415

HEALTH AND ENVIRONMENTAL SCIENCES
Environmental Sciences 0768
Health Sciences
- General 0566
- Audiology 0300
- Chemotherapy 0992
- Dentistry 0567
- Education 0350
- Hospital Management 0769
- Human Development 0758
- Immunology 0982
- Medicine and Surgery 0564
- Mental Health 0347
- Nursing 0569
- Nutrition 0570
- Obstetrics and Gynecology 0380
- Occupational Health and Therapy 0354
- Ophthalmology 0381
- Pathology 0571
- Pharmacology 0419
- Pharmacy 0572
- Physical Therapy 0382
- Public Health 0573
- Radiology 0574
- Recreation 0575
- Speech Pathology 0460
- Toxicology 0383

Home Economics 0386

PHYSICAL SCIENCES
Pure Sciences
Chemistry
- General 0485
- Agricultural 0749
- Analytical 0486
- Biochemistry 0487
- Inorganic 0488
- Nuclear 0738
- Organic 0490
- Pharmaceutical 0491
- Physical 0494
- Polymer 0495
- Radiation 0754

Mathematics 0405
Physics
- General 0605
- Acoustics 0986
- Astronomy and Astrophysics 0606
- Atmospheric Science 0608
- Atomic 0748
- Electronics and Electricity 0607
- Elementary Particles and High Energy 0798
- Fluid and Plasma 0759
- Molecular 0609
- Nuclear 0610
- Optics 0752
- Radiation 0756
- Solid State 0611

Statistics 0463

Applied Sciences
Applied Mechanics 0346
Computer Science 0984
Engineering
- General 0537
- Aerospace 0538
- Agricultural 0539
- Automotive 0540
- Biomedical 0541
- Chemical 0542
- Civil 0543
- Electronics and Electrical 0544
- Heat and Thermodynamics 0348
- Hydraulic 0545
- Industrial 0546
- Marine 0547
- Materials Science 0794
- Mechanical 0548
- Metallurgy 0743
- Mining 0551
- Nuclear 0552
- Packaging 0549
- Petroleum 0765
- Sanitary and Municipal 0554
- System Science 0790

Geotechnology 0428
Operations Research 0796
Plastics Technology 0795
Textile Technology 0994

PSYCHOLOGY
General 0621
Behavioral 0384
Clinical 0622
Developmental 0620
Experimental 0623
Industrial 0624
Personality 0625
Physiological 0989
Psychobiology 0349
Psychometrics 0632
Social 0451

THE UNIVERSITY OF WESTERN ONTARIO
FACULTY OF GRADUATE STUDIES

CERTIFICATE OF EXAMINATION

Chief Advisor

Examining Board

Advisory Committee

The thesis by
Deborah L. Morin

entitled
Research Trends in Epidemiology Doctoral Theses
Examining Illness Outcomes, 1982-1992

is accepted in partial fulfilment of the
requirements for the degree of
Master of Science

Date May 5, 1995

Chair of Examining Board

ABSTRACT

Objectives. To describe the relative frequencies with which study designs were used in epidemiology dissertations and to determine the research topic areas being addressed in these investigations. Secondary objectives included summarizing the research by setting, and type and number of study subjects. ***Methods***. A bibliometric investigation of epidemiologic research was carried out on 758 dissertation abstracts originating from four Canadian and twenty-one American universities providing doctoral education in epidemiology. Abstracts written between January 1982 and December 1992 were selected from within the *Dissertation Abstracts International* database, and were included in the study only when the content described epidemiologic research examining illness outcomes. Other areas of epidemiologic research were excluded. Study designs were classified using a taxonomy developed for this project, and research topic areas were collapsed into ICD-9 chapters. These variables were examined collectively and across four time periods to determine whether time trends were evident. ***Results***. Case-control studies were often used in epidemiology dissertations (32.7%). Intervention studies were infrequently used, accounting for roughly 5% of abstracts. Although the relative frequencies of study designs were found to vary across time, consistent trends were not evident. The research topic area of greatest activity was neoplasms (22.8%). As well, dissertation research described infectious and parasitic diseases (11.5%), maternal and infant health disorders (11.6%), diseases of the circulatory system (10.9%), mental disorders (7.7%) and injuries (6.9%). ***Conclusions***. Observational research methods were predominant in epidemiology dissertations. Of these studies, the case-control design was the most common. A diversity of research topic areas was addressed, however, diseases of primary concern to the population were often examined. To effectively communicate the methods and findings of epidemiology dissertations, it is recommended that graduates prepare a structured abstract. Using this source, future trends in epidemiologic research may be evaluated.

key words: epidemiology, epidemiologic methods, bibliometrics

AKNOWLEDGEMENTS

I would like to thank my advisors, Drs. Truls Østbye, James Robertson and Allan Donner for their guidance and encouragement throughout this project. I am grateful for their assistance. To my friends and family, I extend my thanks and appreciation for their constant support Finally, I would like to thank my parents Audrey and Dennis Morin, for their continuous encouragement and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1 - THE HISTORY OF EPIDEMIOLOGY

1.1 Introduction

From its Greek origins, epidemiology is literally translated as the study (*logos*) of that which is among (*epi*) the people (*demos*) (Detels 1991) More specifically, this discipline is defined as "the study of the distribution and determinants of health-related states or events in specified (human) populations, and the application of this study to control of health problems".[1]

Originally, epidemiology was used to describe the causes of and minimize the spread of infectious disease. However, with the persistent use of sanitary measures, the prevalence of infectious disease declined in the developed countries, resulting in a shift to chronic disease and injury Accordingly, the discipline of epidemiology evolved.

Regardless of the prevailing health problems however, the objectives of epidemiology remain constant: (1) to describe the health (and disease) status of the population; (2) to describe the etiology of disease, (3) to predict the frequency and distribution of disease within the population; and (4) to control the distribution of disease (Kleinbaum, Kupper and Morgenstern 1982).

1.2 History

The history of epidemiologic research, principles and methods has been thoroughly documented elsewhere (MacMahon and Pugh 1970; Lilienfeld and Lilienfeld 1977a, 1977b; Buck et al. (ed.) 1988; White 1992). For this reason, an abbreviated account will be described according to the development of four central ideas: (1) person, place and time; (2) measurement; (3) natural experiments; and (4) intervention studies (MacMahon and Pugh 1970).

1.2.1 Person, Place and Time

In an essay entitled *On Airs, Waters and Places* , the Greek physician Hippocrates described one of the fundamental principles of epidemiology: to describe disease by person (host), place (environment) and time (Buck et al 1988). The following passage is an illustration of this assertion:

> "...Whoever wishes to pursue properly the science of medicine must proceed thus. First he ought to consider what effects each season of the year can produce; for the seasons are not at all alike, but differ widely both in themselves and at their changes...He must also consider the properties of the waters. Therefore, on arrival at a town with which he is unfamiliar, a physician should examine its position with respect to the winds and to the rising of the sun. For a northern, a southern, an eastern, and a western aspect has each its own individual property. He must consider with the greatest care both these things and how the natives are off for water, whether they use marshy, soft waters, or such as are hard and come from rocky heights, or brackish and harsh. The soil too, whether bare and dry or wooded and watered, hollow and hot or high and cold. The mode of life also of the inhabitants that is pleasing to them, whether they are heavy drinkers, taking lunch and inactive, or athletic, industrious, eating much and drinking little."[2]

Today, the insightfulness of Hippocrates is referred to as "descriptive" epidemiology; a practice which is essential for understanding the determinants of disease (Mausner and Kramer 1985, Last 1988). It was not until 1662 however, that John Graunt numerically described disease by person, place and time (MacMahon and Pugh 1970; White, 1992).

1.2.2 Measurement

Graunt (1620-1674) compiled and analyzed the London Bills of Mortality from 1603 to 1662. In his well-known report, *Natural and Political Observations on the Bills of Mortality*, Graunt numerically described variations in mortality by gender, season, and residence (MacMahon and Pugh 1970; Detels and Breslow 1991; White 1992). While he is often credited with being the founder of vital statistics (Last 1986), William Petty, Pierre-Charles Alexandre Louis and William Farr are among other influential contributors (Lilienfeld and Lilienfeld 1977a; White 1992).

William Petty's (1623-1687) contributions included work in vital statistics and economics. He published numerous books; the most distinguished entitled, *Political Arithmetik* (Last 1988; White 1992). In addition, Petty proposed that a central statistical agency be created in which information be collected on vital statistics, disease, occupation, education, income, and housing (White 1992).

Pierre-Charles Alexandre Louis (1787-1872), a French physician and mathematician is recognized for his clinical description of typhoid fever (Lilienfeld and Lilienfeld 1977a), his precision and detail in the observation and measurement of disease (Lilienfeld and Lilienfeld 1977a; Last, 1988), and for his development of the "numerical method" (Lilienfeld and Lilienfeld 1977a). As well, he is known to have instructed William Farr, a brilliant medical statistician (Lilienfeld and Lilienfeld 1977a; White 1992).

During his appointment as Compiler of Statistical Abstracts at the General Register Office in England, Farr (1807-1883) developed and implemented a system of vital statistics (Lilienfeld and Lilienfeld 1977b; Last 1988; White 1992), and, established many of the current methods used to describe and explain the etiology of disease (Lilienfeld and Lilienfeld 1977b; Last 1986). In addition to other concepts, he defined the standardized mortality ratio, the notion of person-years, and the relationship between incidence and prevalence (Last 1986). Farr is credited also, with developing a disease classification system, preserved in the most current revision of the International Statistical Classification of Diseases (ICD-10) (WHO 1992).

1.2.3 Natural Experiments

Arguably the most well-known "natural" or "observational" experiment was carried out by John Snow in London during a cholera outbreak in 1854. Briefly, natural experiments are investigations where study subjects are similar across a number of characteristics yet differ on the basis of exposure to a suspected "risk" factor (Mausner and Kramer 1985; Last 1988). As a result, the relationship between exposure and disease is evaluated by comparing the frequency of disease between groups.

Supplied with data from William Farr (Lilienfeld and Lilienfeld 1977b), Snow observed that cholera mortality was greatest in city districts receiving water from the Southwark and Vauxhall Company compared to areas serviced by the Lambeth Company. Each company drew its water supply from the Thames River; however, the former used as its source, a highly contaminated area, whereas the Lambeth Company used as its source, a less polluted section of the river. Snow speculated that contaminated water was the mode of disease transmission, and provided evidence to support the hypothesis by comparing mortality rates in residential areas serviced by both water companies. The mortality rate was found to be almost nine times greater in homes serviced by the Southwark and Vauxhall Company (4.5 per 1000) compared to homes receiving water from the Lambeth Company (0.5 per 1000) (MacMahon and Pugh 1970; Lilienfeld and Lilienfeld 1977b, Kelsey, Thompson and Evans 1986).

1.2.4 Intervention Studies

In contrast to natural experiments, "intervention" or "experimental" studies evaluate a cause-effect relationship by manipulating a "treatment" factor in the study population (Last 1988). As a result, the "treatment" variable is under the direct control of the investigator and is applied to study groups. An historical example includes the use of fresh fruit in the treatment of scurvy by James Lind in 1747 (MacMahon and Pugh 1970; Last 1986). In this trial, Lind evaluated the efficacy of nutritional therapy by comparing the rates of disease between sailors treated with and without fruit juice.

A more rigorous variation of the intervention study, namely, the randomized clinical trial (RCT), was developed and utilized in 1938 in trials of vaccines for the common cold (Last 1986). In these studies, the vaccine was *randomly* allocated to experimental and control groups, thus reducing the potential for biased findings. After the Second World War, the British Medical Research Council carried out numerous randomized intervention studies (Last 1986), after which the use of the RCT increased

Other noteworthy intervention studies include the fluoridation of drinking water in the prevention of dental caries; and, the Multiple Risk Factor Intervention Trial (MRFIT), a study designed to reduce mortality from coronary heart disease in high-risk subjects (Last 1986; Rothman 1986; Mausner and Kramer 1986).

1.3 The Institutionalization of Epidemiology

Despite this reported research activity, epidemiology was not "institutionalized" as a distinct discipline until this century, with the appointment of Wade Hampton Frost in 1921 as the first full Professor of Epidemiology at the Johns Hopkins University (Fee 1987). The number of faculty members has since increased in the United States to approximately 711 as estimated in a 1992 study (Schoenbach et al. 1994).

Currently, within Canada, the United States, and the United Kingdom, graduate education in epidemiology is provided by more than 80 universities located within schools of public health, medicine and veterinary medicine (Bernier and Mason 1991), and applied training is offered by the Epidemic Intelligence Service Program (EIS) at the Centers for Disease Control. At present, epidemiologic research is actively encouraged, communicated and disseminated by twenty-eight English-speaking organizations (e.g. the International Epidemiologic Association, American College of Epidemiology, Canadian Society of Epidemiology and Biostatistics etc.) and by numerous peer-reviewed journals (e.g. American Journal of Epidemiology, International Journal of Epidemiology, Annals of Epidemiology etc.) (Bernier and Mason, 1991).

1.3.1 The Epidemiologist

In 1942, readers of the *American Journal of Public Health* were invited to submit replies to the question, "What and who is an epidemiologist?"[3] Responses varied but included definitions of epidemiology, descriptions of the practice of epidemiology (e.g. "shoe leather", laboratory-based epidemiology etc.), and the training and experience required of an epidemiologist (e.g. medical training, statistical understanding etc.). More recently, several attempts at enumerating the work force and describing the epidemiologist have been made (Detels 1979; Williams et al. 1988; Gunn et al. 1989; Woernle 1991; Boss

and Foster 1994; Schoenbach et al. 1994). However, enumeration remains difficult since epidemiologists are not licensed, certified or uniformly defined (Williams et al. 1988, Gunn et al. 1989; Woernle 1991, Boss and Foster 1994).

In 1979, Detels assessed the present and ten-year projected need for epidemiologists by polling numerous research agencies, institutes and departments. Among respondents, schools of medicine and agencies within the National Institutes of Health reported the greatest need for epidemiologists. Williams and his co-workers (1988) surveyed the membership of epidemiology associations, and contacted training programs to enumerate the active work force in 1985. The authors reported 4,600 active epidemiologists in the United States, a ratio of 1.94 epidemiologists per 100,000 population. Using population projections, and assuming that this ratio remains constant, the authors estimated the need for epidemiologists to increase by 10%-30% in the year 2010.

Of the active work force enumerated in 1985, 54% were physicians (41% with MPH, 10% with doctoral degree in public health field), while 9% and 37% of epidemiologists were trained at the doctoral- and masters-level respectively. Surveys of state health department epidemiologists have confirmed that most individuals are medically trained (Gunn et al. 1989; Woernle 1991), yet, despite this evidence, the number of physician-trained epidemiologists has been declining and remains an area for recruitment (Williams 1989; Greenberg 1990).

Findings from a national survey carried out in 1983 characterized the state health department epidemiologist as male, with an average age of 41 years, and 9 years of work experience (Gunn et al. 1989). In a similar study of non-infectious disease epidemiologists, more than half of the 260 respondents were male, and over 90% were white (Boss and Foster 1994).

Thacker and his associates collected survey information to assist administrators in the recruitment, preparation, and retention of minority officers within the Epidemic Intelligence Service (Thacker et al. 1992). Compared to non-minority respondents, minority graduates were more likely to be female, and less likely to work within academic institutions, or be appointed to academic positions. When compared by year of

graduation, this last difference was more pronounced in 1980-1989 graduates than in the 1970-1979 cohort. The ethnic distribution of epidemiologists was examined further in a survey of 56 American epidemiology programs. Of the 711 faculty members, 6% were minorities. Of 2,142 students, 5% were Black, 4% were Asian/Pacific Islanders, 4% were Hispanic, and none were Native Americans. In response to this underrepresentation, the American College of Epidemiology is presently developing recommendations to increase minority recruitment in the profession (Schoenbach et al. 1994).

In part, the practice and future of epidemiology will be influenced by the characteristics of the epidemiologist (Terris 1992). Recognizing this fact, Greenberg (1990) has previously recommended that a census of epidemiologists be carried out to better understand the dynamics of the discipline. Still, there are additional factors that will likewise influence epidemiology. These factors include the economic and social climate of society; and the prevailing health problems of the day (Gordis 1988; Williams et al. 1989; Terris 1992).

---

[1] Last JM, editor. A Dictionary of Epidemiology, 2nd ed. New York: Oxford University Press, Inc., 1988: 42.

[2] Hippocrates. Airs, waters, places. WHS Jones eds. Harvard University Press. Cambridge. 1948. Cited in Buck C, Llopis A, Nájera E, Terris M, editors. The Challenge of Epidemiology. Issues and Selected Readings. Washington, D.C.: Pan American Health Organization. 1988: 18-19.

[3] What and who is an epidemiologist? [editorial]. *Am J Public Health* 1941; 32: 414-5.

CHAPTER 2 - REVIEW OF LITERATURE

2.1 Introduction

Of the numerous factors that will influence epidemiology (Gordis 1988; Williams et al. 1989; Terris 1992), perhaps the most significant is the health status of the population. It is important, therefore, to examine the present state of the discipline to determine whether current health issues are being adequately addressed, and to identify trends that may have implications on future research and training. Central characteristics that describe the state of epidemiology include the research areas investigated and the study designs used.

2.2 Research Areas

Most certainly, the prevailing health problems of society have influenced the direction and scope of epidemiologic research. In the developed countries, the prevalence of infectious disease has declined significantly over the years with the establishment of public health interventions (Detels and Breslow 1991). At present, the leading causes of death are chronic, degenerative diseases, such as coronary heart disease and cancer (National Center for Health Statistics 1993). This shift from infectious to chronic disease has been chronicled in the *American Journal of Epidemiology* from its inception in 1921 to 1978 (LaPorte and Cresanta 1980), and similar findings have been reported in a subsequent review of the journal from 1965 to 1989 (Sartwell and Stark 1991). In spite of this overall trend, recent inquiries reflect the current health crisis resulting from the emergence of AIDS. Although not exclusive to epidemiology, the pertinent literature indexed by the National Library of Medicine has grown from 28 records in 1982 to a cumulative total of 29,077 in 1990 (Pratt 1992).

To determine the level of epidemiologic activity within various medical specialties, Dannenberg (1985) carried out a citation analysis using journal articles published between 1974 and 1982. During this time, more than 15,000 citations to and by the *American Journal of Epidemiology* were identified, and each was classified according to the subject

area defined by the referencing or referenced journal. General and internal medicine and public health/epidemiology journals were the two categories most often cited to and by the journal. Other specialty journals included immunology, cancer, pediatrics, cardiovascular system, statistics, and, obstetrics and gynecology. Very few citations to and by the *American Journal of Epidemiology* were observed for allergy, anesthesiology, geriatrics, rheumatology and surgery, suggesting a deficiency of epidemiologic activity within these fields (Dannenberg 1985).

More recently, the subject matter of articles appearing in the Swiss *Journal of Social and Preventive Medicine* from 1980 to 1990 was reviewed and classified. Subject categories were defined according to disease area and common research activity (e.g. health services research, elderly etc.). Of the categories describing various diseases, heart disease, cancer and substance addiction were most common. Low levels of research activity were reported for accidents and maternal and child health problems. Overall, AIDS research was poorly represented in the journal, yet the proportion of related articles increased during the second half of the decade. As well, the proportion of articles devoted to heart disease and cancer was greatest during 1986 to 1990 (Schubert-Subbarathnam and Rougemont 1992).

Epidemiologists working within state health departments in the United States were surveyed to assess their level of participation within specific program areas (Gunn et al. 1989). More than 75% of epidemiologists worked within general epidemiology and communicable disease control, and greater than half (50%-75%) reported activity in cancer control, hypertension and chronic diseases. Less than half of the epidemiologists (25%-50%) participated in programs concerned with maternal and child health, injury control, diabetes, and smoking prevention; and, less than 25% participated in alcohol and drug abuse programs. In a similar study of southern state epidemiologists, the top three areas of activity were infectious disease (39%), environmental epidemiology (21%), and general epidemiology (15%). AIDS represented 33% of the infectious disease activity. Collectively, only 13% of respondents could be accounted for with respect to injuries, occupational health, and maternal and child health (Woernle 1991).

The level and type of epidemiologic activity performed by non-infectious disease epidemiologists was characterized according to risk factors (i.e. environment, occupation, nutrition, tobacco, and substance abuse), diseases (i.e. diabetes, cancer and cardiovascular disease), and health conditions (i.e. injury, birth defects and other reproductive conditions). The greatest amount of activity was reported in environmental epidemiology (55%), followed by cancer (47%), and occupational epidemiology (40%). Nutrition (9%) and substance abuse (10%) represented the areas of least activity (Boss and Foster 1994). Considering current public health priorities (National Center for Health Statistics 1993), the results from these surveys suggest the need for increased epidemiologic activity. Furthermore, epidemiologists will have to address new priorities, as the population ages (e.g. dementia) and familiar health problems like tuberculosis re-emerge (National Center for Health Statistics 1993).

2.3 Study Designs Used In Epidemiology

In epidemiology, the use of a particular design is dependent upon a variety of factors (i.e. incidence of the study disease, availability of study information etc.). Nevertheless, a more rigorous design (e.g. RCT, prospective cohort) will provide the most convincing evidence to support or refute a relationship between an exposure and outcome. In a recent meta-analysis, the type of study design independently predicted the risk estimate for exposure to non-steroidal anti-inflammatory drugs and gastrointestinal disease. The relative risk was lowest in cohort studies (RR=2.0), and greatest in case-control studies (RR=4.4) with hospital controls (Bollini et al. 1992).

The study designs used regularly in epidemiology include the intervention, cohort, case-control, cross-sectional and ecological study. Each of these designs is described below and examples are provided to illustrate its use. A more detailed description of each design, in terms of its strengths and weaknesses can be found in MacMahon and Pugh (1970); Kleinbaum, Kupper and Morgenstern (1982); Kelsey, Thompson and Evans (1986) and Rothman (1986).

2.3.1 Intervention Study

The purpose of an intervention or experimental study is to evaluate the efficacy of a treatment or therapy in slowing (e.g. therapeutic trial) or preventing (e.g. prophylactic trial) a disease process (Mausner and Kramer 1985). Subjects may be selected according to disease or risk status and are assigned to groups receiving the treatment or control factor (e.g. placebo, standard therapy etc.). The study subjects are followed over time and the effectiveness of the treatment is evaluated by comparing the subsequent rates of disease between groups.

When subjects are selected according to their disease status and the study factor is randomly allocated to study groups, the intervention study is known specifically as a *randomized clinical trial* (RCT). One such example of a RCT is the NASCET investigation of surgical versus standard medical treatment in patients at high risk for stroke. In this study, investigators demonstrated a beneficial treatment effect (NASCET 1991).

A second, well-known example of an intervention study is the Multiple Risk Factor Intervention Trial (MRFIT), a study designed to evaluate the efficacy of multiple risk factor modification on subsequent reduction of mortality from coronary heart disease (Rothman 1986; Mausner and Kramer 1985). In this particular study, however, subjects were selected on the basis of risk status and not for clinical presentation of the disease. As a result, this study is known specifically as a *field trial* (Rothman 1986).

A further type of experimental study is the *community intervention trial*. In this design a community or geopolitical division is the study unit in which the preventive or therapeutic factor is applied (Last 1988). In the late 1940's, the drinking water of Newburgh, New York was supplemented with fluoride, while the water supply remained untreated in the nearby community of Kingston. The effectiveness of this intervention was subsequently measured, by comparing the rates of dental caries between the treated and control communities (Last 1986).

2.3.2 Cohort Study

In a cohort study, subjects exposed to varying levels of a suspected risk factor are recruited and followed over time to identify deaths or incident cases of disease. This study is generally referred to as a *prospective* cohort study since subjects are observed forward from exposure to outcome. A well known example of a prospective cohort study is the Framingham Heart Study (Kelsey et al. 1986; Friedman 1987). In this project, residents of Framingham, Massachusetts were examined at baseline and followed up biennially for evidence of coronary heart disease and hypertension.

In some cases, data on individuals in the exposed and unexposed groups are collected from existing records. The current morbidity or mortality status is then identified for both groups. This particular design is often referred to as an *historical* or *retrospective* cohort study since the exposure factor is identified as it occurred in the past. Hoffman et al. (1982) used this particular design in their investigation of mortality in women treated for hyperthyroidism. Medical records were reviewed to establish the exposed (e.g. radioiodine treatment) and unexposed groups (e.g. surgical treatment), and, the cause of death for each subject was subsequently determined from vital records.

2.3.3 Case-Control Study

The case-control study is usually a *retrospective* study because diseased subjects are pursued backwards from outcome to exposure. Specifically, a group of individuals suffering from a particular disease (i.e. cases) is assembled. Likewise, a group of individuals free of the disease (i.e. controls) is constructed from a separate population. Thereafter, comparisons are made between the cases and controls with respect to current or previous exposure to risk factors suspected in causing the disease. Kantor and her colleagues (1984) used this design in their investigation of urinary tract infection and risk of bladder cancer. Cases were defined as newly diagnosed patients with carcinoma of the urinary bladder; controls were individuals randomly selected from the general population. Previous diagnosis of urinary tract infection was subsequently determined through a personal interview with each subject.

A variation of this design is the *nested case-control* study. In this design, both cases and controls are identified from within a cohort study. Kleinbaum and his co-workers (1982) classified this particular design as a hybrid study, since features from each of the cohort and case-control studies are preserved.

2.3.4 Cross-Sectional Study

In a cross-sectional study, subjects are (usually) selected at random from a single target population and exposure and disease status are measured simultaneously, and at one point in time. An example of this type of study is the measurement of maternal smoking status and low infant birth weight in a group of new mothers.

The cross-sectional design may be descriptive or analytic in nature, or may combine aspects from both approaches. A descriptive study provides information detailing the distribution of a disease by person, place and time. In comparison, an analytical study yields information regarding the presence or absence of, as well as the strength of association between study variables (Last 1986, Abramson 1991).

2.3.5 Ecological Study

The ecological study is commonly referred to as an aggregate study because the unit of analysis is a group rather than the individual. More specifically, this type of study is used to investigate and compare variations in the rates of disease between countries or between regions. An example is a study describing variations in rates of cardiovascular disease mortality among industrialized countries.

In addition, the ecological study may be used to investigate time trends. For example, changes in exposure status (e.g. trends in saccharin use) among groups of people may be correlated with observed changes in the rates of a particular disease (e.g. bladder cancer) (Mausner and Kramer 1986).

2.4 Tracking the Use of Study Designs

Several researchers have documented the frequency with which study designs have been reported in the general medical literature (Feinstein 1978; Cole 1979; Fletcher and Fletcher 1979; Bailar et al. 1984; Alvarez-Dardet et al. 1985). These endeavors were carried out to provide readers with an understanding of the editorial process, and to inform medical faculties of the type of instruction and preparation required to accurately interpret and evaluate the printed research (Feinstein 1978). Furthermore, classification was promoted in order to identify the use and underuse of rigorous study designs (Fletcher and Fletcher 1978; Bailar et al. 1984).

Feinstein (1978) classified the study designs reported in original articles appearing in both *The New England Journal of Medicine* and *The Lancet* between October of 1977 and March of 1978. Of 311 clinically relevant articles, 39% described cohort studies and 61% described cross-sectional studies. When defined more precisely, case-control studies accounted for 14% of cross-sectional studies, or 8% of the total sample, and controlled clinical trials represented 44% of cohort studies, or 17% of all articles. In a similar investigation, Bailar and his co-workers (1984) classified the study designs reported in 332 articles appearing in the *New England Journal of Medicine* during 1978 and 1979. 61% of these studies were longitudinal in design and 39% were cross-sectional. Specifically, 33% of articles were intervention studies, 19% were observational cohort studies, and 6% were case-control studies.

Fletcher and Fletcher (1979) classified the study designs described in 612 articles appearing within the *New England Journal of Medicine*, *The Lancet*, and the *Journal of the American Medical Association* from 1946 to 1976. Statistically significant trends were reported for three designs. From 1956 to 1976, the number of cross-sectional studies increased (25%-44%) as did the number of clinical trials (13%-21%). In contrast, the number of cohort studies decreased (59%-34%). When examining 155 articles published in 1976, 44% were cross-sectional studies, of which 8% were case-controls, 34% were cohort studies, and 5% were randomized clinical trials. In a subsequent review of the *New England Journal of Medicine*, researchers did not report significant differences in the frequencies of study designs. Collectively, case-control, cohort and randomized

clinical trials accounted for approximately 25% of articles reviewed in 1975 and 1984 (Alvarez-Dardet et al. 1985).

Contrary to the work of these authors, Cole was interested in documenting the perception of an increased use of the case-control study (Cole 1979). Articles published during 1956-1957 and 1976-1977 in *The Lancet*, *The New England Journal of Medicine*, the *American Journal of Epidemiology* and the *Journal of Chronic Diseases* were classified according to the study design used. Collectively, the proportion of case-control studies increased over time in the general medical literature, from 0.8% to 6%; and, as a percentage of epidemiologic studies, the use of the case-control design increased from 19% to 43%. In the *American Journal of Epidemiology*, the use of this study design was absent in papers appearing between 1956 and 1957; but this number increased to 17 of 201 (8.5%) articles published during 1976 and 1977. A similar increase was not observed for papers appearing in the *Journal of Chronic Diseases*. In addition to these investigations, study designs have been classified in specialty journals such as rheumatology (Ruiz et al. 1990), family medicine (Geyman and Berg 1984; Fromm and Snyder 1984; Marvel et al. 1991), veterinary medicine (Smith 1988); and in the Swiss *Journal of Social and Preventive Medicine* (Schubert-Subbarathnam and Rougemont 1992).

2.5 The Need for Additional Information

The medical literature indicates that the research community has responded to the dynamics of the population's health status (LaPorte and Cresanta 1985; Sartwell and Stark 1991; Pratt 1992). Yet, recent surveys suggest the need for further epidemiologic activity within areas of nutrition, maternal and child health, smoking prevention and substance abuse (Gunn et al. 1989; Woernle 1991; Boss and Foster 1994). Surely, epidemiologic activity will be directed by these current health issues, and will be influenced further as new health problems emerge.

An essential component of epidemiologic activity is the study design. In the general medical and specialty literature, the use of designs has been described (Feinstein 1978; Fletcher and Fletcher 1979; Cole 1979; Bailar et al. 1984; Alvarez-Dardet et al.

1984; Geyman and Berg 1984; Fromm and Snyder 1986; Smith 1988; Ruiz et al. 1990; Marvel et al. 1991). However, the classification systems employed within these studies varied among researchers and across medical disciplines. Furthermore, with the exception of Cole (197:), the use of study designs has not been described recently in journals established specifically for the dissemination of epidemiologic research. Certainly, further studies are required to provide information on current trends in the use of study designs.

One means of addressing this issue is to examine the distribution and content of doctoral dissertations in epidemiology. The doctorate is recognized as the highest earned academic degree in North America and provides an approximate indicator of research growth and activity (Moore and Birren 1971; Wood 1988).

2.6 Doctoral Dissertation Research

The primary method used in dissertation research is bibliometrics, the "application of mathematics and statistical methods to books and other media of communication".[1] This process has been used by researchers within a variety of disciplines to evaluate the scholarly output of a discipline (Clark et al. 1984; Baker 1986; Persson 1986; Wood 1988) and to identify trends in dissertation research (Moore and Birren 1971; Mueller and Birren 1980; Loomis 1985; Brady et al. 1988; Leahy et al. 1988, 1990, 1992; French and Raykovitz 1984; Cleary 1992). Moreover, bibliometrics has been used to describe the literature in the use of particular statistical methods (Concato et al. 1993; Savitz et al. 1994) and in family medicine (Geyman and Berg 1984, Fromm et al. 1986; Marvel et al. 1991), internal medicine (Cooper and Zangwill 1989), nursing (Larson et al. 1991), AIDS (Pratt 1992), library and information science (Dimitroff 1992) and physical therapy (Burnham et al. 1992). One example of a bibliometric study with relevance to epidemiology, is a recent project carried out by Savitz and his co-workers (1994). In this study, authors examined the practice of statistical significance testing and presentation of confidence intervals in articles appearing within the *American Journal of Epidemiology* in 1970, 1980 and 1990.

2.6.1 Describing Doctoral Research Using Dissertation Abstracts

A database appropriate for the study of doctoral research is *Dissertations Abstracts International* (*DAI*), a bibliographic database containing approximately one million records of doctoral and masters-level research (University Microfilms Inc. 1991). The use of this database is facilitated by the availability of quick and efficient on-line and CD-ROM media.

Baker (1991) demonstrated the economic and efficient utility of *DAI* for tracking the growth of social work. Using the on-line version of the database, he identified a substantial growth within the discipline from 1958 to 1980. Similarly, Clark et al. (1984) examined doctoral dissertations on hypnosis produced from 1923 to 1980. The authors reported an increase in the number of dissertations after the Second World War, accompanied by a more pronounced increase in the 1960s and 1970s.

Moore and Birren (1971) described doctoral dissertations produced in the field of gerontology between 1934 and 1969. Using key-word searching, 707 dissertation abstracts were retrieved from *Datrix* and *American Doctoral Dissertations* (previous and companion versions of *DAI*). Information relating to the author's presumed gender, degree-type and university was captured; in addition, to the type of subjects (animal, human, or plant) and the research areas studied. For each five-year interval, the total number of relevant dissertations was described by state of origin; and the total output per university was calculated for ranking purposes. Based on their findings, the researchers suggested that more females be recruited within the field of gerontology (22.5%). In addition, the proportion of dissertations examining "biology of aging" issues was relatively low, suggesting an area of future research. Meuller and Birren (1980) carried out a subsequent study of gerontology dissertations to identify further trends. In France, aging-related dissertations defended in that country from 1971 to 1985 were described. Three comparable dissertations databases were accessed to assemble the group of abstracts for this study, and each was classified by university, subject matter, type of doctoral degree awarded and presumed gender of the author (Faucheux and Minh Lim 1989).

Within the field of rehabilitation, Leahy and his colleagues described the predominant content areas studied in three series of doctoral dissertations produced in the United States (1988; 1990; 1992) and, for each period established an annotated bibliography. The samples were constructed using key-word searching of titles and abstracts contained within the *DAI* database, as well as by direct contact with university departments.

Doctoral dissertations in special education were thoroughly described by Brady and his colleagues in 1988. The authors examined five years of doctoral dissertations originating from American colleges and universities for the years 1981 through 1985. The entire population of relevant dissertations was retrieved from within *DAI* using the descriptor "special education" and degree-types Ph.D., and EdD. Each abstract was reviewed and information was extracted on the dissertation author, age and ethnicity of study subjects, study design, methods of data collection and analysis, and research content. In 94% of abstracts, the authors were able to identify the method of data collection, and in abstracts reporting use of an experimental design, 88% reported sufficient information to characterize the type of design. Based on the number and diversity of dissertations, the authors concluded that current issues within the discipline were adequately addressed; however, concern was expressed over the scarcity of dissertation research on young children and disabled adults (Brady et al. 1988).

Similar studies have been carried out in school psychology (French and Raykovitz 1984), nursing (Loomis 1985), and public administration (Cleary 1992). French and Raykovitz (1984) surveyed the directors of doctoral programs in both the United States and Canada requesting detailed information on doctoral recipients in school psychology for the years 1973 through 1980. The study period was subsequently shortened and the data were supplemented with information from *DAI* as a result of a low response rate. Afterwards, the dissertations were compared to results from an earlier study to assess whether trends in research content, methodology and type of study subjects were evident. In this same study, investigators examined the distribution of subject areas across the various types of doctoral degrees (French and Raykovitz 1984). This distinction is of interest to epidemiology since research- and practice-oriented (i.e. professional degree)

degrees are similarly awarded (e.g. Ph.D., D.Sc. and Dr.P.H). Interestingly enough, Loomis (1985) reported significant differences in research content by the type of doctoral degree awarded in her review of nursing dissertations.

While these studies have numerically described the distribution and content of doctoral research, Cleary (1992) examined the quality of dissertations produced in public administration during 1990. Dissertation abstracts were selected from within DAI and evaluated according to six criteria. The results were then compared with those from an earlier study to assess the degree of change over time.

One aspect that has not been addressed in these investigations is the potential influence of the author's gender on research practices. Although opposing views have been reported on this subject, a recent segment in *Science* suggests the possibility of a "female style" of science (Benditt 1993). Schuckman indirectly examined this issue in an earlier study carried out in 1987. Psychology and biology abstracts were retrieved from within the *DAI* database and classified according to the presumed gender of the author and dissertation advisor. A subsequent five-year search of the databases, *Psychological Abstracts* and *Biological Abstracts* was carried out to determine whether publication rates differed between male and female authors, and between male and female students having advisors of the same and opposite sex (Schuckman 1987).

2.7 Rationale for the Present Study

At present, a study describing the distribution and content of doctoral research has yet to be initiated. Given the relative youth of the discipline, and the dynamics of the population's health status, an investigation of this type will provide useful information in which to assess the diversity of epidemiologic activity. Furthermore, the relative frequencies of study designs can be followed over time to determine the use and underuse of various study methods.

---

[1] Pritchard A. Statistical bibliography or bibliometrics? *J Documen* 1969;25: 348-9.

CHAPTER 3 - MATERIALS AND METHODS

3.1 Introduction

To describe dissertation research in epidemiology, dissertation abstracts completed between 1982 and 1992 were reviewed and classified according to the study design and substantial research topic area (i.e. disease). When possible, variables describing the type and number of study subjects were recorded also. Abstracts were available on CD-ROM from the *Dissertation Abstracts International* database and were included in the study upon meeting established search and inclusion criteria.

3.2 Study Purposes

The major objectives of this study were to identify the relative frequencies with which study designs were used in epidemiology dissertations, and to describe the research topic areas being addressed in these investigations. More specifically, the following research questions were defined:

(1) What was the relative frequency with which study designs were used in epidemiologydissertations?

(2) Were time trends in the specific use of study designs evident in dissertations originating between 1982 and 1992?

(3) What were the research topic areas of interest in epidemiology dissertations?

(4) Collectively, did the research topic areas reflect prevalent diseases of today?

(5) Did the use of study designs vary with the research topic area studied?

(6) Was the use of study designs or selection of research topic areas influenced by the gender of the dissertation author or type of doctoral degree awarded?

Secondary objectives included describing the research further by person (e.g. dissertation author, type and number of study subjects), place (e.g. geographic setting) and time (e.g. 1982 through 1992); determining the inter- and intra-rater agreement for the classification of study designs and research topic areas; and evaluating the quality of the *Dissertation Abstracts International* database as a tool for studying research trends within epidemiology.

3.3 The Doctoral Dissertation Abstract

The doctoral degree is awarded to candidates for evidence of mastery in their subject area and for proficiency in scholarly research. Evidence of the scholarly method is illustrated in the doctoral dissertation, a detailed manuscript which:

> "...must embody the results of extended research, be an original contribution to knowledge and include material worthy of publication. It should demonstrate the candidate's ability to conduct an independent investigation, to abstract principles upon which predictions can be made, and to interpret in a logical manner facts and phenomena revealed by the research."[1]

The dissertation is comprised of numerous chapters describing both the methods and findings of a research study. However, the text is preceded by an abstract, a concise, 350-word summary of the manuscript. Ideally, the abstract describes the objective(s) of the research project, the basic procedures used (e.g. study design, selection, type and number of subjects, analytic methods etc.), the results and the conclusions (Haynes et al. 1990; International Committee of Medical Journal Editors 1991). When sufficient information is presented, the dissertation abstract may be used as a timely and cost-effective source in which to identify and describe research trends within disciplines. Using this source, dissertation research was previously examined in a diverse group of disciplines, including gerontology (Moore and Birren 1971; Meuller and Birren 1980; Faucheux and Minh Lim 1989), school psychology (French and Raykovitz 1984), nursing (Loomis 1985), rehabilitation (Leahy et al. 1988, 1990, 1992), special education (Brady et al. 1988) and public administration (Cleary 1992).

Likewise, in this study, the unit of observation was the dissertation abstract. A preliminary field test was carried out on a sample of 67 abstracts to determine whether the study variables could be easily extracted. In addition to the research topic area, the study design was identified in each of the abstracts reviewed.

3.3.1 Dissertations Database

Dissertation abstracts were available on CD-ROM from University Microfilms Incorporated (UMI), a publishing and information management company founded in 1938 (UMI 1989). Known specifically as *Dissertation Abstracts Ondisc* (*DAO*), the CD-ROM collection was modeled after the on-line version of the database and includes the same information contained within the original print version, *Dissertation Abstracts International* (*DAI*). Both the print and on-line versions of the database are updated monthly, while the *DAO* collection is updated quarterly (UMI 1991). Given the speed, efficiency, and relatively low cost of CD-ROM technology, the *DAO* collection remains the most time- and cost-effective media for storing and retrieving large numbers of dissertation abstracts. Recently, Repp and Glaviano (1987) found that *DAI* and other related UMI products were most often requested and used when scholars and other interested parties were in need of doctoral dissertations.

Research information from approximately 550 universities has been submitted to UMI providing a user with access to more than one million doctoral dissertations and 45,000 master's theses completed since the first American doctoral degree was awarded in 1861. While research information is included for essentially all accredited North American universities and for British graduate schools (beginning in July 1988), information is less complete for institutions elsewhere and for masters degree-level research (UMI 1989).

Included in the dissertations database are bibliographic citations and abstracts (350 words) describing doctoral research completed subsequent to July 1980. Prior to this date, only bibliographic citations are included. Permission to reproduce dissertation abstracts for the purposes of this project was granted by UMI (see Appendix 1). An example of a typical dissertation record is presented in Appendix 2.

3.3.2 Organization of the Dissertations Database

Prior to submitting the dissertation, UMI requests that dissertation authors complete an agreement form which outlines the identifying information required for indexing. The author is requested to include personal information (e.g. author's name, name of institution, type of degree, date degree was awarded etc.) and to select from a list defined by UMI, a maximum of three subject areas (in no explicit order of priority) under which the dissertation may be classified (eg. Health Sciences, Public Health; Biology, Biostatistics etc.). Key words and short phrases that will further assist in describing the research may be included also. Thereafter, officials at the university submit the abstract and completed dissertation in its entirety to UMI for microfilming.

3.3.3 Accessing the *DAO* Database

The *DAO* database is easily accessed using a personal computer and search software provided by UMI. The user may enter a UMI-designated search field (eg. author's name, institution, subject area etc.), corresponding code, or key word which describes the research of interest. Once invoked, the UMI search software compares the search statement with the bibliographic citation and abstract for each record in the database. When no further matches are found, the user is prompted with a message indicating the number of records located within that particular search set. Results are displayed immediately in long (i.e. bibliographic citation and abstract) or short format (i.e. bibliographic citation) and may be printed or downloaded to a diskette for later use.

3.4 Sample Selection

Dissertation abstracts originating from American and Canadian universities and completed between 1982 and 1992 were retrieved from the *DAO* database using a personal computer maintained by the University of Western Ontario Library System. The CD-ROM collection on file at the library included four discs and was arranged by year: 1861-1981, 1982-1987, 1988-1992 and January 1993 through June 1993. Figure 1 illustrates the procedures followed when searching for and including relevant records from within the dissertations database.

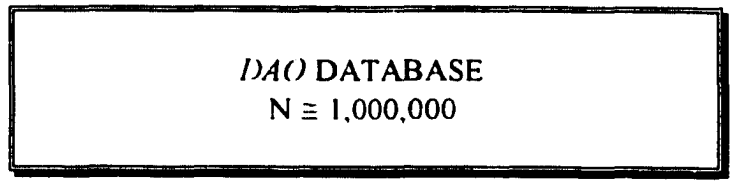

Figure 1. Search and Inclusion Criteria.

3.4.1 Sampling Frame

To identify relevant dissertations, a sampling frame of universities providing doctoral degree programs in epidemiology was constructed using published information from American and Canadian directories (see Appendix 3) (Bernier and Mason 1991; American Council on Education 1992; Association of Universities and Colleges of Canada 1993; AJPH 1993, 1994). Each institution was subsequently specified when searching the *DAO* collection.

3.4.1.1 American Universities

In the United States, doctoral degree programs in epidemiology were identified within 24 schools of public health, 13 medical schools and 13 schools of veterinary medicine (Bernier and Mason 1991; American Council on Education 1992; AJPH 1993, 1994). For this project however, doctoral programs established within schools of public health were included. This restriction was imposed since aggregate information collected by the Association of Schools of Public Health (ASPH) was available to facilitate with the inclusion of epidemiology dissertations. Furthermore, a previous study estimated that roughly 95% of doctoral degrees are awarded to graduates from these institutions (Williams et al. 1989).

3.4.1.2 Canadian Universities

Canadian schools were identified in the 1992-1994 *Directory of Canadian Universities* under the formal subject heading of "Medical/Health Sciences" (Association of Universities and Colleges of Canada 1993). Within the sub-category of "Epidemiology", doctoral degree programs were identified at the Universities of British Columbia, Guelph, McGill, and Western Ontario. In addition, the Universityof Montreal was identified in a review of the 1991 edition of this same directory (Association of Universities and Colleges of Canada 1991). Further programs were identified under the sub-heading of "Public Health and Hygiene" and included the Universities of Manitoba, Memorial, Toronto and Waterloo. A review of universities sub-classified as "Social and Preventive Medicine" did not reveal any other doctoral programs.

In addition to these institutions, the University of Alberta was identified in *Episource*, however, a structured doctoral degree program had not yet been established (Bernier and Mason 1991). As a result, this university was not included in the search and selection process; nor was the University of Waterloo since the program was not specific to epidemiology. Additionally, the Universities of British Columbia and Manitoba were excluded since these programs were not established in time to produce graduates during the study period.

3.4.2 Search Statement

When defining the *DAO* search statement, the search fields, "SCHOOL'" and "SUBJECT" were specified for each of the institutions outlined in Appendix 3. The school field was abbreviated using a UMI-assigned four-digit school code and was combined with the broad subject field of "Health Sciences" using the logical operator "AND" (see Figure 1). Thereafter, abstracts were sorted by the year in which the doctoral degree was awarded, and downloaded in long format to a floppy diskette. The search strategy used in this project was verified by an individual at UMI responsible for database searches, and is consistent with a survey that reported users' success when searching the database by subject area (Repp and Glaviano 1987).

Abstracts indexed as "health sciences" were selected since an individual subject area for "epidemiology" had, unfortunately, not been designated by UMI. As a result, a comprehensive subject category incorporating both physical and mental disease was required to include epidemiologic research otherwise classified Appendix 4 outlines each of the 25 distinct health-related subject areas classified as "health sciences".

For the majority of institutions, UMI had assigned a distinct university-wide school code. However, for the Universities of Alabama, Illinois, Johns Hopkins, Massachusetts, Michigan, Oklahoma, and Tulane, UMI had assigned distinct codes specific to the schools of public health and main campus. Consequently, a separate search statement for each of these additional codes was defined and executed.

3.4.3 Inclusion Criteria

Inclusion criteria were established to refine the sample given the comprehensive search strategy and overlap between epidemiology and other health-related disciplines. Criteria were defined according to the completeness of the record, language of the dissertation abstract, type of doctoral degree awarded, number of study subjects and operational definition of epidemiology

3.4.3.1 Initial Inspection of Dissertation Records

In the initial review, each record was examined and included in the study when the record was complete (i.e. bibliographic citation and abstract present), the abstract was written in English (bibliographic citation excepted), the degree awarded for doctoral research in epidemiology was a Ph.D., Dr.P.H., or Sc.D degree (American Council on Education 1992; Levin 1993, Association of Universities and Colleges of Canada 1991, 1993; Bernier and Mason 1991), and when the author of the dissertation described research carried out on a sample of human subjects (defined as $\geq$ 10 study subjects). Furthermore, American abstracts were included when they were found to be consistent with summary data reported by the ASPH. More specifically, the school of public health at San Diego State University did not award doctoral degrees between 1982 and 1992. Likewise, doctoral degrees were not awarded at the University of South Florida College of Public Health between 1982 and 1991 (Levin 1993). The number of dissertation abstracts included after this first inspection is presented by institution in Appendix 5.

3.4.3.2 Missing Abstracts

When the abstract was missing from the bibliographic citation, the print version of the database was consulted in an attempt to follow it up. Of the 122 missing abstracts, 6 abstracts were located, while 116 could not be found (see Table 1). Consequently, abstracts within this larger group could not be reviewed and were excluded from the total number of records retrieved (1.9%). For Harvard University, UMI did not include dissertation abstracts within the database. Instead, only bibliographic citations were included.

Table 1. Number of missing dissertation abstracts.

School	No. Missing Abstracts	No. Abstracts Found	No. Abstracts Not Found
Harvard University	107	0	107
U of Toronto	1	0	1
U of Washington	2	1	1
U Cal at Berkeley	3	2	1
Tulane University	1	0	1
UNC at Chapel Hill	1	0	1
Yale University	3	0	3
U of Pittsburgh	3	3	0
Johns Hopkins	1	0	1
Total	122	6	116

3.4.3.3 Subsequent Inspection of Dissertation Records

In the second review, the study investigator examined each remaining abstract to determine whether the content was epidemiologic. For this purpose, a suitable definition of epidemiology was required that would differentiate between (a) clinical medicine and epidemiology; and (b) would outline both the internal and external factors influencing health and disease. The following definition was selected from those listed in Appendix 6 and was applied to each abstract:

> "Epidemiology is the study of all factors and their interrelationships which affect the occurrence and course of health and disease in a population. These factors include the characteristics of the host population; the causative agencies-predisposing, precipitating and perpetuating; and the biological, physical and social environment."[2]

Furthermore, in this study, health and disease were operationally defined as objective physiological or psychological illnesses (hereafter referred to as disease), recognized and characterized by a distinct code in the main index of the ninth revision of the International Classification of Diseases and Injuries, Clinical Modification (Commission on Professional and Hospital Activities 1978). Most importantly, the study disease was the outcome described (e.g. distribution of disease by person, place and time

etc.) and/or measured with respect to the frequency (e.g. prevalence, incidence, mortality etc.) or risk of disease (e.g. relative risk).

In summary, each record was included in the study when the content of the abstract was judged to be consistent with the operationalized definition of epidemiology. Several distinctions may require further elaboration. A record was included in the study when the disease was broadly defined as "psychiatric" or "respiratory" disease. Furthermore, a record was included when the research described a cohort or cross-sectional study that was designed to investigate the incidence, prevalence or mortality of numerous, yet distinct diseases (e.g. occupational cohort study measuring all-cause mortality and mortality from specific causes such as coronary heart disease, stroke and cancer). Similarly, records were included when the outcome described "overall mortality". Although smoking and problem drinking are considered risk factors for numerous diseases, they were considered addictions and were classified accordingly in this study. The number of records included after this second review is outlined for each institution and is included in Appendix 5, and three examples of included records are presented in Appendix 7. In total, 758 dissertion records were included in this study.

3.4.3.4 Excluded Records

For comparison, examples of excluded records are presented in Appendix 8. These records are best described as: (1) ill-defined health problems and clinical studies; (2) methodological studies; (3) rehabilitation-related and functional health status research; (4) health services research; and (5) health promotion studies.

This first group of records was excluded since the study disease could not be characterized using the ICD-9-CM. For example, Durand (UMI 1986) examined the association between exposure to asbestos and acute respiratory changes (Appendix 8). While this outcome may be indicative of subsequent disease, it is not specific and could not be appropriately classified. Examples of other ill-defined health problems that could not be assigned a disease code include studies of blood and cell physiology (e.g. red blood cell count, sister chromatid exchange etc.) and the measurement of blood pressure without reference to documented hypertension.

Methodological records were not included in the study when the primary focus of the research was to describe and evaluate a biostatistical procedure or technique. As well, abstracts investigating the utility of data sources, data collection methods, and follow-up procedures were excluded. The remaining groups of abstracts were not included since the outcome measures described rehabilitation and functional health status outcomes (eg. limitation of daily activities, quality of life etc.), utilization of health services (eg. patient-days in hospital), and health knowledge, beliefs, behaviors, and attitudes (eg. perceived risk of HIV infection). In each of these groups, the description and/or measurement of disease was not the study outcome. Furthermore, outcomes describing health services and disability-related research may be more appropriately classified using ICD specialty adaptations (e.g. International Classification of Impairments, Disabilities, and Handicaps).

Although these exclusion criteria may be considered restrictive, the overlap between epidemiology and other disciplines (e.g. nursing, health administration, health education etc.) was minimized. Minimizing this dilution effect was essential in order to establish a sample of abstracts that consistently described epidemiologic research. We do acknowledge however, that by defining our definition of epidemiology in the manner described above, we have limited the representativeness of epidemiologic research as reflected in doctoral dissertations. More specifically, epidemiologic research addressing methodological issues and ill-defined conditions (e.g. acute respiratory changes) and health problems (e.g. limitation of daily activities, elder abuse, family dysfunction etc.) were not described in this study. As a result, the findings from this project are limited to dissertation research describing illness outcomes.

3.4.4 Verification

To measure the reliability of these inclusion criteria, a random sample of 5% (n=38) of included and 1% (n=18) excluded abstracts was reviewed and rated by each of two additional reviewers. Kappa statistics and 95% confidence intervals were tabulated to measure the level of agreement observed beyond chance between the principal investigator and each of these raters. Kappa statistics above .75 indicate excellent agreement beyond

chance, values between .40 and .75 reflect fair to good agreement, and values below .40 suggest that agreement was poor beyond chance (Fleiss 1981)

3.5 Variables

Study variables describing the dissertation author and affiliated institution were extracted directly from the bibliographic citation. Information describing the content of the dissertation was extracted from the abstract and coded using classification procedures developed for this study. Initially, the study investigator and a second reviewer each separately coded a series of abstracts, afterwhich coding practices were established by consensus.

3.5.1 Bibliographic Citation

The following variables were extracted from the bibliographic citation:

(1) presumed gender of the author;

(2) institution;

(3) type of doctoral degree awarded;

(4) year in which the degree was awarded;

(5) author-assigned subject area (i.e. dissertation author)

The author's first name was used to establish gender. The four-digit school code and the year in which the degree was awarded were recorded directly. A categorical variable was created to capture the degree-type. In addition, the author-assigned subject areas were included to assist in developing search strategies for future studies. A coding rule book and data dictionary describing each of these variables is included in Appendix 9.

3.5.2 Abstract

Upon reviewing the content of the abstract, the following variables were extracted:

(1) study design (primary variable);

(2) research topic area (primary variable);

(3) number of study subjects;

(4) ethnicity of study subjects;

(5) gender of study subjects;

(6) age-group of study subjects;

(7) setting of dissertation research

The study design was classified using a taxonomy developed for this project (see Table 2 below). Likewise, the research topic area was identified and classified using the ICD-9 (CM). A data dictionary describing these and other study variables is outlined in Appendix 9. In addition, example records and sample data abstraction are provided in Appendix 10.

3.5.3 Study Design Classification Scheme

The purpose of this classification scheme was to systematically organize study design information into meaningful categories. Ideally, the taxonomy would be characterized by naturalness, exhaustiveness, usefulness, simplicity and constructability (Last 1988). Furthermore, the classification scheme would be uniform across raters and constant over time.

The classification of study designs has been the subject of numerous theoretical discussions (Kramer and Boivin 1987; Greenland and Morgenstern 1988; Feinstein 1988; Miettinen 1988; Maclure 1991). Several axes have been defined to avoid confusion in classification (Kramer and Boivin 1987; Maclure 1991), and unique labels have been established to express information relating to the selection of subjects and direction of observation (e.g. trohoc study) (Feinstein 1978). In this project however, the classification scheme was developed to best represent the study designs used regularly in epidemiology.

For this purpose, information from existing taxonomies described by Fletcher and Fletcher (1979) and Kleinbaum, Kupper and Morgenstern (1982) was combined and modified to produce the taxonomy used in this project. Fletcher and Fletcher identified four basic designs in their classification scheme: cross-sectional studies, case-control studies, cohort studies and clinical trials. Kleinbaum and his colleagues identified three principal categories in their typology of observational designs. These categories included:

basic designs (e.g. cohort study, case-control study and cross-sectional study); hybrid designs (e.g. nested case-control study and others) and incomplete designs (e.g. ecological designs etc.).

Initially, a field test was carried out on a sample of 67 abstracts to evaluate the feasibility of the proposed classification scheme. Although the results were satisfactory, the taxonomy was revised to obtain more precise information as outlined in Table 2.

Table 2. Study design classification scheme.

Study Designs
1. Intervention Study
(a) randomized controlled clinical trial
(b) other intervention study
2. Cohort Study
(a) prospective
(b) retrospective
3. Case-Control Study
4. Cross-Sectional Study
5. Ecological Study
6. Hybrid Study
7. Other Study
8. Cannot Classify

Randomized clinical trials were differentiated from other intervention studies to reflect the scientific rigor of the design (Fletcher and Fletcher 1979; Maclure 1991). Because the number of community intervention studies was not expected to be large, a separate category was not introduced. Instead, these studies were identified as a proportion of all intervention studies.

For each abstract, the investigator reviewed the content and classified the study design according to the taxonomy outlined in Table 2. However, when the dissertation author identified the study design, it was classified accordingly. When an abstract described more than one investigation, each study design was classified, but for the

purpose of describing the remaining variables, the predominant study was selected according to the quality and quantity of information provided. When this selection was not possible, the investigation using the less rigorous of study designs was selected.

In some cases, a cohort study was not described in sufficient detail to characterize the direction of observation. When this occurred, the design was classified simply as a cohort study. Furthermore, when the design could not be classified with confidence, the abstract was forwarded to a second reviewer for consensus. When a decision could not be reached at this stage, the study design was assigned a rating of "cannot classify". Further detail describing the classification of study designs is included in Appendix 9.

3.5.4 Classification of Research Topic Areas

The investigator reviewed the abstract and classified the substantive research topic area (i.e. study disease) using information and conventions outlined in volumes one and two of the ninth revision of the International Classification of Diseases and Injuries, Clinical Modification (ICD-9-CM). The ICD is an internationally accepted disease classification system and is recognized for general epidemiologic purposes (WHO 1992). Although a new revision of the ICD was available (i.e. ICD-10), the alphabetical index referencing the location of codes within the main index had not yet been published.

The main index of the ICD-9-CM includes seventeen chapters (see Appendix 11). Within each, diseases are sub-divided and are identified by a three-digit rubric with further sub-division provided using a two decimal point structure. As well, the ICD-9-CM includes two supplementary chapters: (a) Classification of Factors Influencing Health Status and Contact with Health Services - V Codes; and (b) Classification of External Causes of Injury and Poisoning - E Codes. For this project, the supplementary V-code chapter was not used to classify research areas since inclusion criteria were established to identify abstracts describing or measuring a disease. In contrast, E-codes were used in conjunction with codes from the main index to identify the cause of injury or poisoning.

Research areas were classified with as much detail as that provided by the dissertation author; although, in some cases, a range of three-digit codes was most appropriate. For example, when an author investigated "cancer" without further specification, the research area was classified according to the range of codes identified as "neoplasms" (e.g. 140-239). Furthermore, the ICD-9-CM did not include codes for HIV, AIDS and Hepatitis C infection. Therefore, appropriate codes listed within the ICD-10 were selected.

While every attempt was made to accurately identify and classify the research area, difficulties arose when classifying abstracts describing cohort or cross-sectional studies that were designed to investigate more than one disease. When this occurred, the abstract was identified as studying "multiple diseases" and assigned a summary code appropriate for the group of diseases. More specifically, if a dissertation author carried out a retrospective cohort study to examine the effects of occupation on mortality from cancers of the lung, prostate and bladder, the main research area was classified as "multiple cancers", and each specific disease was documented accordingly. However, when the study diseases extended beyond one chapter in the ICD, the main research area was classified simply as "multiple diseases" and each specific disease was recorded in separate variables

Finally, a summary code for "adverse pregnancy outcomes" was defined. This code was used to classify dissertation abstracts in which there were multiple pregnancy outcomes extending, in most cases, beyond one ICD chapter(e.g. chapter 11 and 15) When summarizing the research areas, this code was included within chapter 15, Certain Conditions Originating in the Perinatal Period.

3.6 Inter- and Intra-Rater Agreement

The primary variables, study design and research topic area were classified independently by a second reviewer and were re-classified by the principal investigator approximately one month following the original coding of abstracts. A random sample of 5% (n=38) of abstracts was selected for this purpose. Kappa statistics and 95% confidence intervals (Fleiss 1981) were tabulated to measure the level of inter- and intra-

rater agreement with respect to the broad and detailed classification of study designs (see Table 2). Because the ICD-9-CM contains several hundred rubrics, percent agreement was used to measure the level of agreement for the research topic area.

3.7 Data Entry and Analysis

SPSS Data Entry II was used to record the study information extracted from the bibliographic citation and abstract (SPSS Inc. 1987). Analyses were carried out using SPSS/PC + (version 5.0.1) (SPSS Inc. 1992).

Because of the descriptive nature of this study and the problems encountered with respect to multiple comparisons, formal statistical analyses were not carried out. Instead, 95% confidence intervals were constructed for each proportion using the following formula defined by Snedecor and Cochran (1989) $p \pm 1.96 \sqrt{p(1-p)/n}$, where p, is the proportionof abstracts describing a particular study design or research topic area, and n, is the total number of abstracts included in the study.

The relative frequencies of study designs and research topic areas were examined overall and across four time periods to determine whether time trends were evident (e.g. 1982-83, 1984-86, 1987-89, 1990-92). For this last purpose, nested case-control studies were collapsed with those reporting case-control designs. The research topic areas were summarized according to the chapters outlined in the ICD-9-CM, and distinct disease areas were examined in further detail. Each variable was further described by gender of the dissertation author and type of doctoral degree awarded (e.g. Ph.D. vs. other degrees). When possible, the age, gender and ethnic distribution of study subjects was described. Finally, the minimum, maximum and median number of subjects was tabulated for intervention, cohort, case-control and cross-sectional studies.

---

[1] Mellon Institute of Science and Carnegie Institute of Technology. Graduate Studies, 1970-1972. Pittsburgh: Carnegie-Mellon University, 1970 as cited in Boyer CJ. The Doctoral Dissertation as an Information Source. A Study of Scientific Information Flow. New Jersey: The Scarecrow Press, Inc., 1973: 1.

[2] As cited in White KL. Healing the Schism. Epidemiology, Medicine, and the Public's Health. New York: Springer-Verlag, 1992: 129.

CHAPTER 4 - RESULTS

4.1 Introduction

The content of dissertation abstracts originating from epidemiology programs was classified according to the study designs used, the research topic areas investigated, and the type of subjects studied. The primary variables were examined collectively and by gender and type of doctoral degree awarded. To examine trends over time, four time periods were established. Research topic areas were collapsed into ICD-9 chapters, and the most common areas were examined in further detail.

4.2 Agreement Measures

To measure the reliability of inclusion criteria, kappa statistics and 95% confidence intervals were calculated for a sub-sample of abstracts reviewed by each of two additional investigators. For data presented in Table 3, the kappa statistic (with 95% confidence interval) was .72 (.46, .98) indicating fair to good agreement (Fleiss 1981) with respect to whether the dissertation abstract met the definition of epidemiology required for inclusion in this study.

Table 3. Sample agreement: rater A by rater B.

		Rater B		
		Include	Exclude	Total
Rater A	Include	34 (.61)	4 (.07)	38 (.68)
	Exclude	3 (.05)	15 (.27)	18 (.32)
	Total	37 (.66)	19 (.34)	56 (1.00)

The kappa statistic tabulated for data presented in Table 4 was .47 (.25, .69). While this value indicates fair to good agreement, the lower limit of the confidence interval suggests that the estimate may range as low as .25. Agreement beyond chance is poor for values below .40 (Fleiss 1981).

Table 4. Sample agreement: rater A by rater C.

		Rater C		
		Include	Exclude	Total
Rater A	Include	22 (.39)	16 (.29)	38 (.68)
	Exclude	0 (0)	18 (.32)	18 (.32)
	Total	22 (.39)	34 (.61)	56 (1.00)

Inter- and intra-rater agreement were assessed for the primary variables, study design and research topic area. A sub-sample of abstracts was separately coded by a second investigator, and was re-coded by the principal investigator. The kappa statistic for inter-rater agreement was .75 (.63, .88) for the classification of detailed study designs (e.g. prospective vs. retrospective cohort studies etc.), and was .77 (.61, .92) when designs were collapsed according to the broad categories summarized in Table 2 (see section 3.5.3). Although these values indicate excellent agreement beyond chance, the lower limits of the confidence intervals suggest that inter-rater agreement may have been in the range of fair to good (Fleiss 1981).

With respect to intra-rater agreement, the kappa statistics were .88 (.76, 1.00) for detailed study designs and .90 (.75, 1.00) for broad categories. Between raters, percent agreement for the research topic areas was 89.5%. Intra-rater agreement for this variable was 97.4%.

4.3 Selected Characteristics of Epidemiology Dissertations

Of the abstracts included, 3.2% (n=24) originated from 4 Canadian universities and 96.8% (n=734) originated from 21 American institutions. The distribution of abstracts across universities is presented in Table 5. The number of doctoral degrees awarded over the study period ranged from 1 to 91. More than half of the universities awarded between 1 and 24 degrees. Six universities awarded more than 50 doctoral degrees and collectively, accounted for 56% of dissertation abstracts.

Table 5. Distribution of epidemiology dissertations across Canadian and American universities, 1982-1992.

Dissertations	University
1-24	Boston University, University of Hawaii, Loma Linda University, University of Massachusetts, Memorial University of Newfoundland, Emory University, University of Toronto, McGill University, University of Western Ontario, University of South Carolina, Tulane University, University of Illinois at Chicago, University of Oklahoma
25-49	University of Alabama, University of California at Los Angeles, Columbia University, University of Michigan, University of Minnesota, Yale University
50-74	University of California at Berkeley, University of Pittsburgh, University of Texas at Houson, University of Washington
75+	Johns Hopkins University, University of North Carolina at Chapel Hill

Table 6 summarizes selected characteristics of epidemiology dissertations. In the United States, the number of dissertations increased steadily from 1982, and most were completed by females (60.2%). With respect to the type of doctoral degree, the Ph.D. degree was awarded to all Canadian graduates and most frequently to American students (82.3%) Males received 35.7% of Ph.D. degrees and 44.1% of all other doctoral degrees. Female graduates were awarded 64.3% of Ph.D. degrees and 55.9% of others.

"Public Health" was the predominant subject area used to classify dissertations (77.8%). However, this finding was far more striking in American dissertations, and when summed across the two remaining subject areas, the percentage of American abstracts with a similar classification increased to 89.2%. Collectively, epidemiology dissertations ranged in length from a minimum of 39 to a maximum of 1,290 pages.

Table 6. Selected characteristics of epidemiology dissertations.

Variable	Canadian Universities		American Universities		Total	
	no.	(%)	no.	(%)	no.	(%)
Gender						
Male	10	(41.7)	261	(35.6)	271	(35.8)
Female	14	(58.3)	442	(60.2)	456	(60.2)
Indeterminate	0	0	31	(4.2)	31	(4.1)
Type of Degree						
Ph.D.	24	(100)	604	(82.3)	628	(82.8)
Sc.D.	0	0	3	(0.4)	3	(0.4)
Dr.P.H.	0	0	127	(17.3)	127	(16.8)
Year						
1982-83	8	(33.2)	100	(13.6)	108	(14.2)
1984-86	7	(29.2)	160	(21.8)	167	(22.0)
1987-89	4	(16.7)	200	(27.2)	204	(26.9)
1990-92	5	(20.8)	274	(37.3)	279	(36.8)
Subject Area						
Public Health	9	(37.5)	581	(79.2)	590	(77.8)
Nutrition	1	(4.2)	23	(3.1)	24	(3.2)
Medicine/Surgery	4	(16.7)	15	(2.0)	19	(2.5)
Mental Health	5	(20.8)	14	(1.9)	19	(2.5)
Others	5	(20.8)	101	(13.8)	106	(14.0)
Total	24	(100)	734	(100)	758	(100)

4.4 Study Designs

The frequency with which study designs were used in Canadian and American epidemiology dissertations is presented in Table 7 and illustrated in Figure 2. Overall, observational studies accounted for more than 80% of dissertations. The use of case-control, cohort and cross-sectional studies did not appear to differ from one another. However, when grouped with hybrid studies (i.e. nested case-control studies), the case-control was the most common design utilized in dissertation research (32.7%; 29.4%-36.1%).

Table 7. Distribution of study designs by Canadian and American universities, 1982-1992.

Study Design	Canadian Universities		American Universities		Total		
	no.	(%)	no.	(%)	no.	(%)	(95% CI)
Intervention Study	3	(12.5)	32	(4.4)	35	(4.6)	(3.1, 6.1)
RCT	*2*	*(66.7)*	*17*	*(53.1)*	*19*	*(54.3)*	*(37.8, 70.8)*
Other Intervention Study	*1*	*(33.3)*	*15*	*(46.9)*	*16*	*(45.7)*	*(29.2, 62.2)*
Cohort Study	9	(37.5)	188	(25.6)	197	(26.0)	(22.9, 29.1)
Prospective Cohort Study	*7*	*(77.8)*	*72*	*(38.3)*	*79*	*(40.1)*	*(33.3, 46.9)*
Retrospective Cohort Study	*2*	*(22.2)*	*66*	*(35.1)*	*68*	*(34.5)*	*(27.9, 41.2)*
Cohort Only	*0*	*0*	*50*	*(26.6)*	*50*	*(25.4)*	*(19.3, 31.5)*
Case-Control	5	(20.8)	223	(30.4)	228	(30.1)	(26.8, 33.3)
Cross-Sectional	4	(16.7)	163	(22.2)	167	(22.0)	(19.1, 25.0)
Ecological	1	(4.2)	38	(5.2)	39	(5.1)	(3.6, 6.7)
Hybrid Study*	1	(4.2)	19	(2.6)	20	(2.6)	(1.5, 3.8)
Other	0	0	7	(1.0)	7	(0.9)	(0.2, 1.6)
Cannot Classify	1	(4.2)	64	(8.7)	65	(8.6)	(6.6, 10.6)
Total	24	(100)	734	(100)	758	(100)	

*Nested case-control study.

Of the cohort studies reviewed, 40.1% were classified as prospective cohort studies, 34.5% were classified as retrospective and 25.4% could not be characterized accordingly. Ecological (5.1%) and intervention studies (4.6%) were infrequently used in epidemiology dissertations. When classified further, randomized clinical trials accounted for 54.3% of intervention studies, while other intervention designs represented 45.7%. Because these particular studies were based on small numbers of dissertations, the confidence intervals were wide and overlapping.

Overall, the frequency of "other" study designs was low in this study and included 2 proportional mortality studies, 4 family studies, and 1 heterosexual partner study. Of all dissertations reviewed, 8.6% were considered ambiguous and could not be classified

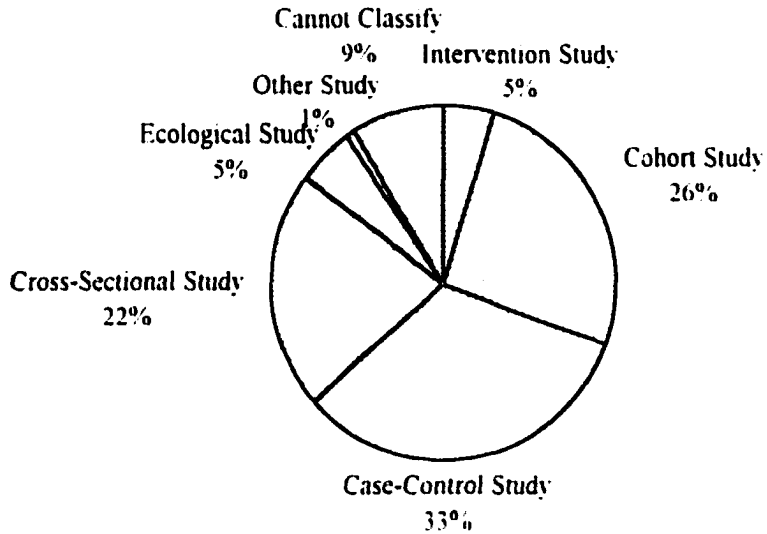

Figure 2. Frequency of study designs. 1982-1992

according to the study design. In these abstracts, the design was most often suspected as cross-sectional or ecological; however, neither design could be classified with certainty.

In dissertations carrying out secondary data analyses, research information was included from well-known studies like the Framingham Study, Multiple Risk Factor Intervention Trial, National Health and Nutrition Examination Survey and the Hypertension Detection and Follow-up Program. Other less known datasets were included also.

The relative frequencies of study designs are summarized in Table 8 and the major study types are illustrated in Figure 3. In this analysis, nested case-control and case-control studies were combined. As well, intervention and cohort studies were not described further because of small cell sizes and substantial number of unclassified cohort studies.

Throughout the study period, intervention and ecological studies remained relatively stable, accounting for roughly 5% of the dissertations. The relative frequency of cohort studies increased from 1982 to 1989, however, there was no observable trend. Similarly, time trends were not evident for other study designs.

Table 8. Relative frequency of study designs used in epidemiology dissertations, 1982-1992

	1982-83			1984-86			1987-89		
Study Design	no.	(%)	(95% CI)	no.	(%)	(95% CI)	no.	(%)	(95% CI)
Intervention Study	5	(4.6)	(0.7, 8.6)	10	(6.0)	(2.4, 9.6)	8	(3.9)	(1.3, 6.6)
Cohort Study	25	(23.1)	(15.2, 31.1)	40	(24.0)	(17.5, 30.4)	58	(28.4)	(22.2, 34.6)
Case-Control Study*	37	(34.3)	(25.3, 43.2)	59	(35.3)	(28.1, 42.6)	63	(30.9)	(24.5, 37.2)
Cross-Sectional Study	29	(26.9)	(18.5, 35.2)	34	(20.4)	(14.3, 26.5)	43	(21.1)	(15.5, 26.7)
Ecological Study	5	(4.6)	(0.7, 8.6)	12	(7.2)	(3.3, 11.1)	8	(3.9)	(1.3, 6.6)
Other	0	0	0	2	(1.2)	(-, 2.8)	2	(1.0)	(-, 2.3)
Cannot Classify	7	(6.5)	(1.8, 11.1)	10	(6.0)	(2.4, 9.6)	22	(10.8)	(6.5, 15.0)
Total	108	(100)		167	(100)		204		

	1990-92			Total		
Study Design	no.	(%)	(95% CI)	no.	(%)	(95% CI)
Intervention Study	12	(4.3)	(1.9, 6.7)	35	(4.6)	(3.1, 6.1)
Cohort Study	74	(26.5)	(21.3, 31.7)	197	(26.0)	(22.9, 29.1)
Case-Control Study*	89	(31.9)	(26.4, 37.4)	248	(32.7)	(29.4, 36.1)
Cross-Sectional Study	61	(21.9)	(17.0, 26.7)	167	(22.0)	(19.1, 25.0)
Ecological Study	14	(5.0)	(2.5, 7.6)	39	(5.1)	(3.6, 6.7)
Other	3	(1.1)	(-, 2.3)	7	(0.9)	(0.2, 1.6)
Cannot Classify	26	(9.3)	(5.9, 12.7)	65	(8.6)	(6.6, 10.6)
Total	279	(100)		758	(100)	

* Includes nested case-control studies

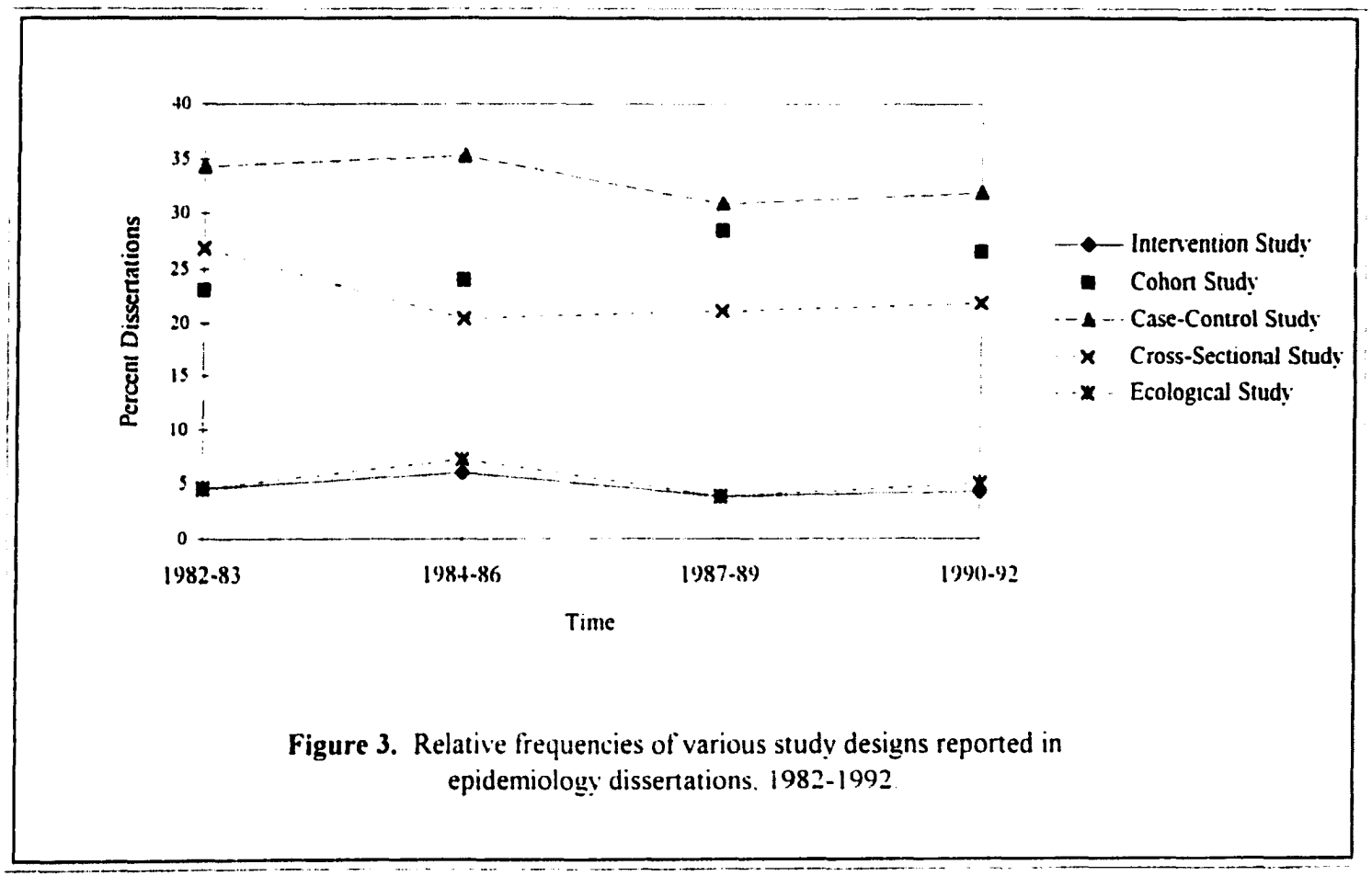

Figure 3. Relative frequencies of various study designs reported in epidemiology dissertations, 1982-1992.

4.4.1 Gender of Dissertation Author

Since the presumed gender of the author could not be determined in 4.1% of abstracts, comparisons were based on a total of 727 abstracts. As shown in Table 9, the use of study designs did not appear to differ between male and female graduates

Table 9. Distribution of study designs by gender of dissertation author.

Study Design	Male			Female			Total	
	no.	(%)	(95% CI)	no.	(%)	(95% CI)	no.	(%)
Intervention Study	12	(4.4)	(2.0, 6.9)	22	(4.8)	(2.9, 6.8)	34	(4.7)
Cohort Study	67	(24.7)	(19.6, 29.9)	123	(27.0)	(22.9, 31.0)	190	(26.1)
Case-Control Study*	83	(30.6)	(25.1, 36.1)	156	(34.2)	(29.9, 38.6)	239	(32.9)
Cross-Sectional	61	(22.5)	(17.5, 27.5)	99	(21.7)	(17.9, 25.5)	160	(22.0)
Ecological Study	19	(7.0)	(4.0, 10.1)	18	(3.9)	(2.2, 5.7)	37	(5.1)
Other Study	1	(0.4)	(-, 1.1)	5	(1.1)	(0.1, 2.1)	6	(0.8)
Cannot Classify	28	(10.3)	(6.7, 14.0)	33	(7.2)	(4.9, 9.6)	61	(8.4)
Total	271	(100)		456	(100)		727	(100)

*Includes nested case-control studies.

4.4.2 Distribution of Study Designs by Type of Doctoral Degree

Because Canadian epidemiology programs award only the Ph.D. degree, the distribution of study designs by degree-type was restricted to American dissertations (n=734) (Bernier and Mason 1991). For this purpose, Dr.P.H. and Sc.D. degrees were combined and compared to the Ph.D. degree. The data in Table 10 indicate that case-control studies were most often carried out by Ph.D. graduates (35.1% vs. 23.1%), while cross-sectional studies were frequently described in dissertations completed by other graduates (33.8% vs. 19.7%). However, as discussed below in section 5.2, the Dr.P.H. and Sc.D. degrees were under-reported in the dissertations database. As a result, these comparisons may not be valid.

Table 10. Distribution of study designs by type of doctoral degree awarded.

Study Design	Ph.D			Other			Total	
	no.	(%)	(95% CI)	no.	(%)	(95% CI)	no.	(%)
Intervention Study	25	(4.1)	(2.6, 5.7)	7	(5.4)	(1.5, 9.3)	32	(4.4)
Cohort Study	161	(26.7)	(23.1, 30.2)	27	(20.8)	(13.8, 27.7)	188	(25.6)
Case-Control Study*	212	(35.1)	(31.3, 38.9)	30	(23.1)	(15.8, 30.3)	242	(33.0)
Cross-Sectional	119	(19.7)	(16.5, 22.9)	44	(33.8)	(25.7, 42.0)	163	(22.2)
Ecological Study	29	(4.8)	(3.1, 6.5)	9	(6.9)	(2.6, 11.3)	38	(5.2)
Other Study	7	(1.2)	(0.3, 2.0)	0	0	0	7	(1.0)
Cannot Classify	51	(8.4)	(6.2, 10.7)	13	(10.3)	(4.8, 15.2)	64	(8.7)
Total	604	(100)		130	(100)		734	(100)

*Includes nested case-control studies.

4.4.3 Multiple Study Designs

Of all abstracts reviewed, 44 (5.8%) described two distinct studies, while 1 described three investigations. As shown in Table 11, secondary investigations were often case-control studies (29.5%), followed by cohort (25.0%), and cross-sectional studies (18.2%). "Other" studies included investigations labelled as incidence studies. Although Last (1988) considers this term a synonym for cohort studies, the information contained within the abstract was ambiguous and could not be re-classified with certainty.

Table 11. Frequency of study designs used in secondary and tertiary investigations, 1982-1992.

Study Designs	Secondary Investigations		Tertiary Investigations	
	no	(%)	no	(%)
Intervention Studies	1	(2.3)	0	0
Cohort Studies	11	(25.0)	0	0
Case-Control Studies	7	(15.9)	0	0
Cross-Sectional Studies	8	(18.2)	0	0
Ecological Studies	4	(9.1)	0	0
Hybrid Studies	6	(13.6)	0	0
Other Studies	2	(4.5)	0	0
Cannot Classify	5	(11.4)	1	(100)
Total	44	(100)	1	(100)

4.5 Research Topic Areas

The research topic areas investigated in epidemiology dissertations are presented in Table 12. The research area of greatest activity was neoplasms (22.8%). Other prominent topic areas included infectious and parasitic diseases (11.5%), diseases of the circulatory system (10.9%), mental disorders (7.7%) and injuries (6.9%). Furthermore, maternal and infant health dissertations were well-represented (11.6%) when the appropriate chapters were collapsed (i.e. complications of pregnancy and childbirth and conditions originating in the perinatal period). Research areas infrequently addressed in epidemiology dissertations included congenital anomalies (1.6%) and diseases of the musculoskeletal system and connective tissue (1.5%). Dissertation research was absent for diseases of the skin and subcutaneous tissue.

For the purposes of this project, two additional chapters were created. The first chapter, "mortality", was created in order to capture information that was concerned with overall mortality. Altogether, this group represented 2% of dissertations reviewed. In comparison, the last chapter, "multiple diseases" was created to capture disease-specific information that extended beyond the limits of one ICD chapter. Five percent of dissertations were accounted for by this group. Infectious diseases, neoplasms, mental

Table 12. Distribution of research topic areas by Canadian and American universities, 1982-1992.

Research Topic Area	Canadian Universities		American Universities		Total		
	no.	(%)	no.	(%)	no.	(%)	(95% CI)
Infectious and Parasitic Diseases	1	(4.2)	86	(11.7)	87	(11.5)	(9.2, 13.7)
Neoplasms	3	(12.5)	170	(23.2)	173	(22.8)	(19.8, 25.8)
Endocrine, Nutritional, Metabolic and Immunity Disorders	0	0	36	(4.9)	36	(4.7)	(3.2, 6.3)
Diseases of the Blood and Blood-Forming Organs	0	0	3	(0.4)	3	(0.4)	(-, 0.8)
Mental Disorders	2	(8.3)	56	(7.6)	58	(7.7)	(5.8, 9.5)
Diseases of the Nervous System and Sense Organs	1	(4.2)	21	(2.9)	22	(2.9)	(1.7, 4.1)
Diseases of the Circulatory System	4	(16.7)	79	(10.8)	83	(10.9)	(8.7, 13.2)
Diseases of the Respiratory System	0	0	23	(3.1)	23	(3.0)	(1.8, 4.3)
Diseases of the Digestive System	3	(12.5)	22	(3.0)	25	(3.3)	(2.0, 4.6)
Diseases of the Genitourinary System	1	(4.2)	18	(2.5)	19	(2.5)	(1.4, 3.6)
Complications of Pregnancy, Childbirth and the Puerperium	1	(4.2)	31	(4.2)	32	(4.2)	(2.8, 5.7)
Diseases of the Skin and Subcutaneous Tissue	0	0	0	0	0	0	0
Diseases of the Musculoskeletal System/Connective Tissue	0	0	11	(1.5)	11	(1.5)	(0.6, 2.3)
Congenital Anomalies	1	(4.2)	11	(1.5)	12	(1.6)	(0.7, 2.5)
Certain Conditions Originating in the Perinatal Period	2	(8.3)	54	(7.4)	56	(7.4)	(5.5, 9.3)
Symptoms, Signs and Ill-Defined Conditions	1	(4.2)	12	(1.6)	13	(1.7)	(0.8, 2.6)
Injuries and Poisoning	3	(12.5)	49	(6.7)	52	(6.9)	(5.1, 8.7)
Mortality	0	0	15	(2.0)	15	(2.0)	(1.0, 3.0)
Multiple Diseases	1	(4.2)	37	(5.0)	38	(5.0)	(3.5, 6.6)
Total	24	(100)	734	(100)	758	(100)	

disorders, injuries and diseases of the circulatory system were the most commonly encountered research topic areas within this category.

Table 13 shows the relative frequencies of the various research topic areas investigated over the eleven-year study period. Although the relative frequency of infectious disease-related dissertations appeared to increase over time, we were unable to observe any consistent trends. Similarly, an increase in maternal and infant health dissertations was not evident.

For descriptive purposes, diseases were rank-ordered within each ICD chapter (Table 14). However, detailed data were not presented when a specific disease was infrequently studied. HIV/AIDS (18.4%) and diarrhea-related dissertations (11.5%) were prominent topics addressed within the infectious and parasitic diseases category. Although not outlined in Table 14, nosocomial infections were also the subject of 5 dissertations. Of all infectious disease dissertations, 4 (4.6%) investigated multiple diseases. One dissertation studied two sexually transmitted diseases, gonorrhea and syphilis, two dissertations described infection with HIV and Hepatitis B or HTLV, and one described five distinct afflictions including tuberculosis and malaria. Because these dissertations investigated more than one infectious disease, the first of these abstracts was classified with the summary code of "sexually transmitted diseases" and the remaining three were classified simply as "infectious disease".

Within the neoplasms chapter, breast (13.9%) and lung cancer (11.6%) were prominent. Other types of cancer that were commonly investigated included leukemias and lymphomas (9.2%), cervical cancer (6.4%), brain cancer (5.2%) and colorectal cancer (4.0%). Of the 22 abstracts that studied multiple cancers, lung, colorectal and bladder cancer were the most common. Three distinct disorders of the blood were described in dissertations and included anemia, thalasemia, and sickle cell disease. Diabetes accounted for more than 75% of diseases classified under endocrine and metabolic disorders and included IDDM, NIDDM, and complications such as diabetic neuropathy and retinopathy. Of the dissertations examining mental disorders, smoking (34.5%) and depression (32.8%) were foremost. Depression was common among the 3 (5.2%) dissertations investigating multiple mental disorders, as were problems of substance abuse. Collectively, Alzheimer's

Table 13. Relative frequencies of research areas investigated in epidemiology dissertations, 1982-1992

	1982-83			1984-86			1987-89		
Research Topic Area	no.	(%)	(95% CI)	no.	(%)	(95% CI)	no	(%)	(95% CI)
Infectious and Parasitic Diseases	8	(7.4)	(2.5, 12.3)	13	(7.8)	(3.7, 11.8)	22	(10.8)	(6.5, 15.0)
Neoplasms	27	(25.0)	(16.8, 33.2)	49	(29.3)	(22.4, 36.2)	48	(23.5)	(17.7, 29.4)
Endocrine, Nutritional and Immunity Disorders	3	(2.8)	(-, 5.9)	7	(4.2)	(1.2, 7.2)	14	(6.9)	(3.4, 10.3)
Diseases of the Blood and Blood-Forming Organs	0	0	0	1	(0.6)	(-, 1.8)	0	0	0
Mental Disorders	8	(7.4)	(2.5, 12.3)	10	(6.0)	(2.4, 9.6)	13	(6.4)	(3.0, 9.7)
Diseases of the Nervous System	4	(3.7)	(0.1, 7.3)	8	(4.8)	(1.6, 8.0)	3	(1.5)	(-, 3.1)
Diseases of the Circulatory System	15	(13.9)	(7.4, 20.4)	17	(10.2)	(5.6, 14.8)	19	(9.3)	(5.3, 13.3)
Diseases of the Respiratory System	6	(5.6)	(1.2, 9.9)	6	(3.6)	(0.8, 6.4)	1	(0.5)	(-, 1.4)
Diseases of the Digestive System	4	(3.7)	(0.1, 7.3)	5	(3.0)	(0.4, 5.6)	5	(2.5)	(0.3, 4.6)
Diseases of the Genitourinary System	3	(2.8)	(-, 5.9)	4	(2.4)	(0.1, 4.7)	6	(2.9)	(0.6, 5.3)
Complications of Pregnancy and Childbirth	3	(2.8)	(-, 5.9)	8	(4.8)	(1.6, 8.0)	9	(4.4)	(1.6, 7.2)
Diseases of the Skin and Subcutaneous Tissue	0	0	0	0	0	0	0	0	0
Diseases of the Musculoskeletal System	3	(2.8)	(-, 5.9)	1	(0.6)	(-, 1.8)	4	(2.0)	(0.1, 3.9)
Congenital Anomalies	1	(0.9)	(-, 2.7)	3	(1.8)	(-, 3.8)	1	(0.5)	(-, 1.4)
Conditions Originating in the Perinatal Period	7	(6.5)	(1.8, 11.1)	10	(6.0)	(2.4, 9.6)	15	(7.4)	(3.8, 10.9)
Symptoms, Signs and Ill-Defined Conditions	1	(0.9)	(-, 2.7)	4	(2.4)	(0.1, 4.7)	3	(1.5)	(-, 3.1)
Injuries and Poisoning	4	(3.7)	(0.1, 7.3)	10	(6.0)	(2.4, 9.6)	21	(10.3)	(6.1, 14.5)
Mortality	3	(2.8)	(-, 5.9)	5	(3.0)	(0.4, 5.6)	1	(0.5)	(-, 1.4)
Multiple Diseases	8	(7.4)	(2.5, 12.3)	6	(3.6)	(0.8, 6.4)	19	(9.3)	(5.3, 13.3)
Total	108	(100)		167	(100)		204	(100)	

Table 13. Relative frequencies of research areas investigated in epidemiology dissertations, 1982-1992 (continued).

Research Topic Area	1990-92 no	1990-92 (%)	1990-92 (95% CI)	Total no	Total (%)	Total (95% CI)
Infectious and Parasitic Diseases	44	(15.8)	(11.5, 20.0)	87	(11.5)	(9.2, 13.7)
Neoplasms	49	(17.6)	(13.1, 22.0)	173	(22.8)	(19.8, 25.8)
Endocrine, Nutritional and Immunity Disorders	12	(4.3)	(1.9, 6.7)	36	(4.7)	(3.2, 6.3)
Diseases of the Blood and Blood-Forming Organs	2	(0.7)	(-, 1.7)	3	(0.4)	(-, 0.8)
Mental Disorders	27	(9.7)	(6.2, 13.1)	58	(7.7)	(5.8, 9.5)
Diseases of the Nervous System	7	(2.5)	(0.7, 4.3)	22	(2.9)	(1.7, 4.1)
Diseases of the Circulatory System	32	(11.5)	(7.7, 15.2)	83	(10.9)	(8.7, 13.2)
Diseases of the Respiratory System	10	(3.6)	(1.4, 5.8)	23	(3.0)	(1.8, 4.3)
Diseases of the Digestive System	11	(3.9)	(1.7, 6.2)	25	(3.3)	(2.0, 4.6)
Diseases of the Genitourinary System	6	(2.2)	(0.4, 3.9)	19	(2.5)	(1.4, 3.6)
Complications of Pregnancy and Childbirth	12	(4.3)	(1.9, 6.7)	32	(4.2)	(2.8, 5.7)
Diseases of the Skin and Subcutaneous Tissue	0	0	0	0	0	0
Diseases of the Musculoskeletal System	3	(1.1)	(-, 2.3)	11	(1.5)	(0.6, 2.3)
Congenital Anomalies	7	(2.5)	(0.7, 4.3)	12	(1.6)	(0.7, 2.5)
Conditions Originating in the Perinatal Period	24	(8.6)	(5.3, 11.9)	56	(7.4)	(5.5, 9.3)
Symptoms, Signs and Ill-Defined Conditions	5	(1.8)	(0.2, 3.3)	13	(1.7)	(0.8, 2.6)
Injuries and Poisoning	17	(6.1)	(3.3, 8.9)	52	(6.9)	(5.1, 8.7)
Mortality	6	(2.2)	(0.4, 3.9)	15	(2.0)	(1.0, 3.0)
Multiple Diseases	5	(1.8)	(0.2, 3.3)	38	(5.0)	(3.5, 6.6)
Total	279	(100)		758	(100)	

Table 14. Frequency of disease areas investigated in epidemiology dissertations, 1982-1992.

Disease Area	no.	(%)
Infectious and Parasitic Diseases	87	(100)
HIV/AIDS	16	(18.4)
Infectious Diarrhea	10	(11.5)
Sexually Transmitted Diseases	7	(8.0)
Malaria	6	(6.9)
Neoplasms	173	(100)
Breast Cancer	24	(13.9)
Lung Cancer	20	(11.6)
Leukemia and Lymphomas	16	(9.2)
Cervical Cancer	11	(6.4)
Brain Cancer	9	(5.2)
Colorectal Cancer	7	(4.0)
Endocrine, Nutritional, and Immunity Disorders	36	(100)
Diabetes	28	(77.8)
Mental Disorders	58	(100)
Smoking	20	(34.5)
Depression	19	(32.8)
Schizophrenia	5	(8.6)
Diseases of the Nervous System	22	(100)
Disorders of the Ear	5	(22.7)
Disorders of the Eye	4	(18.2)
Alzheimer's Disease	4	(18.2)

Table14. Frequency of disease areas investigated in epidemiology dissertations, 1982-1992 (continued)

Disease Area	no.	(%)
Diseases of the Circulatory System	83	(100)
Heart Disease	36	(43.4)
Hypertension	24	(28.9)
Stroke	9	(10.8)
Cardiovascular Disease (incl Heart Disease and Stroke)	55	(66.3)
Diseases of the Respiratory System	23	(100)
Respiratory Disease	18	(78.3)
Diseases of the Digestive System	25	(100)
Diarrhea (unspecified)	9	(36.0)
Dental Disease	9	(36.0)
Diseases of the Genitourinary System	19	(100)
Infertility	3	(15.8)
Pelvic Inflammatory Disease	3	(15.8)
Menstrual Disorders	3	(15.8)
Urinary Tract Infection (unspecified)	3	(15.8)
Complications of Pregnancy, Childbirth, Puerperium	32	(100)
Miscarriage/Spontaneous Abortion	10	(31.3)
Certain Conditions Originating in Perinatal Period	56	(100)
Adverse Pregnancy Outcomes	18	(32.1)
Low Birth Weight	17	(30.4)
Infant Mortality	11	(19.6)
Symptoms, Signs and Ill-Defined Conditions	13	(100)
Sudden Infant Death Syndrome	3	(23.9)
Urinary Incontinence	3	(23.9)
Apnea	3	(23.9)
Injuries and Poisoning	52	(100)
Suicide	4	(7.7)
Motor Vehicle Accidents	3	(5.8)
Head Injury	3	(5.8)

Disease, and various disorders of the eye and ear accounted for almost 59% of diseases of the nervous system

Heart disease (43.4%), hypertension (28.9%), and stroke (10.8%) were frequently included in dissertation research studying diseases of the circulatory system. Furthermore, total cardiovascular disease accounted for 66.3% of abstracts. Of the 3 dissertations including multiple outcomes, cardiovascular disease was investigated in each (i.e. stroke and various forms of heart disease). Almost 80% of dissertations examining diseases of the respiratory system were classified as unspecified respiratory disease, or acute and upper respiratory illnesses. Further detail describing these diseases was not possible with the information provided in the dissertation abstract. Among dissertations examining diseases of the digestive system, unspecified diarrhea (36.0%) and periodontal diseases (e.g. dental caries, enamel fluorosis) (36.0%) were prominent. Of this last group, two dissertations examined more than one type of dental disease. When diarrhea was grouped with infectious diarrhea, 2.5% of all dissertations included in this study were accounted for.

While the overall number of dissertations investigating diseases of the genitourinary system was low in this study, four particular diseases equally accounted for 63% of this category and included infertility, pelvic inflammatory disease, menstrual disorders and unspecified urinary tract infections. Of dissertations examining complications of pregnancy and childbirth, approximately one-third investigated miscarriages and spontaneous abortions. An additional 19% studied unspecified complications. Research activity concerning congenital anomalies represented less than 2% of dissertation research and included various malformations such as spina bifida, pyloric stenosis, cleft palate, and anencephaly. Similar anomalies were investigated in the three dissertations that examined more than one malformation. The primary perinatal condition described in epidemiology dissertations was a group of disorders classified as adverse pregnancy outcomes (32.1%). This group included various perinatal disorders common to the infant (e.g. low birth weight, neonatal complications, congenital anomalies etc.) and mother (e.g. ectopic pregnancy, pre-term delivery etc.). Single perinatal outcomes included low birthweight infants (30.4%) and infant mortality (19.6%).

A diverse group of dissertations classified within the ICD chapter of ill-defined conditions and disorders was encountered, and included research primarily dealing with sudden infant death syndrome (23.9%), urinary incontinence (23.9%), and apnea (23.9%) Similarly, the type of injuries examined in epidemiologic studies was diverse since the dissertation author described the injury according to the site or cause. Nevertheless, suicide (7.7%), head injuries (5.8%), and motor vehicle accidents (5.8%) were among the most commonly reported injuries and accidents. 34.6% of dissertations investigated occupational injuries occurring among pilots, painters, firefighters, farmers, nurses and other employees. In this group, injuries due to fire and falls were described, as well as hand and lower back injuries. Another 14% of dissertations investigated pediatric injuries, of which 1 specifically described eye injuries

Table 15 describes the distribution of study designs by ICD-9 chapters. Cross-sectional studies were often utilized when studying infectious diseases (41.4%) and endocrine and immunological disorders (38.9%). Case-control studies were more prominent in dissertation research examining various cancers (71.7%). Furthermore, this design accounted for almost half of the dissertation research describing diseases of the nervous system (50.0%), genitourinary system (47.4%), complications of pregnancy and childbirth (46.9%), musculoskeletal and connective tissue (54.5%), and congenital anomalies (50.0%) In comparison, cohort studies appeared to be the primary study design utilized in dissertations investigating diseases of the respiratory system (52.2%), overall mortality (73.3%) and multiple diseases (65.8%). Of the other study designs, family studies were used in two dissertations concerning cancer, one concerning congenital anomalies and one investigation of heart disease. Proportional mortality studies were used to investigate mortality from cancer and from other various diseases.

Table 15. Distribution of study designs by research topic areas in epidemiology dissertations, 1982-1992

Research Area *	Intervention Study		Cohort Study		Case-Control Study**		Cross-Sect Study		Ecological Study		Other Study		Cannot Classify		Total	
	no.	(%)	no.	(%)	no.	(%)	no.	(%)	no.	(%)	no.	(%)	no.	(%)	no.	(%)
1	4	(4.6)	22	(25.3)	14	(16.1)	36	(41.4)	1	(1.1)	1	(1.1)	9	(10.3)	87	(100)
2	1	(0.6)	21	(12.1)	124	(71.7)	8	(4.6)	9	(5.2)	3	(1.7)	7	(4.0)	173	(100)
3	5	(13.9)	7	(19.4)	7	(19.4)	14	(38.9)	2	(5.6)	0	0	1	(2.8)	36	(100)
4	1	(33.3)	0	0	0	0	2	(66.7)	0	0	0	0	0	0	3	(100)
5	9	(15.5)	16	(27.6)	4	(6.9)	17	(29.3)	1	(1.7)	0	0	11	(19.0)	58	(100)
6	2	(9.1)	4	(18.2)	11	(50.0)	5	(22.7)	0	0	0	0	0	0	22	(100)
7	5	(6.0)	28	(33.7)	14	(16.9)	23	(27.7)	7	(8.4)	1	(1.2)	5	(6.0)	83	(100)
8	0	0	12	(52.2)	2	(8.7)	5	(21.7)	0	0	0	0	4	(17.4)	23	(100)
9	0	0	7	(28.0)	5	(20.0)	8	(32.0)	2	(8.0)	0	0	3	(12.0)	25	(100)
10	1	(5.3)	5	(26.3)	9	(47.4)	3	(15.8)	0	0	0	0	1	(5.3)	19	(100)
11	1	(3.1)	11	(34.4)	15	(46.9)	3	(9.4)	0	0	0	0	2	(6.3)	32	(100)
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	(100)
13	1	(9.1)	1	(9.1)	6	(54.5)	3	(27.3)	0	0	0	0	0	0	11	(100)
14	0	0	0	0	6	(50.0)	2	(16.7)	1	(8.3)	1	(8.3)	2	(16.7)	12	(100)
15	1	(1.8)	16	(28.6)	11	(19.6)	13	(23.2)	6	(10.7)	0	0	9	(16.1)	56	(100)
16	2	(15.4)	3	(23.1)	4	(30.8)	3	(23.1)	1	(7.7)	0	0	0	0	13	(100)
17	2	(3.8)	8	(15.4)	14	(26.9)	17	(32.7)	5	(9.6)	0	0	6	(11.5)	52	(100)
18	0	0	11	(73.3)	2	(13.3)	1	(6.7)	0	0	0	0	1	(6.7)	15	(100)
19	0	0	25	(65.8)	0	0	4	(10.5)	4	(10.5)	1	(2.6)	4	(10.5)	38	(100)
Total	35	(4.6)	197	(26.0)	248	(32.7)	167	(22.0)	39	(5.1)	7	(0.9)	65	(8.6)	758	(100)

*See Appendix 12 for an outline of ICD Chapter Headings
** Includes nested case-control studies

4.5.1 Gender of the Dissertation Author and Research Topic Areas Investigated

Table 16 presents the distribution of research topic areas by the gender of the dissertation author. With the exception of maternal and infant health dissertations, there did not appear to be any real differences between male and female graduates. Men appeared to more frequently address research issues concerning diseases of the respiratory system (5.5% vs. 1.5%), however, this comparison was based on very small numbers, and therefore, is not likely reliable.

4.5.2 Type of Doctoral Degree and Research Topic Areas Investigated

The distribution of ICD chapters by doctoral degree was restricted to American dissertations. As shown in Table 17, PhD graduates appeared to examine cancer-related dissertations more frequently than other graduates (25.2% vs. 13.8%). However, this comparison may not be valid as a result of the limited reporting of other doctoral degree-types in the dissertations database (see section 5.2).

4.5.3 Research Areas Examined in Dissertations With Secondary and Tertiary Investigations

Secondary studies were described in 44 (5.8%) abstracts; one dissertation abstract described three distinct investigations. Of these studies, 51 research topic areas were of interest to dissertation authors (see Table 18), the most common of which included cancer (31.4%), infectious and parasitic diseases (11.8%), and diseases of the circulatory system (11.8%).

Table 16. Distribution of research topic areas by gender of dissertation author.

Research Topic Area	Male			Female			Total	
	no.	(%)	(95% CI)	no.	(%)	(95% CI)	no	(%)
Infectious and Parasitic Diseases	39	(14.4)	(10.2, 18.6)	38	(8.3)	(5.8, 10.9)	77	(10.6)
Neoplasms	64	(23.6)	(18.6, 28.7)	104	(22.8)	(19.0, 26.7)	168	(23.1)
Endocrine and Immunological Disorders	11	(4.1)	(1.7, 6.4)	25	(5.5)	(3.4, 7.6)	36	(5.0)
Diseases of Blood/Blood-Forming Organs	1	(0.4)	(-, 1.1)	1	(0.2)	(-, 0.6)	2	(0.3)
Mental Disorders	14	(5.2)	(2.5, 7.8)	43	(9.4)	(6.7, 12.1)	57	(7.8)
Diseases of the Nervous System	5	(1.8)	(0.2, 3.4)	17	(3.7)	(2.0, 5.5)	22	(3.0)
Diseases of the Circulatory System	39	(14.4)	(10.2, 18.6)	43	(9.4)	(6.7, 12.1)	82	(11.3)
Diseases of the Respiratory System	15	(5.5)	(2.8, 8.3)	7	(1.5)	(0.4, 2.7)	22	(3.0)
Diseases of the Digestive System	8	(3.0)	(0.9, 5.0)	16	(3.5)	(1.8, 5.2)	24	(11.3)
Diseases of the Genitourinary System	5	(1.8)	(0.2, 3.4)	14	(3.1)	(1.5, 4.7)	19	(2.6)
Complications of Pregnancy and Childbirth	4	(1.5)	(0.0, 2.9)	28	(6.1)	(3.9, 8.3)	32	(4.4)
Diseases of the Skin/Subcutaneous Tissue	0	0	0	0	0	0	0	0
Diseases of the Musculoskeletal System	3	(1.1)	(-, 2.4)	8	(1.8)	(0.5, 3.0)	11	(1.5)
Congenital Anomalies	1	(0.4)	(-, 1.1)	10	(2.2)	(0.8, 3.5)	11	(1.5)
Conditions Originating in the Perinatal Pd.	16	(5.9)	(3.1, 8.7)	37	(8.1)	(5.6, 10.6)	53	(7.3)
Symptoms, Signs, Ill-Defined Conditions	4	(1.5)	(0.0, 2.9)	8	(1.8)	(0.5, 3.0)	12	(1.7)
Injuries and Poisoning	19	(7.0)	(4.0, 10.1)	31	(6.8)	(4.5, 9.1)	50	(6.9)
Mortality	6	(2.2)	(0.5, 4.0)	6	(1.3)	(0.3, 2.4)	12	(1.7)
Multiple Diseases	17	(6.3)	(3.4, 9.2)	20	(4.4)	(2.5, 6.3)	37	(5.1)
Total	271	(100)		456	(100)		727	(100)

Table 17. Distribution of research topic areas by type of doctoral degree awarded

	Ph.D			Other			Total	
Research Topic Area	no.	(%)	(95% CI)	no	(%)	(95% CI)	no.	(%)
Infectious and Parasitic Diseases	69	(11.4)	(8.9, 14.0)	17	(13.1)	(7.3, 18.9)	86	(11.7)
Neoplasms	152	(25.2)	(21.7, 28.6)	18	(13.8)	(7.9, 19.8)	170	(23.2)
Endocrine and Immunological Disorders	30	(5.0)	(3.2, 6.7)	6	(4.6)	(1.0, 8.2)	36	(4.9)
Diseases of Blood/Blood-Forming Organs	1	(0.2)	(-, 0.5)	2	(1.5)	(-, 3.7)	3	(0.4)
Mental Disorders	46	(7.6)	(5.5, 9.7)	10	(7.7)	(3.1, 12.3)	56	(7.6)
Diseases of the Nervous System	20	(3.3)	(1.9, 4.7)	1	(0.8)	(-, 2.3)	21	(2.9)
Diseases of the Circulatory System	69	(11.4)	(8.9, 14.0)	10	(7.7)	(3.1, 12.3)	79	(10.8)
Diseases of the Respiratory System	17	(2.8)	(1.5, 4.1)	6	(4.6)	(1.0, 8.2)	23	(3.1)
Diseases of the Digestive System	17	(2.8)	(1.5, 4.1)	5	(3.8)	(0.5, 7.2)	22	(3.0)
Diseases of the Genitourinary System	15	(2.5)	(1.2, 3.7)	3	(2.3)	(-, 4.9)	18	(2.5)
Complications of Pregnancy and Childbirth	26	(4.3)	(2.7, 5.9)	5	(3.8)	(0.5, 7.2)	31	(4.2)
Diseases of the Skin/Subcutaneous Tissue	0	0	0	0	0	0	0	0
Diseases of the Musculoskeletal System	9	(1.5)	(0.5, 2.5)	2	(1.5)	(-, 3.7)	11	(1.5)
Congenital Anomalies	10	(1.7)	(0.6, 2.7)	1	(0.8)	(-, 2.3)	11	(1.5)
Conditions Originating in the Perinatal Pd.	38	(6.3)	(4.4, 8.2)	16	(12.3)	(6.7, 18.0)	54	(7.4)
Symptoms, Signs, Ill-Defined Conditions	11	(1.8)	(0.8, 2.9)	1	(0.8)	(-, 2.3)	12	(1.6)
Injuries and Poisoning	36	(6.0)	(4.1, 7.8)	13	(10.0)	(4.8, 15.2)	49	(6.7)
Mortality	12	(2.0)	(0.9, 3.1)	3	(2.3)	(-, 4.9)	15	(2.0)
Multiple Diseases	26	(4.3)	(2.7, 5.9)	11	(8.5)	(3.7, 13.2)	37	(5.0)
Total	604	(100)		130	(100)		734	(100)

Table 18. Research topic areas investigated in secondary and tertiary studies, 1982-1992

Research Topic Area	Secondary Investigation		Tertiary Investigations	
	no	(%)	no	(%)
Infectious and Parasitic Diseases	6	(11.8)	0	0
Neoplasms	16	(31.4)	0	0
Endocrine, Nutritional and Imm. Disorders	3	(5.9)	0	0
Diseases of the Blood	0	0	0	0
Mental Disorders	0	0	0	0
Diseases of the Nervous System	2	(3.9)	0	0
Diseases of the Circulatory System	6	(11.8)	0	0
Diseases of the Respiratory System	1	(2.0)	0	0
Diseases of the Digestive System	1	(2.0)	0	0
Diseases of the Genitourinary System	0	0	0	0
Complications of Preg. and Childbirth	5	(9.8)	0	0
Diseases of the Skin/Subcutaneous Tissue	0	0	0	0
Diseases of the Musculoskeletal System	0	0	0	0
Congenital Anomalies	1	(2.0)	1	(100)
Conditions Originating in Perinatal Period	4	(7.8)	0	0
Ill- Defined Symptoms and Signs	1	(2.0)	0	0
Injuries and Poisoning	1	(2.0)	0	0
Mortality	4	(7.8)	0	0
Multiple Diseases	0	0	0	0
Total	51	(100)	1	(100)

Note: Because the number of abstracts in this group was small, tabulation was permitted for each research area. As a result, more than one disease area was associated with several abstracts.

4.6 Type and Number of Study Subjects

When possible, the study subjects included in dissertation research were described by gender, age and ethnicity. As well, the geographic setting was summarized. To describe the number of subjects, the minimum, maximum and median number were determined for each study design.

Of the 758 dissertations, 5.4% could not be classified with respect to the gender of study subjects. In the remaining abstracts, 68.2% described subjects of both genders, 11.9% described males, and 19.9% described female subjects.

352 abstracts provided information describing the age of study subjects. Of these abstracts 28.1% included infants (birth through 12 months), 24.7% selected children (13 months through 17 years), 47.2% described adults between the ages of 18 and 44 years, 29.8% included subjects aged 45-65 years, and 13.9% included older adults aged 65 years and over. Because the age of study subjects often extended beyond one age category (e.g. 25 to 74 years), the total does not sum to 100%.

The ethnicity of study subjects was not reported in 63.9% of dissertation abstracts, and in the remaining 274, it was frequently necessary to classify ethnicity on the basis of geography. The breakdown of ethnicity in these abstracts was 61.7% White, 39.4% Black, 17.2% Hispanic, 10.2% Asian, and 1.5% Native Americans. Since dissertations may have reported more than one ethnic group, the total does not sum to 100%. Additionally, 25.6% of abstracts incompletely reported ethnicity. In these dissertations, the author did not describe the ethnic distribution of subjects. Instead ethnicity was reported when statistical significance was observed for a particular group.

The geographic setting of dissertations was described in 57.3% of abstracts. For abstracts without description, it was assumed that the study was carried out in the province or state of the affiliated institution. Furthermore, in 18 abstracts, more than one research setting was described, resulting in a total of 796 distinct geographic research centers. Of these settings, research was carried out in the United States (86.6%), Canada (3.5%), and in countries abroad (i.e. Africa, Asia, Central and South America, Mexico, Europe, Caribbean, and Australia) (9.9%).

In Canada, dissertation research most often occurred within the provinces of Ontario (50%) and Quebec (25%). In the United States, epidemiology dissertations were frequently carried out in California (15.2%), Maryland (10.6%), and North Carolina (8.9%). Dissertation research in other countries originated from American doctoral programs and were most common at the Universities of North Carolina, Texas, and

Tulane. The research area frequently addressed within these settings was infectious and parasitic diseases (34.7%), of which diarrhea and malaria were common.

The number of study subjects was described for dissertations with complete information and included intervention, cohort, case-control, nested case-control, and cross-sectional study designs. Because the number of "other" studies was small (see Table 4.3) and ecological studies included groups as the unit of analysis, the number of study subjects was not summarized. Similarly, the number of subjects was not described for designs categorized as "cannot classify". As illustrated in Table 19, the median number of study subjects was greatest in cohort studies and smallest in intervention studies. Of the case-control studies carried out (n=248), 13 included two or more case-groups, and 25 included 2 or more control groups. The type of subjects identified as controls included population, hospital and clinic, family, neighborhood and deceased subjects.

Table 19. Minimum, maximum and median number of study subjects included in epidemiology dissertations

Study Design	Minimum	Maximum	Median
Intervention Study (n=28)	24	2,021	93.5
Cohort Study (n=162)	18	270,464	854.5
Case-Control Study (186)*	60	9,418	420
Cross-Sectional Study (n=123)	29	375,767	534
Total (n=499)	18	375,767	509

CHAPTER 5 - DISCUSSION

5.1 Introduction

Overall, the results from this study indicate that observational studies have been predominant in dissertation research, with the case-control study being carried out most often. Among the research areas investigated, epidemiology dissertations have addressed diseases of greatest concern to society (e.g. cancer, cardiovascular disease, AIDS), and reflect research activity in other significant areas like mental disorders, injuries, and maternal and infant health. This chapter will discuss the implications of these results, review the limitations of the study methods and provide recommendations for structured abstracts and future research.

5.2 Epidemiology Dissertations

In this study, dissertation abstracts were unevenly distributed between Canadian and American universities because the number of epidemiology programs within the United States far exceeded the number of programs in Canada (Bernier and Mason 1991). As a result, American abstracts accounted for more than 95% of dissertations.

Dissertations originating in the United States increased in frequency throughout the eleven-year study period, and this finding is consistent with an increasing number of doctoral graduates reported by the Association of Schools of Public Health (ASPH) (Levin 1993). A similar pattern was not evident for Canada as a result of the small number of dissertations available for study.

In Canada, the Ph.D. is awarded to graduates of doctoral programs in epidemiology (Bernier and Mason 1991; Association of Universities and Colleges of Canada; 1993), while in the United States, one of three doctoral degrees may be awarded depending on the institution. The Ph.D. is most common among graduates of schools of public health (Levin 1993), and was well represented in this study. The Sc.D. and Dr.P.H. degrees were under-reported. Three schools award the Sc.D. degree in epidemiology and include Boston (e.g. Sc.D.), Harvard (e.g. Sc.D. and Dr.P.H.) and the Johns Hopkins

University (e.g. Ph.D., Sc.D. and Dr.P.H.) (Bernier and Mason 1991). The low frequency of this degree-type may be explained, in part, by the absence of research information from Harvard, in addition to the inconsistent reporting practices by the Johns Hopkins University. More specifically, the Ph.D. program is administered by a university graduate board and students completing this program are required to submit dissertations to UMI. Graduates completing other degree programs are not obligated to submit dissertations since these programs are administered separately by the school of public health (S. Eichner, Johns Hopkins School of Public Health and Hygiene, Baltimore, Maryland 1994).

5.3 Study Designs

Study designs were identified in the dissertation abstract and classified using a classification scheme developed for this project. For this purpose, information from existing taxonomies was combined to reflect the principal research methods used in epidemiology (see Table 2). Although we tried to ensure that the categories were mutually exclusive, there is some overlap between cross-sectional and ecological studies. Distinguishing between these particular designs was difficult when the information presented in the abstract was unclear.

To evaluate the reliability of this taxonomy, study designs were classified by each of two investigators using a subsample of abstracts. The results revealed strong agreement (kappa=0.75). Most disagreement was due to the assignment of "cannot classify". Intra-rater agreement was likewise measured, and excellent agreement was observed (kappa=0.85).

5.3.1 Classification of Study Designs - Limitations

Six percent of dissertation abstracts reported more than one study. When this occurred, the predominant investigation was classified and reported in the overall results. Secondary and tertiary studies were reviewed separately. We followed this procedure for two reasons. The reporting of information in the less predominant of studies was often unclear, and it was apparent in several of the abstracts, that these studies were carried out to assist with the design and implementation of the primary investigation. As a result, the

dissertation author tended to emphasize the methods and results of the primary study, while very briefly describing the less predominant investigation. While this was not the case for all of these abstracts, to remain consistent we reported the predominant study throughout. Because the number of abstracts in this group was small, the results would not have not been different if we had allowed more than one study design per abstract.

Second, we were unable to classify the study design in roughly 9% of abstracts because the information described within was insufficient or ambiguous. Of this group, cross-sectional and ecological studies were most often suspected. To the extent that these designs were represented in this group of abstracts, the frequency of cross-sectional and ecological studies has been under-reported. We do not believe that our proportion of unclassified abstracts is unreasonable when compared to a similar study using both abstracts and journal articles. Ruiz et al (1991) were unable to classify 6% of articles appearing in rheumatology journals. Further difficulties were encountered when classifying cohort studies. Twenty-five percent of these abstracts could not be characterized as prospective or retrospective.

5.3.2 Frequency of Study Designs - Main Findings

Our results show that observational studies were frequently used in epidemiology dissertations. The infrequent use of intervention studies, and more specifically, randomized clinical trials is consistent with reports from the general medical (Fletcher and Fletcher 1979; Alvarez-Dardet et al. 1985) and specialties literature (Geyman and Berg 1984; Fromm and Snyder 1986; Marvel et al. 1991; Schubert- Subbarathnam and Rougemont et al. 1992), but also may reflect the ethical constraints that limit its feasibility in epidemiology (Rothman 1986). Although variations over time were observed throughout the study period, we were unable to detect any trends. However, the study period defined in this project was relatively short when compared to the 20-year period examined by Cole (1979), and the 30-year perspective assessed by Fletcher and Fletcher (1979).

When grouped with nested case-control studies, the case-control design was the most common design utilized in epidemiology dissertations Historically, case-control methods have been refined (Lilienfeld and Lilienfeld 1979), and not surprisingly its use in the general medical literature has increased substantially (Cole 1979). The strengths of this design have been reported elsewhere (Kleinbaum, Kupper and Morgenstern 1982; Kelsey, Thompson and Evans 1986; Rothman 1986), however, it may best be described by its effectiveness in studying rare diseases and for its efficiency in measuring the effects of numerous exposure factors.

In addition to the case-control study, the cohort study was well-represented in epidemiology dissertations, accounting for roughly one-quarter of the research. Unlike a previous study (Fletcher and Fletcher 1979), there was no evidence of a trend over time. The cohort study has certain similarities with intervention studies, and therefore, is effective in establishing cause and effect (Kleinbaum et al. 1982; Rothman 1986).

Cross-sectional studies have been widely reported in both the general medical (Feinstein 1978) and specialties literature (Fromm et al. 1986, Schubert et al. 1992), and in one review was shown to increase significantly over time (Fletcher and Fletcher 1979). The use of this design was likewise popular in dissertations In epidemiology, cross-sectional studies are used to determine the prevalence of a given disease, and are helpful in identifying risk factors for further study (Kelsey et al. 1986). Yet, unlike the cohort and case-control study, the cross-sectional study is unable to determine whether cause preceded effect since both exposure and outcome are measured simultaneously and at a single point in time (Kelsey et al. 1986; Rothman 1986).

5.3.3 Study Designs by Gender and Doctoral Degree

To determine whether research practices varied by gender and type of doctoral degree, we compared the distribution of study designs between male and female authors, and, between Ph.D. and other doctoral degrees. Generally, the use of study designs was similar between men and women.

With respect to the type of doctoral degree, the case-control study was more common among Ph.D. dissertations. Cross-sectional studies were used more frequently among graduates of other doctoral degree programs. While these comparisons are interesting, the results may not be reliable since Dr.P.H. and Sc.D. degrees were under-reported.

5.4 Research Topic Areas

In previous studies, bibliometric analyses have relied upon the subject categories outlined in the study journals to categorize the research activity (Dannenberg 1985, Schubert et al. 1992), and some have defined groups to reflect common disease areas (LaPorte and Cresanta 1980, Woernle et al. 1991). Still, other studies have based their classification of epidemiologic activity upon established programs within the state health departments (Gunn et al. 1989), and by risk factor and disease area (Boss and Foster 1994). In this study however, a working definition of epidemiology was established at the onset to include "disease-specific" abstracts. For this reason, the ICD-9 (CM) was selected to classify the disease areas described in epidemiology dissertations.

5.4.1 Research Topic Areas - Limitations

The research topic areas were classified with as much detail as that provided by the dissertation authors. However, we encountered several difficulties. First, 2% of abstracts described "overall mortality" as the outcome of interest, and 5% described more than one study disease, each of which extended beyond one ICD chapter. To more accurately classify these abstracts, we created two categories in addition to the seventeen chapters outlined within the ICD (i.e. mortality, multiple diseases). In this last group, the distinct disease areas were each recorded to allow for description; however, the overall research area was summarized as "multiple diseases", and, it was this category that was included in the overall tabulations. Because only 5% of abstracts were included in this group, we do not feel that the limitations imposed by this classification have substantially effected the overall results. Instead, the classification of these abstracts was simplified, and each research area was associated with one abstract.

Second, we created one final code in which to classify abstracts that described adverse outcomes of pregnancy. Although the disease areas included within this group extended beyond one ICD chapter, they were similar to one another and reflect the overlap among categories defined by the ICD (i.e. complications of pregnancy and childbirth and conditions originating in the perinatal period). Specifically, abstracts in this group described several perinatal outcomes which came from these different categories and were related to both mother and child. Notwithstanding these difficulties, there was strong agreement (percent agreement=89.5%) between the two investigators when classifying the research topic areas, and intra-rater agreement was found to be stable over time (percent agreement=97.4%).

5.4.2 Research Topic Areas - Main Findings

In terms of the research topic areas, our results indicated that a variety disease areas were addressed in epidemiology dissertations. Yet, established research areas like cancer, infectious disease and diseases of the circulatory system (i.e. cardiovascular disease) were common. When the appropriate ICD chapters were collapsed, maternal and infant health dissertations were also well-represented. A relatively small proportion of dissertations accounted for research areas concerned with congenital anomalies, and diseases of the blood and musculoskeletal systems.

Among the leading causes of death in Canada and the United States, cardiovascular disease and cancer rank first and second (Canada Yearbook 1985, 1987, 1989, 1991; National Center for Health Statistics, 1993). The significance of these health problems has been reflected in the medical literature (LaPorte and Cresanta 1980, Dannenberg 1985; Richards and Baker 1988), and clearly, in epidemiology dissertations as well. Of these research areas, cancer was the most common, accounting for more than 20% of dissertations. Within this category, breast and lung cancer were prominent.

Overall, infectious disease research represented almost 12% of dissertations. AIDS and HIV dissertations were not represented until the second half of the study period and explain in part, the large number of infectious disease dissertations during 1990-92. Approximately one-third of this research activity was accounted for by HIV and AIDS dissertations.

5.4.3 Study Designs by Research Areas

In epidemiology, the use of a particular design is dependent upon several factors (e.g. relative frequency of study disease, availability of study information etc.). Clearly, a prospective cohort study is an inefficient design for studying a rare disease (Rothman 1986). To determine how the study designs varied by the research areas studied, we examined the distribution of study designs by ICD-9 chapters. The most striking finding was observed for neoplasms; more than 70% of these dissertations reported using a case-control study. Of the other disease areas frequently studied, cross-sectional studies were common in dissertations describing infectious and parasitic diseases.

5.4.4 Research Areas by Gender and Doctoral Degree

To determine whether the research areas varied by gender and degree type, we compared the distribution of ICD-9 chapters between male and female authors, and between Ph.D. and other doctoral degrees. Maternal and infant health problems appeared to be more common among female graduates.

In terms of the doctoral degree, Ph.D. degrees were more common among dissertations investigating neoplasms. However, because of the under-reporting of other doctoral degree-types, this finding may not be valid.

5.5 Type and Number of Study Subjects

Information describing the age and ethnicity of study subjects was poorly reported in dissertation abstracts. Often, the author did not provide a description of the study sample, and in some cases, information was incomplete (e.g. "an increase in risk was evident for black subjects" etc.). Furthermore, in several abstracts, it was necessary to

use geography to classify the predominant ethnic group. As a result, this variable could not be classified with precision. Considering these shortcomings, the quality of this information is questionable. The abstract is often used by researchers as a primary source of information in which to assess the external validity of the study. Therefore, the study sample must be adequately described for effective communication.

Despite these limitations, dissertation research targeted both male and female subjects, and frequently included individuals ranging in age from 18 to 64 years. Of dissertations reporting ethnicity, White subjects were described most frequently. Dissertation research was commonly based on subjects from the provinces and states in which the doctoral degree was completed. However, approximately 10% of dissertation research occurred in settings beyond Canada and the United States. The number of subjects varied remarkably within study designs. For cohort and cross-sectional studies, this variability may be explained in part, by the use of data from large birth and death files, and from existing datasets like the Framingham and MRFIT studies.

5.6 Study Considerations

First, the selection of American dissertations was restricted to epidemiology programs within the schools of public health. We excluded other programs since "demographic" data from the Association of Schools of Public Health (ASPH) was available to assist with the inclusion of relevant dissertation abstracts (Levin 1993). Because most graduates are from these institutions (Williams et al. 1988), we do not feel that this restriction has influenced the results significantly.

Second, epidemiology can be seen as a discipline bordering on and overlapping with many others (e.g. medicine, public health, sociology/demography, biostatistics, microbiology etc.). We have considered it to be a medical discipline, and for the purposes of this study, it was assumed that dissertations would be classified accordingly. Nevertheless, a search of the database for "Health Science" dissertations precluded the selection of epidemiology abstracts simply classified, as "Sociology" or "Psychology". To the extent that these dissertations were excluded, the total number of abstracts eligible for inclusion in this study may be considered to be underestimated.

Third, an operational definition of epidemiology was established and applied to each abstract in the event that dissertations reflected research from disciplines other than epidemiology. Generally, the kappa statistics indicated fair to good agreement (Fleiss 1981) with respect to the content of abstracts, however, a value of .47 suggests room for improvement in the training and application of the working definition. A more suitable procedure would be to include dissertations after a consensus is reached among a panel of two or more individuals. We considered this option in the present study, however, the number of abstracts remaining after the first inspection was too large (n=2513) to be practically managed.

Certainly, the definition of epidemiology would be simplified further if the names of graduates were readily available from each university and the database were searched accordingly. However, to attempt this type of data collection is both costly and time consuming and cannot be justified without a working knowledge of the dissertations database. Furthermore, regardless of the search strategy, the unit of observation remains the dissertation abstract. For this reason, the quality of the dissertations database was the primary concern.

5.6.1 Quality of the Dissertations Database

To evaluate the database, quality was defined according to (1) the organization and management of the database, and, (2) the communication of research information presented in the dissertation abstract.

In terms of coverage, the database was limited by the absence of dissertation abstracts from Harvard University, and the incomplete reporting of doctoral degrees other than the Ph.D. For this reason, it is recommended that UMI consolidate dissertation abstracts from each participating university in CD-ROM format and ensure that common reporting practices are consistently followed.

Furthermore, if a subject area for epidemiology were created and included under the heading of "Health Sciences", the selection and retrieval of dissertations would be facilitated in future studies. If this were the case, researchers may be more inclined to use the database for regular bibliometric analyses. Another suggestion to assist with the

selection of epidemiology dissertations, is to include within the bibliographic citation, the name of the university department from which the dissertation originated (e.g. epidemiology and biostatistics). This information is presently collected by UMI, but it is limited to accounting records (Delphine Lewis, UMI, Michigan 1994). This addition would make the dissertations database more consistent with Medline, which currently includes the address of authors.

As a result of insufficient or ambiguous information, roughly 9% of abstracts included in this study could not be classified according to the study design. Moreover, information describing the type and number of study subjects was often incomplete. For these reasons, it is recommended that doctoral recipients prepare structured abstracts. In doing so, authors will be reminded to include information describing the number and type of subjects upon which the results are based. In addition, precise documentation of the study design will facilitate classification, and assist researchers in interpreting the results (Bailar et al., 1984; Maclure 1991). In this study, less than 45% of dissertation authors explicitly stated the study design. Structured abstracts have been adopted by several journals (e.g. Canadian Medical Association Journal, American Journal of Public Health etc.) and guidelines have been published (Haynes et al. 1990). An example of a structured abstract prepared by a graduate from the University of Washington is presented in Appendix 12. Dissertation authors may also consider including "epidemiology" as a key word within the abstract when submitting their dissertation to UMI.

Notwithstanding these limitations, the strengths and advantages of the database are obvious. UMI-assigned search fields and codes permitted the user to quickly and efficiently identify and retrieve relevant research information, and with the simplicity and speed of CD-ROM technology data collection was further enhanced. Our study would have been much more cumbersome to perform if we had not had access to the electronic database, but rather were restricted to the printed version. Most importantly, an existing, unobtrusive and cost-effective source of information was available in which to address the research questions defined in this project.

5.7 Future Research

In carrying out this study, we identified several research areas requiring further investigation. First, to provide a more comprehensive review of dissertation research, we recommend that a similar study be carried out broadening the definition of epidemiology, and including dissertations originating from each American epidemiology program and from the Universities of British Columbia and Manitoba (i.e. once graduates become available).

Second, epidemiologic methods are being increasingly applied in other health disciplines (e.g. health promotion, health services research, program evaluation etc.). For this reason, a bibliometric study using the dissertations database may provide valuable information describing the extent to which epidemiology has been applied.

Third, epidemiology is a discipline that has as one of its bases, the techniques and methods of biostatistics. The use of these procedures may be tracked within dissertations to determine not only the amount of use, but also the specific research application. A similar study was carried out in the general medical literature to determine the frequency of multivariable analysis (Concato et al. 1993), and in the *American Journal of Epidemiology*, the practice of statistical significance testing and presentation of confidence intervals was described (Savitz et al. 1994).

Fourth, in the present study, we did not attempt to formally measure the quality of epidemiology dissertations. For this purpose, a sound knowledge of each content area would be required (Brady et al. 1988), and one can question if the quality of the dissertations can be judged based on the abstract only. However, to globally assess the quality of dissertations in public administration, Cleary (1992) applied six evaluative criteria to abstracts selected from within the dissertations database. Administrators of epidemiology programs may be interested in pursuing a similar study.

Finally, epidemiologists may be interested in the more general applications of the database. Prospective doctoral candidates may find it helpful to review dissertations completed under the supervision of faculty members sharing similar research interests. The names of dissertation advisors were made available in the bibliographic citation

beginning in July of 1988 (UMI 1989). In addition, the database may be reviewed to help define research questions or identify solutions to research problems.

5.8 Conclusions

In this study, we described the central characteristics of epidemiology dissertations, namely, study design and research topic area. Our review of dissertation research was limited to theses examining illness outcomes. Observational studies were predominant in epidemiology dissertations. Case-control, cohort and cross-sectional studies were well-represented in dissertation research. However, the case-control design was more common when grouped with the nested case-control study. A diversity of research areas were addressed in dissertation research. Neoplasms, maternal and infant health disorders, infectious and parasitic diseases and diseases of the circulatory system were among the research topic areas that were commonly investigated.

To further examine epidemiology dissertations, we have suggested several potential research projects. At the center of these studies is the dissertations database. *DAI* provides a valuable source of information in which to examine research patterns and trends within epidemiology, and, with the advent of CD-ROM technology, the time- and cost-effectiveness of the database have been significantly enhanced. To further improve the utility of this database, we have made several recommendations regarding coverage, indexing and structured abstracts. With the adoption of these suggestions, researchers will be able to more quickly and efficiently monitor the developments in and progress of epidemiology as reflected in doctoral dissertations.

APPENDIX 1. Permission to Reprint

UMI

A Bell & Howell Company

August 17, 1993

Deborah L. Morin, MSc Candidate
Department of Epidemiology & Biostatistics
Kresge Building
The University of Western Ontario
London, Ontario
CANADA N6A 5C1

Dear Ms. Morin,

We are pleased to grant you permission to use dissertation abstracts and citations found in Dissertation Abstracts International for your thesis.

Information about the appropriate rights and permissions is enclosed. As long as the citations and abstracts are not altered in any way, we are happy to grant you permission to reprint the requested information. Please state that full text copies are available from University Microfilms International. We would like to request the inclusion of our UMI order number as well.

Please feel free to contact me on our toll-free number, 1-800-521-0600, extension 3708, if I can be of further assistance.

Sincerely,

Candace C. Wise
Manager, Database Licensing

Enclosure

Appendix 1. Permission to Reprint (continued)

Rights and Permissions

Dissertation Abstracts International

Please include the following statements somewhere in the section in which the abstracts are located:

"The dissertation titles and abstracts contained here are published with permission of University Microfilms Inc., publishers of *Dissertation Abstracts International* (copyright © _____*
by University Microfilms Inc.), and may not be reproduced without their prior permission."

You may also wish to include information about the availability of full-text copies:

"Copies of the dissertations may be obtained by addressing your request to:

University Microfilms Inc.
300 North Zeeb Road
Ann Arbor, Michigan 48106

or by telephoning (toll-free) 1-800-521-3042."

U·M·I

*Insert correct copyright year for materials used.

APPENDIX 2. Sample Record

Order No: AAC 9000304 ProQuest - Dissertation Abstracts
Title: A CASE-CONTROL STUDY OF ORAL CANCER AND PREDIAGNOSTIC CONCENTRATIONS OF SELENIUM, IRON, ZINC, AND CALCIUM IN NAIL TISSUE (CANCER)
Author: ROGERS, MARY A. M.
School: UNIVERSITY OF WASHINGTON (0250) Degree: PHD Date: 1989
pp: 144
Advisor: THOMAS, DAVID B.
Source: DAI-B 50/08, p. 3428, Feb 1990
Subject: HEALTH SCIENCES, PUBLIC HEALTH (0573); HEALTH SCIENCES, NUTRITION (0570); HEALTH SCIENCES, MEDICINE AND SURGERY (0564)

Abstract: A case-control study of the relationships between nail tissue levels of selenium, iron, zinc and calcium and cancer of the oral cavity was conducted. Oral cancer cases were identified through a population-based registry. Only individuals who were diagnosed from 9/83 through 2/87, were between the ages of 20 and 74, and who resided in King, Pierce and Snohomish counties of Washington State were included. Controls were selected by telephone using random digit dialing and were frequency matched by sex and age-group to the cases. Subjects participated in a personal interview and submitted clippings from the nails of each great toe. Concentrations of selenium, iron, zinc and calcium were measured in the nail samples by neutron activation analysis. Male oral cancer cases had lower selenium levels than did the controls, but femal cases did not. The association between low selenium concentrations and oral cancer was particularly strong in young men. For those individuals who had a previous history of low ascorbic acid intake, low selenium levels were positively associated with oral cancer in both sexes. Cases were more likely to have low zinc levels than controls, while there was suggestive evidence that controls had lower calcium levels than cases. There was no association between low or high levels of iron and oral cancer. However, individuals with high iron concentrations accompanied by low zinc levels were 4 times more likely to have oral cancer than subjects with low iron and high zinc levels. There was a positive association between tonsillar carcinoma and low concentrations of selenium and iron. Stratification by time interval from diagnosis to interview suggested that the differences in mineral content of the nail tissue in cases and controls were a result of the disease, rather than related to etiology.

APPENDIX 3. Sampling Frame

University	UMI Code	Doctoral Degrees in Epidemiology	Participation With UMI
Canadian Schools [1]			
University of Guelph	0081	Ph.D	1968
Memorial University of Newfoundland	0306	Ph.D	1970
McGill University	0781	Ph.D	1968
University of Montreal	0992	Ph.D	unavailable
University of Toronto	0779	Ph.D	1962
University of Western Ontario	0784	Ph.D	1971
American Schools [2]			
University of Alabama at Birmingham	1039 0005	Ph.D., Dr.P.H.	1970
Boston University	0017 0822	Sc.D	1957
University of California at Berkeley	0028	Ph.D., Dr.P.H.	1962
University of California at Los Angeles	0031	Ph.D., Dr. P.H.	1962
Columbia University	0054	Ph.D., Dr. P.H.	1952
Emory University	0665	Ph.D.	1958
Harvard University	0084	Dr.P.H., Sc.D.	1982
University of Hawaii	0085	Ph.D., Dr.P.H.	1960
University of Illinois at Chicago	0793 0806	Ph.D.	1981 1954
Johns Hopkins University	1046 0098	Ph.D., Dr.P.H., Sc.D.	1958
Loma Linda University	0106	Dr.P.H.	1958
University of Massachusetts	2389 0118	Ph.D.	1961
University of Michigan	0766 0127	Ph.D., Dr.P.H.	1962 1952

APPENDIX 3. Sampling Frame (continued)

University	UMI Code	Doctoral Degrees in Epidemiology	Participation With UMI
University of Minnesota	0130	Ph.D.	1976
University of North Carolina at Chapel Hill	0153	Ph.D., Dr P.H	1957
University of Oklahoma	0361	Ph.D., Dr P.H.	1975
	0169		1954
University of Pittsburgh	0178	Ph.D., Dr P.H.	1953
San Diego State University	0220	Ph.D.	unavailable
	0385		
University of South Carolina	0202	Ph.D., Dr P.H	1953
University of South Florida	0206	Ph.D.	1972
University of Texas at Houston	0219	Ph.D., Dr.P.H	1972
Tulane University	0325	Ph.D., Dr.P.H.	1974
	0225		1957
University of Washington	0250	Ph.D.	1952
Yale University	0265	Ph.D., Dr.P.H.	1963

Source:

[1] Association of Universities and Colleges of Canada. Directory of Canadian Universities. Ottawa, Ontario. 1991 and Association of Universities and Colleges of Canada. 1992-1994 Directory of Canadian Universities. 29th edition. Ottawa, Ontario. 1993.

[2] Bernier RH, Mason VM, editors. Episource. A Guide to Resources in Epidemiology. Roswell, GA Epidemiology Monitor, 1991: 405-12.

[3] University of Alabama at Birmingham School of Public Health Catalog 1990-1992, Volume 23, No 4. Birmingham, Alabama, Fall 1990.

[4] University of Texas School of Public Health at Houston 1991-1993 Catalog Houston, Texas, 1993

[5] American Council on Education. American Universities and Colleges, 14th edition. New York: Walter de Gruyter, Inc., 1992.

[6] U.S. Schools of Public Health and Graduate Programs Accredited by the Council on Education for Public Health. *Am J Public Health* 1993; 83(3): 447-8 and U.S. Schools of Public Health and Graduate Programs Accredited by the Council on Education for Public Health. *Am J Public Health* 1994; 84(4): 694-5.

[7] University Microfilms Inc. DAI Cumulative Author Index. Volume 45 Ann Arbor, MI, 1985

APPENDIX 4. Subject Areas Classified As "Health Sciences"

Subject Area	UMI Subject Code
Audiology	0300
Chemotherapy	0992
Dentistry	0567
Education	0350
General	0566
Hospital Management	0769
Human Development	0758
Hygiene	0568
Immunology	0982
Medicine and Surgery	0564
Mental Health	0347
Nursing	0569
Nutrition	0570
Obstetrics and Gynecology (new field in 1991)	0380
Occupational Health and Therapy (new field in 1981)	0354
Opthamology (new field in 1991)	0381
Pathology	0571
Pharmacology	0419
Pharmacy	0572
Physical Therapy (new field in 1991)	0382
Public Health	0573
Radiology	0574
Recreation	0575
Speech Pathology	0460
Toxicology (new field in 1991)	0383

Source: University Microfilms Inc. User Guide. Dissertation Abstracts Online. How to Use the Online Dissertation Database Step-by-Step. Ann Arbor, MI., 1991.

APPENDIX 5. Number of Dissertation Records Included

Universities	Number Retrieved	No. Included in First Pass	No. Included in Second Pass
Canadian Universities			
University of Guelph*	76	5	0
McGill University	176	25	7
Memorial University of Newfoundland	29	4	3
University of Montreal	3	0	0
University of Toronto	268	55	11
University of Western Ontario	33	10	3
Sub-Total	585	99	24
American Schools of Public Health			
University of Alabama (1039)	19	19	14
University of Alabama (0005)	307	30	11
Boston University (0017)**	247	28	1
Boston University Graduate School (0822)	7	0	0
University of California at Berkeley	305	165	59
University of California at Los Angeles	369	176	48
Columbia University	186	74	26
Emory University***	83	3	1
Harvard University	237	0	0
University of Hawaii	71	36	9
University of Illinois (0793)	7	0	0
University of Illinois (0806)	317	155	23
Johns Hopkins University (1046)	6	6	0
Johns Hopkins University (0098)	282	154	91
Loma Linda University	37	21	8
University of Massachusetts (2389)*	0	0	0
University of Massachusetts (0118)*	144	37	5
University of Michigan SPH (0766)	16	13	2
University of Michigan (0127)	474	205	41
University of Minnesota	483	185	38

APPENDIX 5. Number of Dissertation Records (continued).

Universities	Number Retrieved	No. Included in First Pass	No. Included in Second Pass
University of North Carolina at Chapel Hill	474	248	84
University of Oklahoma (0361)	130	55	18
University of Oklahoma (0169)	28	16	0
University of Pittsburgh	413	257	55
San Diego State University (0202)""	1	0	0
University of San Diego and San Diego State Univ. (0385)""	4	0	0
University of South Carolina	151	90	16
University of South Florida"""	54	4	0
University of Texas at Houston	227	186	74
Tulane University (0325)	2	1	0
Tulane University (0235)	49	23	12
University of Washington	325	163	61
Yale University	132	64	37
Sub-Total	5587	2414	734
TOTAL	**6172**	**2513**	**758**

* The doctoral program in epidemiology at the University of Guelph was not established until 1987.[1]

** Boston University SPH did not award any doctoral degrees prior to the 1986/87 academic year.[2]

*** Emory University SPH was not in operation until the 1990/91 academic year, and there were no doctoral degrees awarded at this time or during the 1991/92 time period. However, dissertation abstracts dated 1992 were considered to account for the first half of the 1992/93 academic year and were subject to review in the second inspection.[2]

\# University of Massachusetts SPH did not award any doctoral degrees until 1984-85.[2]

\## San Diego State University did not award doctoral degrees during the study period. However, dissertation abstracts dated 1992 were considered to account for the first half of the 1992/93 academic year and were subject to review in the second inspection.[2]

\### University of South Florida SPH was not in operation until 1986-87, and there were no doctoral degrees awarded between 1986 and 1992. However, dissertation abstracts dated 1992 were considered to account for the first half of the 1992/93 academic year and were subject to review in the second inspection.[2]

APPENDIX 5. Number of Dissertation Records (continued).

Source:

[1] Canadian Society of Epidemiology and Biostatistics News. November 1994.

[2] Levin M. Data Report on Applicants. New Enrollments and Students. Fall 1992 and Graduates and Expenditures. 1991-92 with Trends analysis for 1974-1975 through Fall 1992. Washington. D.C. Association of Schools of Public Health. 1993.

APPENDIX 6. Definitions of Epidemiology

(1) Last JM, editor. A Dictionary of Epidemiology, 2nd ed. New York: Oxford University Press, Inc., 1988.

- epidemiology is the study of the distribution and determinants of health-related states or events in specified populations, and the application of this study to control of health problems.

(2) Holland WW, Detels RG, Knox G, Fitzsimmons B, Gardner L, editors. Oxford Textbook of Public Health. 2nd ed. Volume 2. Methods of Public Health. New York: Oxford University Press 1991.

- epidemiology represents a philosophical method of studying a health problem and can be applied to a wide range of problems, ranging from transmission of an infectious disease agent to the design of a new strategy of health-care delivery.

- objectives of epidemiology:

  (a) describe the spectrum of disease
  (b) describe the natural history of disease
  (c) identify factors which increase the risk of acquiring disease
  (d) predict disease trends
  (e) test the efficacy of intervention strategies
  (f) elucidate mechanisms of disease-transmission
  (g) evaluate intervention programs (public health efficacy)
  (h) identify the health needs of a community
  (i) evaluate public health programs

(3) Kleinbaum DL, Kupper LL, Morgenstern H. Epidemiologic Research. Principles and Quantitative Methods. California: Lifetime Learning Publications, 1982.

- the study of health and illness in human populations

- objectives of epidemiologic research:

  (a) describe the health status of populations by enumerating the occurrence of disease, obtaining the relative frequencies within groups, and discovering important trends
  (b) explain the etiology of diseases by determining factors that cause specific diseases or trends and by discovering modes of transmission

APPENDIX 6. Definitions of Epidemiology (continued)

(c) predict the number of disease occurrences and the distribution of health status within populations

(d) control the distributions of diseases in the population by prevention of new occurrences, eradication of existing cases, prolongation of life with the disease, or otherwise improving the health status of afflicted persons

(4) Mausner JS, Kramer S. Epidemiology- An Introductory Text. Philadelphia: W.B. Saunders Company, 1985.

- the study of the distribution and determinants of diseases and injuries in human populations
- concerned with the frequencies and types of illnesses and injuries in groups of people and with the factors that influence their distribution

(5) Friedman GD. Primer of Epidemiology. 3rd ed. New York: McGraw-Hill, Inc., 1987.

- study of disease occurrence in human populations with the primary unit of concern being groups of persons
- study the etiology and prevention of disease and the allocation of effort and resources in health care facilities and communities
- concerned with disease patterns in natural populations such as communities or nations

(6) Rothman KJ. Modern Epidemiology. Boston: Little, Brown and Company, 1986.

- the study of the occurrence of illness
- as a branch of science, epidemiology deals with the evaluation of scientific hypotheses

(7) Kelsey JL, Thompson WD, Evans AS. Methods in Observational Epidemiology. New York: Oxford University Press, Inc., 1986.

- the study of the occurrence and distribution of diseases and other health-related conditions in populations

APPENDIX 6. Definitions of Epidemiology (continued)

- objectives: to determine the magnitude and impact of diseases or other conditions in populations or in selected subgroups of populations so that this information can be used in setting priorities for investigation and for control, in deciding where preventive efforts should be focused, in evaluating the efficacy of therapeutic procedures, and in determining what type of treatment facilities are needed

(8) White KL. Healing the Schism. Epidemiology, Medicine, and the Public's Health. New York: Springer-Verlag, 1992.

- From the Institute of Medicine: the mother science of public health is epidemiology i.e. the systematic, objective study of the natural history of disease within populations and the factors that determine its spread.. epidemiology is the 'glue' that holds public health's many professions together, and they rest upon the scientific core of epidemiology

- Epidemiology is the study of all factors and their interrelationships which affect the occurrence and course of health and disease in a population. These factors include the characteristics of the host population; the causative agencies-predisposing, precipitating and perpetuating; and the biological, physical and social environment (as cited in White, p. 129).

- The objective is to discover the causes of the disease process and to determine the points in its natural history where interruptions may be accomplished in man's favor.

APPENDIX 7. Included Records

Order No: AAC 8413501 ProQuest - Dissertation Abstracts
Title: THE ASSOCIATIONS OF CIGARETTE SMOKING, HERPES SIMPLEX VIRUS TYPE 2 INFECTION, AND CANCER OF THE CERVIX
Author: MAYBERRY, ROBERT MORRIS
School: UNIVERSITY OF CALIFORNIA, BERKELEY (0028) Degree: PH.D.
Date: 1983 pp: 163
Source: DAI-B 45/03, p. 836, Sep 1984
Subject: HEALTH SCIENCES, PUBLIC HEALTH (0573)

Abstract: The objective of this epidemiologic study was to assess the relationship of cigarette smoking, herpes simplex virus type 2 (HSV-2) infection, and cervical cancer. The retrospective case-control study design compared women with abnormal cervical smears (n = 210) seen at the dysplasia clinic and women without abnormal smears (n = 317) attending gynecologic and birth control clinics at the University of California Medical Center, San Francisco. In this unmatched case-control study, factors associated with disease in univariate analyses of all cases of cervical abnormalities (including dysplasia, carcinoma in situ, and invasive carcinoma) were also more markedly observed in the analyses of histologically confirmed severe dysplasia, carcinoma in situ, and invasive carcinoma (n = 35). Cases were older, less well-educated, more likely to be black, and more likely to have had multiple marriages, pregnancies, and live births. Cases and controls did not differ by early age of first sexual intercourse, multiple sex partners, or antibodies to HSV-2. The prevalence of cigarette smoking was higher among cases than controls. The crude risk of cervical abnormality was 1.87 for current smokers and 1.27 for past smokers. The risk of severe disease was 2.92 for current smokers and 2.94 for past smokers. Adjusting for several potential confounding factors did not alter the significance of the smoking/cervical disease association. Women who smoked and were infected with HSV-2 were not at a greater risk of cervical abnormality than women who smoked and were not infected with HSV-2. However, cigarette smoking and HSV-2 were associated. The prevalence of smoking was higher among women with antibodies to HSV-2 than women without antibodies to HSV-2. Logistic analyses to assess the independent influence of various factors in the development of cervical abnormality indicated that race, education, number of live births, and cigarette smoking were independent risk factors for the predominantly dysplasia case group. The risk of cervical abnormality was 1.40 for past smokers and 1.97 for current smokers. Age, education, and cigarette smoking were the only independent risk factors for severe disease. Relative risks were 1.94 for past smokers and 3.74 for current smokers. No significant interaction between cigarette smoking and HSV-2 was observed in logistic analyses.

APPENDIX 7. Included Records (continued)

Order No: AAC 8712603 ProQuest - Dissertation Abstracts
Title: DIETARY RELATIONSHIPS TO FATAL BREAST CANCER AMONG SEVENTH-DAY ADVENTISTS
Author: MILLS, PAUL KEVIN
School: THE UNIV. OF TEXAS H.S.C. AT HOUSTON SCH. OF PUBLIC HEALTH(0219) Degree: PH.D. Date: 1986 pp: 132
Source: DAI-B 48/03, p. 721, Sep 1987
Subject: HEALTH SCIENCES, PUBLIC HEALTH (0573)

Abstract: Laboratory experiments in animals, correlational and migrant studies in humans suggest a role for diet in the etiology of breast cancer. Data gathered from individuals via case-control studies are less consistent. Seventh-day Adventist women experience lower mortality from breast cancer than comparable U.S. populations and this decrease is thought to be, at least in part, related to dietary practices (half are vegetarian). In 1960, 25,000 California Seventh-day Adventists completed a questionnaire which included a 21 item food frequency section. Cancer mortality in this population was monitored between 1960 and 1980 and the relationship of high fat food intake and fatal breast cancer was evaluated. Although established risk factors for breast cancer were observed in this population (e.g. age at menarche, age at first pregnancy, age at menopause and obesity) consumption of high fat foods were not observed to exert a strong influence on fatal breast cancer risk. Odds ratios (O.R.) for fatal breast cancer among non-vegetarians was 1.2. Increasing meat consumption bore little relation to risk; O.R. = 1.0, 1.2, 1.1 for consumption categories of none/occasional, 1-3 days/week and 4+ days/week respectively. Nor did the consumption of other high fat foods of animal origin (e.g. butter, cheese, milk, eggs) show any relationship to risk. These results remained unchanged after simultaneously controlling for the effect of other, potentially confounding variables (menstrual characteristics, obesity) via logistic regression analysis.

APPENDIX 7. Included Records (continued)

Order No: AAC 9136174 ProQuest - Dissertation Abstracts
Title: CHEMOTHERAPY AND RADIATION IN THE TREATMENT OF BREAST CANCER AND THE RISK OF A SECOND PRIMARY OF THE BREAST
Author: NICOLAOU, ANDREAS LEON
School: YALE UNIVERSITY (0265) Degree: PH.D. Date: 1991 pp: 230
Advisor: THOMPSON, W. DOUGLAS
Source: DAI-B 52/10, p. 5057, Apr 1992
Subject: BIOLOGY, BIOSTATISTICS (0308); HEALTH SCIENCES, PUBLIC HEALTH (0573); HEALTH SCIENCES, CHEMOTHERAPY (0992)

Abstract: Among women who have had cancer of the breast, the risk of developing a second primary breast cancer can be up to five times that of women in the general population. It has been suggested that adjuvant chemotherapy and radiation treatments for the initial cancer can affect the development of a second. The effect each of these two factors may have on contralateral breast cancer was studied in a case-control study. Known and suspected risk factors were treated as confounders in the analyses. Through the Connecticut Tumor Registry, 309 cases and 325 controls were identified between July 1, 1983 and September 30, 1986. Data were collected through interviews and abstracted from medical records. The use of chemotherapy was significantly protective against a second primary of the opposite breast (aOR = 0.52) This effect was most evident during the years closest to the treatment of the first breast cancer, and among women who were older ($\geq$50) at first diagnosis. Radiation to the breast also seemed to have a significant overall protective effect (aOR = 0.56). Here, too, the protection was strongest for those diagnosed with the first breast cancer at a later age. During the years most distant from initial diagnosis, radiation to the breast may actually increase a woman's risk for contralateral breast cancer. The protection seen during the early years after treatment is questionable because of lack of biological plausibility. The greatest weakness of this study is that of missing values for variables that are essential to the analysis. Since proper adjustment of potential confounders was not always feasible, interpretations should be made cautiously. Although the results suggest associations exist, further research must emphasize complete data collection.

APPENDIX 8. Excluded Records

Example 1a: Non-Specific Research Outcomes

Order No: NOT AVAILABLE FROM UMI ProQuest - Dissertation Abstracts
Title: ETUDE SUR LES CHANGEMENTS RESPIRATOIRES AIGUS CHEZ LES RAVAILLEURS DES SALLES DE CUVES DE L'INDUSTRIE PRIMAIRE DE L'ALUMINIUM (COMPLEXE INDUSTRIEL DE JONQUIERE)
Author: DURAND, PIERRE
School: MCGILL UNIVERSITY (CANADA) (0781) Degree: PH.D Date: 1985
Source: DAI-B 46/11, p. 3804, May 1986
Subject: HEALTH SCIENCES, PUBLIC HEALTH (0573)

Abstract: Acute respiratory changes have previously been observed in aluminum potroom workers. The main purposes of this study were to determine the prevalences of such changes amongst potroom workers assigned respectively to either one of the plant electrolytic processes, namely the Soderberg and the Prebake, and to compare the prevalence in each exposure group to that in unexposed workers of the same plant. Pulmonary function tests were administered four times on the first working shift of the week to 385 of the 412 subjects who had also answered a respiratory symptoms questionnaire. An index of acute changes was devised. Among the 279 smokers, those exposed to the Prebake process and with high exposure to the Soderberg had a higher index on average than those with low exposure to the Soderberg process and unexposed workers. Among the 106 workers who had never smoked, no significant differences appeared.

APPENDIX 8. Excluded Records (continued).

Example 1b: Clinical Studies

Order No: AAC 8426864 ProQuest - Dissertation Abstracts
Title: CYTOGENETIC STUDIES OF WORKERS EXPOSED TO ETHYLENE OXIDE
Author: MEINHARDT, THEODORE JOSEPH
School: UNIVERSITY OF CALIFORNIA, BERKELEY WITH SAN FRANCISCO STATE UNIV. (0515) Degree: PH.D. Date: 1984 pp: 284 Source:DAI-B 45/09, p. 2885, Mar 1985
Subject: HEALTH SCIENCES, PUBLIC HEALTH (0573)

Abstract: The effects of occupational exposure to ethylene oxide(EtO) and cigarette smoking on sister chromatid exchanges (SCEs) in circulating lymphocytes were assessed in an epidemiologic study. Workers in three occupational environments were compared: those unexposed to chemicals ("comparison subjects"); those in production units manufacturing EtO ("production workers"); and those in the quality assurance laboratory, where there were many chemical exposures with EtO one of the more common ("laboratory workers"). Two analytical approaches were used to measure these effects: (1) comparing the mean SCE frequency among groups with different exposures, and (2) comparing the distribution of SCE frequency per chromosome of each subject with an expected distribution based on the comparison subjects. Subjects with significantly elevated SCE levels based on this second criterion were considered to have "high SCE distributions". Cigarette smoking was associated with a 27.6 percent increase in group mean SCE frequency per chromosome in the comparison subjects. Among nonsmokers, employment in the production units was associated with a marginally significant elevation of 7.4 percent in group mean SCE frequency ($p = 0.06$). Among production workers who smoke cigarettes, the elevation of group mean SCE frequency was not statistically significant, although it was equal in absolute value (i.e. 7.2 percent) to that observed in nonsmokers. The same evaluation of laboratory workers relative to the appropriate comparison individuals revealed significant elevations in group mean SCE frequency of 42.3 percent in nonsmokers and 11.5 percent in cigarette smokers. The odds ratio for high SCE distributions of cigarette smokers versus nonsmokers was 42.2 among comparison individuals. Production workers who did not smoke had a 7.7 odds ratio for high SCE distributions relative to comparison nonsmokers. Cigarette smokers in the production units had an odds ratio for high SCE frequencies of 5.2 relative to comparison cigarette smokers. The following significant odds ratios were observed in laboratory workers relative to the appropriate comparison individuals: 69.0 for nonsmokers and 30.3 for cigarette smokers. A comparison of independent and joint effects provided some evidence of interaction between factors.

APPENDIX 8. Excluded Records (continued)

Example 2a: Methodological Studies

Order No: AAC NN69180 ProQuest - Dissertation Abstracts
Title: HE EFFECTS OF ERRORS IN MEASUREMENT IN SURVIVAL ANALYSIS (AIDS, IMMUNE DEFICIENCY)
Author: RABOUD, JANET MARIE
School: UNIVERSITY OF TORONTO (CANADA) (0779) Degree: PHD Date: 1991 pp: 146 Source:DAI-B 53/08, p. 3876, Feb 1993
Subject: BIOLOGY, BIOSTATISTICS (0308), HEALTH SCIENCES,PUBLIC HEALTH
(0573) ISBN:0-315-69180-8

Abstract: This work was motivated by the analysis of data from the Toronto Sexual Contact Study, a longitudinal study which followed male sexual contacts of men diagnosed with Acquired Immunodeficiency Syndrome (AIDS) or AIDS Related Complex (ARC). A series of immunologic markers was measured repeatedly over time as part of the study protocol and one of the problems of interest was to find out which of the markers best predicted progression to AIDS among HIV infected cohort members. The immunologic markers are known to fluctuate considerably over time due to minor infections and to laboratory measurement error, even among healthy individuals. In this thesis we investigate several aspects of the effect of measurement errors on estimates of relative risk, using the proportional hazards model. Errors in two types of covariates are considered: controlled and uncontrolled covariates. Controlled covariates are determined in advance, unlike uncontrolled covariates which are allowed to vary freely. Berkson (1950) writes, "A controlled observation is made when, instead of wanting to know the value of some unknown u$\sb{i}$, we wish to bring the quantity to a value u$\sb{i}$." In uncontrolled experiments one tries to measure the true covariate as accurately as possible. In contrast, in controlled experiments the experimenter attempts to bring the true value of the covariate to the value of the observed covariate, which is fixed in advance. The use of smoothing techniques to reduce error in covariates measured repeatedly over time is investigated. In addition to errors in the covariates, the time from infection with HIV to diagnosis of AIDS was measured imprecisely. This imprecision is due to the fact that the dates of infection of men who were infected with HIV before they were enrolled into the study had to be estimated and that the dates of diagnosis with AIDS were difficult to determine exactly due to the progressive nature of the disease. We consider the effects of errors in survival times on estimates of survival probabilities and on the logrank statistic. Methods developed throughout the thesis are illustrated with data from the Toronto Sexual Contact Study.

APPENDIX 8. Excluded Records (continued)

Example 2b: Methodological Studies

Order No: AAC 8416945 ProQuest - Dissertation Abstracts
Title: ASSESSMENT OF THE COMPARABILITY OF FREQUENCY AND QUANTITATIVE DIETARY INTAKE MEASUREMENTS FOR EPIDEMIOLOGIC STUDIES OF DIET-DISEASE ASSOCIATIONS
Author: CHU, SUSAN YING
School: UNIVERSITY OF HAWAII (0085) Degree: PHD Date: 1982 pp: 187
Source: DAI-B 45/04, p. 1161, Oct 1984
Subject: HEALTH SCIENCES, PUBLIC HEALTH (0573)

Abstract: The extent of agreement between frequency and quantitative dietary intake assessments was measured to determine whether frequency intake data can substitute for quantitative intake data without affecting the statistical conclusions in diet-disease associations. Frequency and quantitative intakes of 342 male subjects were obtained using a recall interview method designed to assess the subject's past, usual dietary habits. The extent of agreement between frequency and quantitative intakes of various dietary components (44 food items, 20 food groups, and 8 nutrients) was determined at the group and individual levels, and with regard to the specific study objective. This study found that for certain situations, e.g., ecologic analyses comparing mean intakes of various food items, frequency intake data can be used interchangeably with quantitative measurements and give fairly comparable results. However, for most purposes, e.g., for comparisons of mean nutrient intakes or when estimates of individual intakes are required, the extent of agreement was not sufficient to provide clear confidence that frequency dietary data can be used in place of quantitative data without changing the measured diet-disease associations.

APPENDIX 8. Excluded Records (continued)

Example 3: Rehabilitation-Related/Functional Health Status Studies

Order No: AAC 8610028 ProQuest - Dissertation Abstracts
Title: DETERMINANTS OF FUNCTIONAL HEALTH STATUS IN THE ELDERLY (DISABILITY, PHYSICAL FUNCTIONING, AGING)
Author: GURALNIK, JACK MICHAEL
School: UNIVERSITY OF CALIFORNIA, BERKELEY (0028) Degree PHD Date 1985 pp 376
Source: DAI-B 47/03, p. 1014, Sep 1986
Subject: HEALTH SCIENCES, PUBLIC HEALTH (0573); GERONTOLOGY (0351)

Abstract: This study describes physical functioning and disability and evaluates determinants of functional status in a representative sample of Alameda County, California residents followed from 1965 to 1984. The cohort investigated here was born prior to 1920, with survivors being aged 65 and over at the time of the 1984 follow-up survey. The 1984 survey was done by telephone and included a comprehensive set of questionnaire items to assess activities of daily living, mobility, physical performance of more strenuous activities and exercise, active recreation and sports The functional level of surviving members of the cohort was found to be quite similar to that of other large cohorts surveyed in comparable ways. Functional level remained high in a substantial percentage of those age 65 to 79. Compared to this group, in those age 80 and above a far greater percentage of individuals had difficulty or needed help in performing the activities being queried Contrary to common beliefs about the elderly, a large proportion of the sample engaged in exercise, active recreation and sports. A scale of physical functioning was developed and subsequently used to score participants in the 1984 survey. Two analytic approaches were used to evaluate variables predictive of this function score. In the first, multiple regression analysis was used to assess, among surviving participants, the association of variables collected in previous surveys with the 1984 function score. In the second, the entire 1965 cohort was dichotomized in terms of eventual outcome. Those surviving until 1984 with high function were compared with the remainder of their cohort using multiple logistic analytic techniques Adjustment for baseline status was done using three variables representing 1965 health status Precisely controlling for baseline physical function score was not possible, as the variables comprising this scale were not all collected in 1965. A variety of independent variables, most of which were collected in 1965, were assessed for their relationship with functional outcome in 1984 Five general domains hypothesized to relate to functional outcome included: sociodemographic status, health-related practices, psychological factors, energy level and active lifestyle and other health-related variables In one or more of the analyses, a significant and independent association with higher function was found for younger individuals, men, non-blacks, higher family income levels, being married in 1965 and 1974, non-smoking, higher physical activity level, weight status, positive feelings, high energy level, busy spare time, involvement in groups and normal blood pressure. It is suggested that measures of functional status offer a valuable means of assessing health in older persons and that this work offers insight into factors associated with healthy aging.

APPENDIX 8. Excluded Records (continued)

Example 4: Health Services Research

Order No: AAC 8314267 ProQuest - Dissertation Abstracts
Title: SELECTIVE ENROLLMENT IN AND DISENROLLMENT FROM HMOS BY MEDICAID RECIPIENTS
Author: DESHARNAIS, SUSAN IRENE LIEBERMAN
School: THE UNIVERSITY OF MICHIGAN (0127) Degree: PH.D. Date: 1983
pp: 236
Source: DAI-B 44/02, p. 463, Aug 1983
Subject: HEALTH SCIENCES, PUBLIC HEALTH (0573)

Abstract: Medicaid recipients could choose to obtain health care from either of two Health Maintenance Organizations (HMOs) or from fee-for-service providers. Lower hospital use among Medicaid HMO families led to speculation that a large proportion of lower-risk families had chosen HMOs. Some HMO physicians saw both HMO and non-HMO Medicaid families and had financial incentives to selectively influence enrollment and disenrollment decisions. The purpose of this study was to assess evidence of selective enrollment and disenrollment, and evidence of physician influence. Enrollment and utilization data were analyzed for cases with continuous Medicaid eligibility by using Contingency tables, Discriminant Analysis and Logit Analysis. HMO families had lower than average per capita expenses during the three months preceding enrollment. Over 69 percent of these HMO families had no physician contacts during this pre-enrollment period. Among those families who saw HMO physicians exclusively prior to joining the HMO, there was no clear evidence that the lower-risk families were influenced to join the HMOs or the higher-risk families were retained as private patients. Post-disenrollment hospital utilization and total expenses were higher than average during the three months after families left the HMOs. However, most disenrolled families either began to see non-HMO physicians or received no medical care during this time period. Thus, there was no evidence that HMO physicians had influenced higher-risk families to disenroll so that they then could continue to see them as private patients. It would seem that selective enrollment occurred not because of physician influence, but because the HMOs were more successful in enrolling families that did not have ongoing relationships with physicians. Because those without physicians were probably healthier than average, this resulted in favorable recruitment by the HMOs. Reasons for relatively high use following disenrollment are more difficult to interpret, and need to be explored using more direct methods such as interviews.

APPENDIX 8. **Excluded Records (continued)**

Example 5a: Health Education/Promotion Studies

Order No:AAC 8821534 ProQuest - Dissertation Abstracts
Title: THE INFLUENCE OF ATTITUDE, SOCIAL INFLUENCES AND KNOWLEDGE N PHYSICIANS' DECISIONS TO OFFER MATERNAL SERUM ALPHA FETOPROTEIN SCREENING FOR NEURAL TUBE DEFECTS
Author: SABLE, MARJORIE ROSE
School: THE UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL (0153)
Degree: DR.P.H. Date: 1987 pp: 135
Advisor: PEOPLES-SHEPS, MARY
Source:DAI-B 49/09, p. 3707, Mar 1989
Subject:HEALTH SCIENCES, PUBLIC HEALTH (0573); EDUCATION, HEALTH (0680)

Abstract: The purposes of this study were (1) to examine the relationships among attitude, perceived social influences, knowledge and physician intention with regard to maternal serum alpha-fetoprotein (MS-AFP) screening for neural tube defects (NTDs), and (2) to identify which beliefs and social referents differentiated those physicians who intend to screen from those who do not. The study hypotheses were based on Fishbein and Ajzen's Theory of Reasoned Action (1975).A cross-sectional survey was conducted to examine the factors related to physicians' decisions to offer MS-AFP screening. Data from 186 physicians who provide prenatal care in Missouri reveal that physicians intending to offer MS-AFP screening had more favorable attitudes towards screening and perceived that their social referents felt that they should offer screening. High intenders were more likely than low intenders to perceive that screening is dictated by current medical-legal standards of care, that their partners and colleagues think they should offer screening and that the American College of Obstetricians and Gynecologists (ACOG) recommend screening. Physician attitude toward abortion and location of practice were important predictors of screening intention. High intenders held a more liberal attitude toward abortion and practiced in urban areas. Although there was a wide range of scores on the knowledge quiz, knowledge did not predict screening intention over and above the other variables in the model. While certain beliefs about the consequences of screening differentiated between high and low intenders, neither group believed that MS-AFP screening would reduce the incidence of neural tube defects, although 75% of the physicians felt that abortion was permissible in circumstances of a serious birth defect. The results of this study suggest the MS-AFP screening intention is based more on a desire to comply with medical-legal standards of care than on belief that screening will lead to a reduction in the incidence of NTDs. These findings may reflect the growing influence of medical liability on medical practice

APPENDIX 8. Excluded Records (continued)

Example 5b: Health Education/Promotion Studies

Order No: AAC 9224428 ProQuest - Dissertation Abstracts
Title: THE SOCIAL CONSTRUCTION OF RISK: AIDS KNOWLEDGE AND RISK PERCEPTION AMONG THE UNITED STATES POPULATION (IMMUNE DEFICIENCY, MASS MEDIA)
Author: SWEAT, MICHAEL DAVID
School: EMORY UNIVERSITY (0665) Degree: PHD Date: 1992 pp: 214
Advisor: LEVIN, MARTIN L.
Source: DAI-A 53/04, p. 1281, Oct 1992
Subject: SOCIOLOGY, GENERAL (0626); HEALTH SCIENCES, PUBLIC HEALTH (0573); EDUCATION, HEALTH (0680)

Abstract: This is a study of how people perceive risk. Economic, cultural, psychological, and public health theories of risk perception are reviewed and a model tracing the relationship between risk perception and the standard public health intervention approach is developed. This model is tested using data from the National Health Interview Survey AIDS Knowledge and Attitudes Supplement of 1988 and 1989, one of the few surveys to collect information about AIDS from a representative sample (N = 71,379) of the U.S. population. Analysis revealed that mass media was the most common source of information about AIDS, with exposure to AIDS information from mass media reported most often by those who were young and more highly educated. Knowledge about AIDS clustered into four distinct factors (HIV transmission, AIDS definition, technical issues, and commonly known facts about AIDS). The best predictor of knowledge about AIDS was exposure to AIDS mass media campaigns. Perception of risk of acquiring AIDS was rare among the U.S. population, yet many people misinterpreted their potential for getting AIDS. Among those with no reported risk behavior, overestimation of risk was associated with lower education, male gender, single marital status, Non-White race, and knowing a person with AIDS. For those at risk for AIDS, underestimation of risk was associated with higher education, female gender, and being married. A path analysis showed that perception of risk for acquiring AIDS is best predicted by demographic factors and social networking factors, rather than knowledge about AIDS. Mass media campaigns to educate the public about AIDS have increased knowledge, but have failed to increase accurate risk perception. Many people in the United States continue to misinterpret their risk of acquiring AIDS due to biases about who is likely to get infected with the AIDS virus.

APPENDIX 9. Coding Rules and Data Dictionary

I. Bibliographic Citation

(a) Variable: IDNO Variable Label: ID Number

- The seven-digit order number included in the bibliographic citation was used as the identification number for each dissertation abstract.
- Order numbers for Canadian abstracts were preceded by the letter N - (example: NN59110).
- When the order number was unavailable, a unique seven-digit identification number was assigned (e.g. Canadian records - NN00001, US records - 9990001).

(b) Variable: GENDER Variable Label: Gender of Study Subject

1=male 2=female
8=indeterminate

- The presumed gender of the author was determined using the first name of the doctoral recipient.

(c) Variable: SCHOOL Variable Label: School

- The UMI-assigned four-digit school code was entered for each Canadian and American institution.

(d) Variable: Degree Variable Label: Type of Degree Awarded

1=Ph.D. 2=Sc.D.
3=Dr.P.H. 8=unknown

- A categorical variable was created to capture information regarding the type of doctoral degree awarded.

(e) Variable: Year Variable Label: Year Degree Awarded

- The last two digits of the year in which the doctoral degree was awarded were captured (e.g. 82-92).

(f) Variable: NOPAGES Variable Label: Number of Pages

- The total number of dissertation pages was recorded. When this information was unavailable, NOPAGES was recorded as missing.

APPENDIX 9. Coding Rules and Data Dictionary (continued)

(g) Variable: SUBAREA1-SUBAREA3
Variable Label: Dissertation Author-Assigned Subject Area(s)

- The four-digit author-assigned subject areas were abstracted as reported in the bibliographic citation (e.g. 0573=Health Sciences, Public Health; 0347=Health Sciences, Mental Health).

II. Abstract

(a) Variable: IDNO Variable Label: ID Number

- The seven-digit order number included in the bibliographic citation was used as the identification number for each dissertation abstract.
- Order numbers for Canadian abstracts were preceded by the letter N - (example: NN59110).
- When the order number was unavailable, a unique seven-digit identification number was entered (e.g. Canadian records - NN00001, US records - 9990001).

(b) Variable: DESIGN1 - DESIGN5 Variable Label: Study Design

1=Intervention Study	2=Cohort Study
3=Case-Control Study	4=Cross-Sectional Study
5=Ecological Study	6=Hybrid Study
7=Other	88=cannot classify

- When the study design was stated explicitly by the dissertation author, the appropriate classification was coded; otherwise, the design was classified based on information described in the abstract.
- Intervention and cohort studies were classified more precisely when sufficient information was provided (e.g. prospective or retrospective cohort study). Unless "randomized", "random allocation", or "random assignment" was specified, the design was classified as "Other Intervention Study". When a cohort study could not be classified as prospective or retrospective, the design was classified simply as a cohort study.
- When the dissertation author carried out a secondary data analysis, and the original study design was described, the design was classified accordingly. If the previously collected data were gathered using a survey approach, then the design was coded as a cross-sectional study; or, if the data were analyzed using a different study design approach (e.g. nested case-control analyses), then the latter design was classified.

APPENDIX 9. Coding Rules and Data Dictionary (continued)

- When there were multiple designs, the predominant study design was classified and all corresponding study variables were coded. Remaining designs were captured in DESIGN2 to DESIGN5.
- The nested case-control design was coded as "Hybrid Design".
- "Other" designs included proportional mortality studies, family studies, heterosexual partner studies, and incidence studies. Separate categories were not created since the total number of these studies was small.

(c) Variable: OUTCOME Variable Label: Research Topic Area

- Research outcomes were classified according to the conventions outlined in the ICD-9-CM.
- When the dissertation author studied a single disease, the appropriate code was recorded.
- When the author investigated more than one disease, each outcome was classified accordingly, and was described by a single "summary" code (e.g. lung cancer, breast cancer, and colon cancer wereeach recorded, yet the summary code is "cancer"). However, if the study diseases extended beyond one ICD chapter (e.g. mortality from coronary heart disease, cancer and suicide), the summary code of "multiple diseases" was assigned and the dissertation was included in a separate and newly defined chapter heading of the same name.
- Dissertations investigating "risk of death", and "overall mortality" were classified simply as "mortality" and were included in a separate and newly defined chapter heading of the same name.
- Unless specified as "benign", all cancer outcomes were coded according to "Primary, Malignant" at the three-digit rubric level.
- When the outcome was a manifestation of the underlying disease, the latter was coded.
- Unspecified complications of pregnancy, labor and delivery were classified under one heading and included in chapter 11 of the ICD-9-CM.
- A separate category was created for adverse outcomes of pregnancy and was included in chapter 15 of the ICD-9-CM.
- When there were multiple designs, the corresponding outcomes were classified accordingly.

(d) Variables: AGEGRP1-AGEGRP5 Variable Label: Age of Study Subjects

1=0-12 months	2=13 months-17 years
3=18-44 yrs	4=45-65 yrs
5=65 + years	8=unknown

- The age-group was recorded for the study subjects in which the outcome was being measured (e.g. risk of birth defects AGEGRP=1.)

APPENDIX 9. Coding Rules and Data Dictionary (continued)

- When the age-group of study subjects was reported, the corresponding code was entered directly in the numeric variable AGEGRP.
- When the age-group was described in the text as "children", the corresponding code was entered (e.g. AGEGRP=2).
- When the age-group was not provided AGEGRP was classified as unknown.

- Other AGEGRP procedures:
- a university-based sample was coded as 18-44 years
- the age of veterans was coded as 8 (unknown) unless the specific war was identified
- when the dissertation author included current employees as study subjects, age was coded as 18-44 and 45-65 years unless otherwise specified
- the age of subjects in an occupational cohort study was coded as 8 (unknown) when current vs. retired employee status could not be distinguished.
- when the age of study subjects was reported as 5-19 years, the predominant age-group was coded (e.g. 13 months-17 years)
- when age was reported as less than 21 years, the predominant age-group (e.g. 13 months-17 years) was recorded.
- the age of study subjects was coded as 8 (unknown) when described broadly (e.g. 18 years and over, less than 50 years etc.).
- pregnant women were classified as 18-44 years
- when the outcome was classified as spontaneous abortion or miscarriage, age was coded as 18-44 years
- both the age of the infant and mother was recorded when the study disease was adverse pregnancy outcomes (e.g. 0-12 months, 18-44 years)

Note: Age-groups were initially defined according to the conventions outlined in the National Library of Medicine's Index Medicus. However, the age of study subjects was not described in sufficient detail and was revised to reflect the amount and quality of information presented in the dissertation abstract.

(e) Variables: ETHNIC1-ETHNIC5 and ETHCOMM1-ETHCOMM5
Variable Label: Ethnicity of Study Subjects

1=White	2=Black
3=Hispanic	4=Asian
5=Native Americans	8=unknown

- The ethnic distribution of study subjects was classified according to the categories outlined by the United States National Center for Health Statistics (1993).
- If a distinct ethnic group was reported, then the corresponding code was entered.

APPENDIX 9. Coding Rules and Data Dictionary (continued)

- If an ethnic group other than White, Black, Hispanic, Asian, or Native American was described, then the code for "Other" was entered in the numeric variable ETHNIC and the text was captured directly in the string variable ETHCOMM. The string variables were subsequently recoded to reflect the predominant ethnic group. Also, geography was used to assign the predominant group in abstracts not reporting this information (e.g. Puerto Rico - coded ethnic group as Hispanic).
- When ethnicity could not be determined, ETHNIC=8 (unknown).
- When the ethnicity of study subjects was not described but instead mentioned in the results section for one particular ethnic group, then ETHNIC1 was coded accordingly, and ETHNIC2 was coded as unknown.

(f) Variable: STUDYSEX Variable Label: Gender of Study Subjects

1=male	2=female
3=both	8=unknown

- When the gender of study subjects was provided, the appropriate code was entered.
- When the gender was not provided but could be inferred (with confidence) from the study disease (e.g. breast cancer) the corresponding code was recorded.
- When the gender was not provided and the study disease did not provide sufficient information to make a judgment, STUDYSEX=3 (males and females).
- When the author selected subjects from a traditionally male-dominated occupation (e.g. petroleum refinery) yet did not specifically state the subjects' gender, STUDYSEX was coded as unknown since gender could not be coded as "males" or "males and females" with confidence.
- When the outcome was spontaneous abortion or miscarriage, STUDYSEX was recorded as females.
- When the outcome was adverse pregnancy outcomes, gender was classified for both infants and mothers (e.g. STUDYSEX=3).

(g) Variable: NOSUBJ Variable Label: Number of Study Subjects

- The total number of subjects participating in the study was recorded.
- When the author restricted the analysis to a reduced sample size then that number was recorded.
- When the number of study subjects varied across data collection components, the largest number was recorded (e.g. questionnaire data was completed for 3000 study subjects, but serological data was complete for only 2500 subjects - number of subjects is 3000).

APPENDIX 9. Coding Rules and Data Dictionary (continued)

- When the number of study subjects was not provided, or, could not be deciphered from the information provided, NOSUBJ was coded as unknown. Similarly, NOSUBJ was coded as unknown when the overall sample size was not reported in dissertations merging information from two or more existing datasets.
- When the study design was a case-control study, and the number of cases, or, number of controls was incomplete, NOSUBJ was coded as unknown.

(h) Variable: NOCASES and NOCTRLS
Variable Labels: Number of Cases/Number of Controls

- When the dissertation abstract described a case-control or nested case-control study, the corresponding number of cases and controls was entered when reported. When this information was incomplete, NOCASES and NOCTRLS was coded unknown.

Note: Information on group size was inconsistently reported for intervention (i.e. treatment and control groups) and cohort studies (i.e. exposed vs. unexposed group). Therefore, group size was limited to case-control and nested case-control studies.

(i) Variable: GEORSH1 - GEORSH8
Variable Label: Place Where Research was Carried Out

- When information was provided which described the setting in which the research was carried out, the state/province or country was recorded.
- When the dissertation author used existing datasets to answer his/her research question and the location of the research was reported, GEORSH was classified accordingly (e.g. Framingham Study - GEORSH=Massachusetts).
- When information was not provided for this variable GEORSH=1 (i.e. province/state of affiliated institution).
- For settings outside of Canada and the United States, the countries were coded to reflect the region of the world most appropriate (e.g. Zaire = Africa).

APPENDIX 10. Data Abstraction

Order No: NOT AVAILABLE FROM UMI ProQuest Dissertation Abstracts
Title: THE COMMUNITY HYPERTENSION MANAGEMENT PROJECT
Author: MCALISTER, NEIL HARDING
School: UNIVERSITY OF TORONTO (CANADA) (0779) Degree: PHD Date: 1986
Source: DAI-B 47/06, p. 2367, Dec 1986
Subject: HEALTH SCIENCES, MEDICINE AND SURGERY (0564)

Abstract: A clinical trial tested the hypothesis that general practitioners using computer-aided follow-up and treatment of hypertension could obtain better patient outcomes than doctors who did not use computer support. Ninety family physicians were randomized to three treatment strategies. Test physicians (T) completed a data collection sheet after each visit of a hypertensive patient, and mailed these forms to the test center for processing. Computer feedback was mailed to the doctors. It encouraged application of the "stepped care" protocol, furnished charts of diastolic pressure versus time for every patient, and percentile ranked patients' diastolic blood pressures (DBP). Appointment reminder letters were mailed to patients. Attention Control doctors (AC) filled out the same data collection forms as Test physicians, but neither doctors nor patients received computer feedback. Physicians using the computer handled more patients per practice than Attention Control doctors. (50 in T; 40 in AC). For all subjects the length of followup was significantly longer in T than AC. (199 days T; 167 AC). Additional benefits of computer-assisted hypertension management were demonstrated for "moderate" cases with baseline DBP >= 105 mm Hg, and for cases newly-detected during the study. For "moderate" cases, the mean score last recorded DBP was below the goal of 90 mm Hg in Test practices (88.5), but it failed to reach goal in Attention Controls (93.3). Greater average DBP reduction was achieved in T than AC (21.7 mm Hg T; 16.7 AC). While "moderate" cases were better-controlled in T than AC practices, these patients visited their doctors less frequently in T. (13.3 visits per patient-year in T; 17.4 AC). Among newly-detected cases, test practices achieved greater DBP reduction than controls (15.1 mm Hg in T vs. 11.3 AC), and more sustained hypertension control (323 days per patient-year with DBP <= 90 mm Hg in T, vs. 259 AC). In terms of clinical effect per unit cost, the computer-assisted strategy is more "cost-effective" than the control maneuver. This margin increases with the scale of implementation.

Data Abstraction

(1) agegrp1=8=unknown
(2) ethnic1=8=unknown
(3) studysex=3=both sexes
(4) nosubj=90
(5) design1=RCT
(6) outcome=hypertension
(7) georsh=1

APPENDIX 10. Data Abstraction (continued)

Order No: .AAC 8819066 ProQuest Dissertation Abstracts
Title: A LONGITUDINAL STUDY OF A COHORT WITH PAST EXPOSURE TO RADAR. THE MIT RADIATION LABORATORY FOLLOW-UP STUDY (EPIDEMIOLOGY)
Author: HILL, DOREEN G.
School: THE JOHNS HOPKINS UNIVERSITY (0098) Degree: PHD Date: 1988 pp: 387
Source: DAI-B 49/07, p. 2592, Jan 1989
Subject: HEALTH SCIENCES, PUBLIC HEALTH (0573); ENVIRONMENTAL SCIENCES (0768); BIOLOGY, BIOSTATISTICS (0308)

Abstract: The purpose of the study was to assess the long-term health effects among participants in a radar research and development project at the Massachusetts Institute of Technology (MIT) during .World War II. The project was conducted in what was known as the Radiation Laboratory (Rad Lab). The study was of a nonconcurrent prospective design. The overall mortality experience, the mortality from specific causes of death, and longevity were examined. Radar refers to the use of electromagnetic energy for detecting and locating reflecting objects. Radars generally are within the microwave portion (0.3 to 3.5 GHz) of the electromagnetic spectrum. Exposure to the radiation from radar systems is, thus, exposure to nonionizing radiation. The study population included all professional staff members of the MIT Rad Lab. The group largely consisted of scientists and engineers. Analyses were restricted to white males. The mortality experience of the cohort was ascertained after extensive tracing and follow-up. Of the total population of 1592 persons, less than five percent were not traced beyond the 1940's. The majority were followed into the 1980's with an average survival of 36 years. Standardized mortality ratios (SMRs) were calculated and demonstrated a lower than expected mortality based on rates for U.S. white males. A very strong healthy worker effect was evidenced. Certain underlying causes of death were elevated, e.g., Hodgkin's Disease, but none were statistically significant. The population was also compared to a cohort of physicians. The overall survival of the Rad Lab cohort was better than survival among physicians. With the exception of a slightly increased risk of Hodgkin's Disease and cirrhosis of the liver in the Rad Lab population, the cohorts were generally similar with respect to specific causes of death.

Data Abstraction

(1) agegrp1=3=18-44
(2) agecomm1=employee
(3) agegrp2=4=45-65
(4) agecomm2=employee
(5) ethnic1=1=white
(6) studysex=1=male
(7) nosubj=1592
(8) design1=retrospective cohort study
(9) outcome=multiple diseases - mortality
(10) georsh=Mass

APPENDIX 10. Data Abstraction (continued)

Order No: AAC 9000304 ProQuest Dissertation Abstracts
Title: A CASE-CONTROL STUDY OF ORAL CANCER AND PREDIAGNOSTIC CONCENTRATIONS OF SELENIUM, IRON, ZINC, AND CALCIUM IN NAIL TISSUE (CANCER)
Author: ROGERS, MARY A. M.
School: U NIVERSITY OF WASHINGTON (0250) Degree: PHD Date: 1989 pp: 144
Advisor:THOMAS, DAVID B.
Source: DAI-B 50/08, p. 3428, Feb 1990
Subject: HEALTH SCIENCES, PUBLIC HEALTH (0573); HEALTH SCIENCES, NUTRITION (0570); HEALTH SCIENCES, MEDICINE AND SURGERY (0564)

Abstract: A case-control study of the relationships between nail tissue levels of selenium, iron, zinc and calcium and cancer of the oral cavity was conducted. Oral cancer cases were identified through a population-based registry. Only individuals who were diagnosed from 9/83 through 2/87, were between the ages of 20 and 74, and who resided in King, Pierce and Snohomish counties of Washington State were included. Controls were selected by telephone using random digit dialing and were frequency matched by sex and age-group to the cases. Subjects participated in a personal interview and submitted clippings from the nails of each great toe. Concentrations of selenium, iron, zinc and calcium were measured in the nail samples by neutron activation analysis. Male oral cancer cases had lower selenium levels than did the controls, but female cases did not. The association between low selenium concentrations and oral cancer was particularly strong in young men. For those individuals who had a previous history of low ascorbic acid intake, low selenium levels were positively associated with oral cancer in both sexes. Cases were more likely to have low zinc levels than controls, while there was suggestive evidence that controls had lower calcium levels than cases. There was no association between low or high levels of iron and oral cancer. However, individuals with high iron concentrations accompanied by low zinc levels were 4 times more likely to have oral cancer than subjects with low iron and high zinc levels. There was a positive association between tonsillar carcinoma and low concentrations of selenium and iron. Stratification by time interval from diagnosis to interview suggested that the differences in mineral content of the nail tissue in cases and controls were a result of the disease, rather than related to etiology.

Data Abstraction

(1) agegrp1=3=18-44	(6) agecomm3=20-74	(11) noctrls=8=unknown
(2) agecomm1=20-74	(7) ethnic1=8=unknown	(12) design1=case-control
(3) agegrp2=4=45-65	(8) studysex=3=both sexes	(13) outcome=oral cancer
(4) agecomm2=20-74	(9) nosubj=8=unknown	(14) georsh=Wash. State
(5) agegrp3=5=65+	(10) nocases=8=unknown	

APPENDIX 10. Data Abstraction (continued)

Order No. AAC 8628531 ProQuest Dissertation Abstracts
Title: AN EPIDEMIOLOGIC INVESTIGATION OF CENTRAL NERVOUS SYSTEM CANCERS AMONG WORKERS AT DIVERSE NUCLEAR FACILITIES (BRAIN CANCER, RADIATION)
Author: CARPENTER, ARVIND VITHALBHAI
School: UNIVERSITY OF ALABAMA AT BIRMINGHAM (0005) Degree: DRPH Date: 1986 pp: 198
Source: DAI-B 47/08, p. 3316, Feb 1987
Subject: HEALTH SCIENCES, PUBLIC HEALTH (0573)

Abstract: Results of at least 15 epidemiologic investigations have shown that persons employed in specific industries, including the nuclear industry, may be at increased risk of developing central nervous system (CNS) cancers. A nested case-control study was conducted in three cohorts of workers employed between 1943 and 1977 in two nuclear facilities to investigate the possible association of primary CNS cancers with occupational exposure to radiation from external and internal sources, to chemicals, and/or to other nonoccupational risk factors such as history of epilepsy and head injury, and blood type. The study was motivated, in part, by a report of excess CNS cancer deaths in one of the three cohorts included in the study. The vital status for these cohorts was determined through 1979. Eighty-nine white male and female workers who, according to the information on death certificates, died of primary CNS cancers were identified as cases. For each case, four controls, who were alive at the time of death of the case, were individually matched on sex, cohort, year of birth, and year of hire. External radiation exposure data were available from film badge readings for individual workers, whereas radiation dose to lung from internally deposited radionuclides, mainly uranium, was estimated from area and personnel monitoring data and was used in analyses as a surrogate for the dose to brain assuming a proportionality between the lung dose and the brain dose. Twenty-six chemicals were identified and subjectively assessed for potential exposure to each job title in all departments. Plant medical records provided information about the nonoccupational factors for some of the cases and controls. Although we failed to find an association between CNS cancers and exposure to radiation from external or internal sources, such an association could not be entirely ruled out because the number of monitored subjects was small and their doses were low. A strong association between exposure to any of the 26 chemicals and CNS cancers was not found, but small effects may have been missed because of possible misclassification in exposure assessment. We found a moderately strong association between CNS cancer occurrence and history of epilepsy, but failed to find an association with major blood groups or history of head injury.

Data Abstraction

(1) agegrp1=8=unknown	(5) nocases=89	(9)georsh=1
(2) ethnic1=1=white	(6) noctrls=366	
(3) studysex=3=both sexes	(7) design1=nested case-control	
(4) nosubj=455	(8) outcome=CNS Cancer	

APPENDIX 10. Data Abstraction (continued)

Order No: AAC 9029188 ProQuest Dissertation Abstracts
Title: AN EPIDEMIOLOGIC STUDY OF SMOKING AND INFERTILITY IN WOMEN
Author: LAURENT, SHERRY LYNN
School: UNIVERSITY OF SOUTH CAROLINA (0202) Degree: PHD Date: 1990 pp: 118
Source: DAI-B 51/05, p. 2308, Nov 1990
Subject: HEALTH SCIENCES, PUBLIC HEALTH (0573), WOMEN'S STUDIES (0453)

Abstract: The purpose of this study is to determine if there is an association between exposure to cigarette smoke as measured by smoke intensity and cumulative smoke exposure and the outcomes of primary infertility, secondary infertility, and time to conception. The 3,402 subjects of this study are selected from the control group used for the Cancer and Steroid Hormone (CASH) Study which was coordinated by the Reproductive Health Division of the Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control. All information for this study was obtained through individual interview with life calendars from menarche to menopause constructed for each woman. The major findings of this study indicate that both smoking intensity (one pack of cigarettes per day, O.R. = 1.37) and cumulative smoke exposure (smoking initiated prior to 18 years of age, O.R. = 1.34) significantly contribute to the odds of primary infertility. Secondary infertility demonstrates similar findings with an odds ratio of 1.26 for smoking one pack of cigarettes per day and 1.15 for 5 pack years of smoke exposure. These significant associations remain unchanged after adjustment for age at infertility/first pregnancy, age at first intercourse, education, history of pelvic inflammatory disease, Quetelet Index, and history of ectopic pregnancy. Women most at risk in both categories of infertility are black, less educated, low income, and older aged. Life-table and Proportional Hazards analysis indicate that smoking intensity and cumulative smoke exposure are not associated with the length of time required to conceive. The age of women when infertility occurs is significantly associated with the time required for conception in both primary and secondary infertility categories (fecundability ratio = 0.93). In addition to age, within the secondary infertility category the age of intercourse, a history of pelvic inflammatory disease, and race are significantly associated with conception time.

Data Abstraction

(1) agegrp1=3=18-44
(2) ethnic1=2=black
(3) ethnic2=8=unknown
(4) studysex=2=female
(5) nosubj=3402
(6) design1=cross-sectional
(7) outcome=infertility
(8) georsh=1

APPENDIX 10. Data Abstraction (continued)

Order No: AAC 8527297 ProQuest Dissertation Abstracts
Title: APPARENT CONSUMPTION OF ALCOHOL AND VIOLENT AND ACCIDENTAL DEATH IN NORTH CAROLINA: A TIME SERIES ANALYSIS, 1956-1981 (SUICIDE, MOTOR VEHICLE, HOMICIDE)
Author: LESSNE, CYNTHIA WILLEY
School: THE UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL (0153) Degree: DRPH Date: 1985 pp: 387
Source: DAI-B 46/11, p. 3806, May 1986
Subject: HEALTH SCIENCES, PUBLIC HEALTH (0573)

Abstract: Rapid rises in per capita apparent consumption of alcohol occurred in North Carolina during the decades of the 1960's and 1970's. The purpose of this research is to examine the consequences of this rapid rise, particularly with respect to its impact upon mortality from violence and accidents. Monthly per capita apparent consumption of wine, beer, and spirits during 1956-1981 was correlated with mortality from non-traffic accidents, motor vehicle accidents, homicide, suicide, and cirrhosis of the liver among six different age-race-sex groupings. Seasonal trends and autocorrelation in the time series were modeled and controlled for using Box-Jenkins time series models. Results of the investigation suggest different effects for the three major alcoholic beverages. Beer and wine consumption showed strong correlations with non-traffic accident mortality among several specific age-race-sex groupings. Apparent consumption of wine and beer will also highly associated with male motor vehicle mortality. Spirits consumption was not associated with non-traffic accident mortality, but showed strong and consistent associations with motor vehicle accident mortality among all age-race-sex groups examined. Apparent consumption of all three beverages was associated with cirrhosis mortality at both early lags (0-1 month) and later lags (approximately one year). None of the three beverages showed associations with mortality from homicide or suicide. Differences in the results for each of the three alcoholic beverages are discussed with respect to differences in beverage specific patterns of alcohol consumption. The usual time and place of consumption of each beverage as well as differences in the popularity of each beverage among the various age-race-sex groupings examined are believed to account for some of the beverage specific differences observed.

Data Abstraction

(1) agegrp1=8=unknown	(7) outcome1=non-traffic accidents
(2) ethnic1=8=unknown	(8) outcome2=MVTA
(3) studysex=3=both sexes	(9) outcome3=homicide
(4) nosubj=8=unknown	(10) outcome4=suicide
(5) design1=ecological	(11) outcome5=cirrhosis of the liver
(6) outcome=multiple accidents	(12) georsh=North Carolina

APPENDIX 10. Data Abstraction (continued)

Order No: AAC 9200822 ProQuest Dissertation Abstracts
Title:IMPLICATIONS OF LOW BIRTH WEIGHT AND OTHER RISK FACTORS FOR THE PREVENTION OF MILD MENTAL RETARDATION
Author: MCDERMOTT, SUZANNE WALLMAN
School: UNIVERSITY OF SOUTH CAROLINA (0202) Degree: PHD Date: 1991 pp: 232
Source: DAI-B 52/07, p. 3558, Jan 1992
Subject: HEALTH SCIENCES, PUBLIC HEALTH (0573)

Abstract: This dissertation analyzes risk factors for mild mental retardation and then presents a dynamic model which can simulate the impact of changes in population parameters on the proportion of mentally retarded individuals in the total population. The Child Health and Development Studies data, from the Kaiser Health Plan in California, are used to analyze the relationship of low birth weight and other risk factors on mild mental retardation. The increase in risk for mild mental retardation is approximately two fold for low birth weight children at school entry compared to children born at normal birth weight. However, by age 9-11 years, this increased risk is no longer evident. College education of the parents had a strong protective effect for mild mental retardation (Odds Ratio = 0.3 (0.2, 0.5) at age 5 years, O.R. = 0.4 (0.2, 0.5) at age 9-11 years). The risk factors that were associated with increased risk of mild mental retardation at age 5 years were low birth weight, mothers education of less than a high school, black race, and low family income. None of the pregnancy, labor and delivery, or early childhood medical factors were significant risks. At age 9-11 parents education less than high school and black race were associated with mild mental retardation. A computer generated model was developed to illustrate and predict the impact of changing population characteristics on the proportion of mentally retarded people in the population. The model was implemented using California vital statistics for 1960, 1970 and 1980 and the Child Health and Development Study data. The simulations indicate that the most important prevention strategy for mild mental retardation is targeted programs for a subset of the poor where a deprived environment causes normally born children to score in the mild mental retardation range when they are school-age. The relevance of this tactic is the large number of children affected. Poverty affects 12-20 percent of children. Thus, the influence of deprivation on a portion of children living in poverty has a large impact on the proportion of individuals with mental retardation in the population.

Data Abstraction

(1) agegrp1=2=children
(2) ethnic1=2=black
(3) ethnic2=8=unknown
(4) studysex=3=both sexes
(5) nosubj=8=unknown
(6) design1=cannot classify
(7) outcome=mental retardation
(8) georsh=California

APPENDIX 11. ICD-9-CM Chapter Headings and Codes

ICD-9 Chapter	Chapter Title	Code
1	Infectious and Parasitic Diseases	001-139
2	Neoplasms	140-239
3	Endocrine, Nutritional, and Metabolic Diseases and Immunity Disorders	240-279
4	Diseases of the Blood and Blood-Forming Organs	280-289
5	Mental Disorders	290-319
6	Diseases of the Nervous System and Sense Organs	320-389
7	Diseases of the Circulatory System	390-459
8	Diseases of the Respiratory System	460-519
9	Diseases of the Digestive System	520-579
10	Diseases of the Genitourinary System	580-629
11	Complications of Pregnancy, Childbirth and the Puerperium	630-676
12	Diseases of the Skin and Subcutaneous Tissue	680-709
13	Diseases of the Musculoskeletal System and Connective Tissue	710-739
14	Congenital Anomalies	740-759
15	Certain Conditions Originating in the Perinatal Period	760-779
16	Symptoms, Signs, and Ill-Defined Conditions	780-799
17	Injuries and Poisoning	800-999
E-CODE	Classification of External Causes of Injury and Poisoning	E 800-998

Source: Commission on Professional and Hospital Activities. International Classification of Diseases. 9th Revision. Clinical Modification. Volume 1 & 2. Ann Arbor, Michigan: Edwards Brothers, Inc., 1978.

APPENDIX 12. Structured Abstract

Order No: AAC 9211585 ProQuest - Dissertation Abstracts
Title: A CASE-CONTROL STUDY OF CHLAMYDIA PNEUMONIAE STRAIN TWAR INFECTION AND ANGIOGRAPHICALLY DEMONSTRATED CORONARY ARTERY DISEASE
Author: THOM, DAVID HINTON
School: UNIVERSITY OF WASHINGTON (0250) Degree: PHD Date: 1991 pp: 160
Advisor: SISCOVICK, DAVID S.
Source: DAI-B 53/01, p. 224, Jul 1992
Subject: HEALTH SCIENCES, PUBLIC HEALTH (0573); HEALTH SCIENCES, MEDICINE AND SURGERY (0564)

Abstract: Objective. To evaluate the association between past infection with Chlamydia pneumoniae, as measured by antibody, and coronary artery disease (CAD).Study design. A case-control study was conducted from December 15, 1988 to April 30, 1991 among members of Group Health Cooperative of Puget Sound, a Seattle-based health maintenance organization. Enrollment was limited to men 55 years of age and younger and women 65 years and younger. The study enrolled 171 cases, defined as subjects with at least one coronary artery lesion occupying $\ge$50% of the luminal diameter. The population control group consisted of 212 Group Health members without known coronary heart disease enrolled from September, 1987 to April, 1991. In addition, the study enrolled 80 angiographically negative controls, defined as subjects with no coronary artery lesion occupying $>$20% of the luminal diameter. All subjects had blood drawn and completed an interview. Sera were tested for IgG antibody to C. pneumoniae using the species-specific assay. The exposure of interest was the presence of IgG antibody at a titer of $\ge$ 1:8, compared to antibody $<$1:8. Adjusted odds ratios, estimated by logistic regression, were used to approximate the relative risk of disease in exposed, compared to non-exposed, subjects. Results. Because timing of blood draw was found to confound the association of interest, analysis was limited to the cases (n = 171) and those population controls (n = 120) drawn over the same time period. After adjusting for age, gender and quarter of blood draw, the odds ratio (OR) for CAD associated with the presence of antibody was 2.6 (95% confidence interval (CI) = 1.4$-$4.8). When cases and controls were further restricted to those assayed concurrently, the adjusted OR was 4.2 (95% CI = 1.8$-$10.0). Smoking status was found to be a significant effect modifier of the association between the presence of antibody and CAD after adjustment for age and gender. Among never-smokers, the OR was 0.8 (95% CI = 0.3$-$1.9) while among ever-smokers, the OR was 3.5 (95% CI = 1.7$-$7.0). There was no association between the presence of antibody and CAD using the angiographically negative controls (OR = 1.1, adjusted for age and gender).Conclusions. These results generally support the previously reported association between C. pneumoniae infection and coronary heart disease. However, caution should be used in interpreting the basis for the association.

REFERENCES

Abramson JH. Cross-sectional studies. In: Holland WW, Detels RG, Knox G, Fitzsimmons B, Gardner L, editors. Oxford Textbook of Public Health. 2nd ed. Volume 2. Methods of Public Health. New York: Oxford University Press 1991: 107-20

Alvarez-Dardet C, Gascon E, Mur P, Nolasco A. 10-year trends in the *Journal's* publications. *N Engl J Med* 1985; 312(23): 1521-2.

American Council on Education. American Universities and Colleges, 14th edition. New York: Walter de Gruyter, Inc., 1992.

Association of Universities and Colleges of Canada. Directory of Canadian Universities. Ottawa, Ontario, 1991.

Association of Universities and Colleges of Canada. 1992-1994 Directory of Canadian Universities, 29th edition. Ottawa, Ontario, 1993.

Bailar JC III, Louis TA, Lavori PW, Polansky MA. A classification for biomedical research reports. *N Engl J Med* 1984; 311: 1482-7.

Baker DR. On-line bibliometric analysis for researchers and educators. *Journal of Social Work Education* 1991; 27(1): 41-7.

Benditt J, features editor. Gender and the culture of science. *Science* 1993; 260: 383-93.

Bernier RH, Mason VM, editors. Episource. A Guide to Resources in Epidemiology. Roswell, GA: Epidemiology Monitor, 1991: 405-12.

Bollini P, Garcia Rodriquez LA, Perez Gutthann S, Walker AM. The impact of research quality and study design on epidemiologic estimates of the effect of nonsteroidal anti-inflammatory drugs on upper gastrointestinal tract disease. *Arch Int Med* 1992; 152: 1289-95.

Boss LP, Foster LR. Survey of state health agencies' staff who practice the epidemiology of noninfectious diseases and conditions. *Public Health Rep* 1994; 109(1): 112-7.

Boyer CJ. The Doctoral Dissertation as an Information Source: A Study of Scientific Information Flow. New Jersey: The Scarecrow Press, Inc., 1973.

Brady MP, Williams RE, Bailey K. A quantitative analysis of doctoral dissertation research in special education. *Teacher Education and Special Education* 1988; 11(1): 3-8.

Buck C, Llopis A, Nájera E, Terris M, editors. The Challenge of Epidemiology. Issues and Selected Readings. Washington, D.C.: Pan American Health Organization, 1988.

Burnham JF, Shearer BS, Wall JC. Combining new technologies for effective collection development: a bibliometric study using CD-ROM and a database management program *Bull Med Libr Assoc* 1992; 80(2): 150-6.

Clark BD, Hungerford LE, Bailey RR. Doctoral dissertation on hypnosis: 1923-1980. *Psychol Rep* 1984; 54: 203-9.

Cleary RE. Revisitng the doctoral dissertation in public administration: an exmination of the dissertations of 1990. *Public Adm Rev* 1992; 52(1): 55-61.

Cole P. The evoloving case-control study. *J Chronic Dis* 1979; 32: 15-27.

Commission on Professional and Hospital Activities. International Classification of Diseases, 9th Revision. Clinical Modification. Volume 1 & 2. Ann Arbor, Michigan: Edwards Brothers, Inc., 1978.

Concato J, Feinstein AR, Holford TR. The risk of determining risk with multivariable models. *Ann Intern Med* 1993; 118: 201-10.

Cooper GS, Zangwill L. An analysis of the quality of research reports in the *Journal of General Internal Medicine*. *J Gen Intern Med* 1989 May/June; 4: 232-6.

Dannenberg AL. Use of epidemiology in medical specialties: an examination by citation analysis. *Am J Epidemiol* 1985; 121(1): 140-51.

Detels R. The need for epidemiologists. *JAMA* 1979; 242(15): 1644-6.

Detels R, Breslow L. Current scope and concerns in public health. In: Holland WW, Detels RG, Knox G, Fitzsimmons B, Gardner L, editors. Oxford Textbook of Public Health. 2nd ed. Volume 1. Influences of Public Health. New York: Oxford University Press 1991: 49-65.

Detels R. Epidemiology: the foundation of public health. In: Holland WW, Detels RG, Knox G, Fitzsimmons B, Gardner L, editors. Oxford Textbook of Public Health. 2nd ed. Volume 2. Methods of Public Health. New York: Oxford University Press 1991: 285-91.

Dimitroff A. Research in health sciences library and information science: a quantitative analysis. *Bull Med Libr Assoc* 1992; 80(4): 340-6.

Faucheux BA, Minh Lim D. An analysis of doctoral training and research in gerontology in France, from dissertations defended in institutions of higher learning, 1971-1985. *Educ Gerontol* 1989; 15: 297-319.

Fee E. Disease and Discovery. A History of the Johns Hopkins School of Hygiene and Public Health, 1916-1939. Baltimore: The Johns Hopkins University Press, 1987.

Feinstein AR. Clinical biostatistics. XLIV. A survey of the research architecture used for publications in general medical journals. *Clin Pharmacol. Ther* 1978; 24(1): 117-25.

Feinstein AR. [editoral comment]. Classification of epidemiologic research. *J Clin Epidemiol* 1988; 41(8): 805.

Fleiss JL. Statistical Methods for Rates and Proportions. New York: John Wiley & Sons, 1981.

Fletcher RH, Fletcher SW. Clinical research in general medical journals. A 30-year perspective. *N Engl J Med* 1979; 301: 180-3.

French JL, Raykovitz J. Dissertation research in school psychology, 1978-1980. *J Sch Psychol* 1984; 22: 73-82.

Friedman GD. Primer of Epidemiology. 3rd ed. New York: McGraw-Hill, Inc., 1987.

Fromm BS, Snyder VL. Research design and statistical procedures used in *The Journal of Family Practice*. *J Fam Pract* 1986; 23(6): 564-566.

Geyman JP and Berg AO. *The Journal of Family Practice* - 1974-1983: analysis of an evolving literature base. *J Fam Pract* 1984; 18(1): 47-51.

Gordis L. Challenges to epidemiology in the next decade. *Am J Epidemiol* 1988; 128(1): 1-9

Greenberg RS. The future of epidemiology. *Ann Epidemiol* 1990; 1(2): 213-4.

Greenland S, Morgenstern H. Classification schemes for epidemiologic research designs. *J Clin Epidemiol* 1988; 41(8): 715-6.

Gunn RA, White MC, Miller GB, Conrad JL, Tyler CW Jr. State epidemiology programs and state epidemiologists: results of a national survey. *Public Health Rep* 1989; 104(2): 170-7.

Haynes RB, Mulrow CD, Huth EJ, Altman DG, Gardner MJ. More infomative abstracts revisited. *Ann Intern Med* 1990; 113(1): 69-76.

Hill DG. A longitudinal study of a cohort with past exposure to radar. The MIT radiation laboratory follow-up study (epidemiology) [dissertation abstract]. University Microfilms Inc. 1989.

Hippocrates. Airs, waters, places. WHS Jones eds. Harvard University Press, Cambridge, 1948.

Hoffman DA, McConahey WM, Diamond EL, Kurkland LT. Mortality in women treated for hyperthyroidism. *Am J Epidemiol* 1982; 115(2): 243-54.

International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals. *N Engl J Med* 1991; 324(6): 424-8.

Kantor AF, Hartge P, Hoover RN, Narayana AS, Sullivan JW, Fraumeni Jr JF. Urinary tract infection and risk of bladder cancer. *Am J Epidemiol* 1984; 119(4): 510-5.

Kelsey JL, Thompson WD, Evans AS. Methods in Observational Epidemiology. New York: Oxford University Press, Inc., 1986.

Kleinbaum DL, Kupper LL, Morgenstern H. Epidemiologic Research. Principles and Quantitative Methods. California: Lifetime Learning Publications, 1982.

Kramer MS, Boivin JF. Toward an "unconfounded" classification of epidemiologic research design. *J Chron Dis* 1987; 40(7): 683-8

LaPorte RE, Cresanta JL. Research trends in the *American Journal of Epidemiology*. *Am J Epidemiol* 1980; 111(2): 137-41.

Larson E, Dear M, Keitkemper MM. Comparison of two schema for classifying nursing research. *Image J Nurs Sch* 1991; 23(3): 167-70.

Last JM. Epidemiology and health information. In: Last JM, Chin J, Fielding JE, Frank AL, Lashof JC, Wallace RB, editors. Maxcy-Rosenau Public Health and Preventive Medicine. Twelfth edition. Connecticut: Appleton-Century-Crofts, 1986: 9-74.

Last JM, editor. A Dictionary of Epidemiology, 2nd ed. New York: Oxford University Press, Inc., 1988.

Leahy MJ, Habeck RV, Fabiano RJ. Doctoral dissertation research in rehabilitation: 1982-1983. *Rehab Couns* 1988; 32: 161-7.

Leahy MJ, Van Tol B, Habeck RV, Fabiano RJ. Doctoral dissertation research in rehabilitation: 1986-87. *Rehab Couns* 1990; 33(4): 315-25.

Leahy MJ, Habeck RV, VanTol B. Doctoral dissertation research in rehabilitation: 1988-1989. *Rehab Couns* 1992; 35(4): 253-63.

Levin M. Data Report on Applicants, New Enrollments and Students, Fall 1992 and Graduates and Expenditures, 1991-92 with Trends Analysis for 1974-1975 through Fall 1992. Washington, D.C.: Association of Schools of Public Health, 1993.

Lilienfeld AM and Lilienfeld DE. What else is new? An historical excursion. *Am J Epidemiol* 1977; 105(3): 169-179.

Lilienfeld DE, Lilienfeld AM. Epidemiology: a retrospective study. *Am J Epidemiol* 1977; 106(6): 445-59.

Lilienfeld AM, Lilienfeld DE. A century of case-control studies: progress? *J Chronic Dis* 1979; 32: 5-13.

Logan RFA, Kay CR. Oral contraception, smoking and inflammatory bowel disease-findings in the Royal College of General Practitioners Oral Contraception Study. *Int J Epidemiol* 1989; 18(1): 105-7.

Loomis ME. Emerging content in nursing: an analysis of dissertation abstracts and titles. 1976-1982. *Nurs Res* 1985; 34(2): 113-9.

Maclure M. Taxonomic axes of epidemiologic study designs: a refutationist perspective. *J Clin Epidemiol* 1991; 44(10): 1045-53.

MacMahon B, Pugh TF. Epidemiology. Principles and Methods.Boston: Little, Brown and Company, 1970.

Marvel MK, Staehling S, Hendricks B. A taxonomy of clinical research methods: comparisons of family practice and general medical journals. *Fam Med* 1991; 23(3): 202-7.

Mausner JS, Kramer S. Epidemiology- An Introductory Text. Philadelphia: W.B. Saunders Company, 1985.

Mellon Institute of Science and Carnegie Institute of Technology, Graduate Studies, 1970-1972. Pittsburgh: Carnegie-Mellon University, 1970.

Miettinen OS. Striving to deconfound the fundamental of epidemiologic study design. *Journal of Clinical Epidemiology* 1988; 41(8): 709-13.

Minister of Supply and Services. Canada Yearbook. Ottawa, Ontario, 1985, 1987, 1989, 1991.

Moore JL, Birren JE. Doctoral training in gerontology: an analysis of dissertations on problems of aging in institutions of higher learning in the United States, 1934-1969. *J Gerontol* 1971; 26(2): 249-57.

Mueller JE, Birren JE. An analysis of doctoral dissertations on aging written in institutions of higher learning in the United States, 1969-1975. *J Gerontol* 1980; 35(5): 788-93.

NASCET Collaborators. Beneficial effect fo carotid endarterectomy in symptomatic patients with high-grade carotid stenosis. *N Engl J Med* 1991; 325: 445-53.

National Center for Health Statistics. Health, United States, 1992. Hyattsville, Maryland Public Health Service. 1993.

Personal communication with Suzanne Eichner, Registrar's Office, John's Hopkins School of Public Health and Hygiene, Baltimore, 1994.

Personal communication with Delphine Lewis, Manager, Manuscript Services, University Microfilms Inc., Ann Arbor, Michigan, 1994.

Persson O. Online bibliometrics. A research tool for every man. *Scientometrics* 1986; 10(1-2): 69-75.

Pratt GF. A decade of AIDS literature. *Bull Med Libr Assoc* 1992; 80(4): 380-1.

Pritchard A. Statistical bibliography or bibliometrics? *J Docum* 1969; 25: 348-9.

Repp JM, Glaviano C. Dissertations: a study of the scholar's approach. *Coll Res Libr* 1987 March; 148-159.

Rothman KJ. Modern Epidemiology. Boston: Little, Brown and Company, 1986.

Ruiz MT, Alvarez-Dardet C, Vela P, Pascual E. Study designs and statistical methods in rheumatological journals: an international comparison. *Br J Rheumatol* 1991; 30: 352-5.

Sartwell PE, Stark F. *American Journal of Epidemiology*: its evolution since 1965. *Am J Epidemiol* 1991; 134(10): 1041-6.

Savitz DA, Tolo KA, Poole C. Statistical significance testing in the *American Journal of Epidemiology*, 1970-1990. *Am J Epidemiol* 1994; 139(10): 1047-52.

Schoenbach VJ, Reynolds GH, Kumanyika SK. Racial and ethnic distribution of faculty, students, and fellows in US epidemiology degree programs, 1992. *Ann Epidemiol* 1994; 4: 259-65.

Schubert-Subbarathnam H, Rougemont A. A descriptive analysis of articles in the *Journal of Social and Preventiv Medicine (Soz Praventivmed)*: 1980-1990. *Soz Praventivmed* 1992; 37: 151-8.

Schuckman H. PhD recipients in psychology and biolgoy. Do those with dissertation advisors fo the same sex publish scholarly papers more frequently? *Amer Psychol* 1987; 41(11): 987-92.

Smith RD. Veterinary clinical research: a survey of epidemiologic study designs and clinical issues appearing in a practice journal. *Acta Vet Scand* 1988; 84(Suppl.): 504-6

Snedecor GW, Cochran WG Statistical Methods, eighth edition. Ames, Iowa: Iowa State University Press, 1989.

SPSS Inc. SPSS Data Entry II. Chicago, Illinois, 1987.

SPSS Inc. SPSS/PC+, V5.0. Chicago, Illinois, 1992.

Terris M. The Society for Epidemiologic Research (SER) and the future of epidemiology *Am J Epidemiol* 1992; 136(8): 909-15.

Thacker SB, Mayberry RM, Herndon JL, Warren RC. Survey of graduates of the epidemic intelligence service as an approach to enhancing ethnic diversity among the nation's epidemiologists. *Public Health Rep* 1992; 107(6): 718-23.

University Microfilms Inc. Dissertation Abstracts Ondisc. Ann Arbor, MI., 1989.

University Microfilms Inc. User Guide. Dissertation Abstracts Online. How to Use the Online Dissertation Database Step-by-Step. Ann Arbor, MI., 1991.

U.S. Schools of Public Health and Graduate Programs Accredited by the Council on Education for Public Health. *Am J Public Health* 1993; 83(3): 447-8

U.S. Schools of Public Health and Graduate Programs Accredited by the Council on Education for Public Health. *Am J Public Health* 1994; 84(4): 694-5.

What and who is an epidemiologist? [editorial]. *Am J Public Health* 1941; 32: 414-5.

White KL. Healing the Schism. Epidemiology, Medicine, and the Public's Health. New York: Springer-Verlag, 1992.

Williams SJ, Tyler CW Jr, Clark L, Coleman L, Curran P. Epidemiologists in the United States: an assessment of the current supply and the anticipated need. *Am J Prev Med* 1988; 4(4): 231-8.

Woernle CH. Staff patterns of epidemiologists in the health departments of 12 southern states. *Public Health Rep* 1991; 106(5): 583-6.

Wood JB. The growth of scholarship: an online bibliometric comparison of dissertations in the sciences and humanities. *Scientometrics* 1988; 13(1-2): 53-62.

World Health Organization. International Statistical Classification of Diseases and Health Related Problems. Tenth Revision. Volume 2. Instruction Manual. Geneva, 1993.

VITA

NAME:	Deborah Lynn Morin
PLACE OF BIRTH:	Sudbury, Ontario
YEAR OF BIRTH:	1967
POST-SECONDARY EDUCATION AND DEGREES:	University of Waterloo Waterloo, Ontario Sept 1986 - May 1991 B.Sc. (Health Studies) The University of Western Ontario London, Ontario Sept 1991 - May 1995 M.Sc (Epidemiology)
RELATED WORK EXPERIENCE:	Research Assistant Department of Epidemiology and Biostatistics The University of Western Ontario 1993-1994 Teaching Assistant Department of Epidemiology and Biostatistics The University of Western Ontario 1992-1993

PUBLICATIONS:

Borrie MJ, Østbye T, Morin DL and Hunskaar S. Urinary and fecal incontinence in Canadian long-term care institutions prevalence and associated factors. May 1995. (submitted for publication).

SCIENTIFIC PRESENTATIONS AND ABSTRACTS:

Morin DL, Østbye T, and Robertson J. "The Epidemiology of Epidemiology Doctoral Research Trends in Epidemiology, 1982-1992", Presented at the 122nd Annual Meeting and Exhibition of the American Public Health Association, Public Health and Diversity. Opportunities for Equity, Washington, DC USA, October 30-November 3, 1994.

END

20-11-95

FIN