MASTER'S THESIS GUIDE

This guide incorporates both Epidemiology Department and Graduate School requirements. Discussion includes; topic development, Human Subjects training, roles of the thesis committee and chair, formatting, writing and revising, submission.

Developing and Completing Your Epidemiology MS or MPH Thesis (and Surviving to Tell About It)

Table of Contents

Developing and Completing Your Epidemiology M.S. or M.P.H. Thesis (and Surviving to Tell About It)*

S.M. Schwartz (Professor) and T.D. Koepsell (Professor Emeritus)
Department of Epidemiology
School of Public Health
University of Washington

I. Introduction

A. Learning objectives for the M.S. and M.P.H. include acquiring

1. Breadth of knowledge about health and health care (M.P.H. only)
2. Depth in a particular area of the field relevant to the student's interests and career path
3. Skills, including analytic and communicative

B. SPH-wide coursework and practicum requirements (M.P.H.) are aimed at providing breadth
C. Department of Epidemiology requirements are aimed at providing depth and skills
D. Your thesis is a culminating project meant to demonstrate your command of the learning objectives, particularly with respect to depth of knowledge and methodologic skills

II. Practical Issues Before You Begin Your Thesis Research

A. What is an acceptable thesis in the Department of Epidemiology? Your thesis committee ultimately decides. But,

1. In broad terms, the thesis must:

a. Involve generation of new knowledge
b. Answer one or more significant research questions about disease etiology, prognosis, early detection, etc.
c. Apply epidemiologic principles and methods, including study design and data analysis;

2. The SPH defines the MPH thesis as follows:

"The MPH thesis is an original research study that uses rigorous methods that are appropriate to the research question, generates new knowledge, applies concepts and methods from disciplines relevant to public health, and is presented in a scholarly format. The thesis demonstrates the student's comprehensive knowledge of the substantive area of the study and the research methods used. It also

represents the culminating product of the master's program in which students are expected to integrate and apply the concepts and methods learned in coursework."

3. A thesis may not be:

    a. A literature review (although a formal meta-analyses may be acceptable)
    b. Qualitative research
    c. A group effort (other than being provided with feedback by your mentors)

4. A thesis does not need to, but may:

    a. Involve new data collection
    b. Be a significant extension of a class project (e.g., EPI514)

B. How do I find a topic?

1. Should be something you're very interested in

2. Meet and talk over possibilities with faculty working in the general area of your interests

    a. Can start with your academic advisor or program director
    b. Get names of additional possibly helpful people from each person you talk with
    c. Often possible to piggy-back onto an existing project or use previously collected data
    d. Can get useful feedback about possible research questions and study designs
    e. Process also helps you decide who you want as members of your committee
    f. See the Epidemiology MPH or MS Student Handbook for the Master's Supervisory Committee requirements.

3. Worth spending the effort to develop one or a few (e.g., 1-4) well-focused research questions

    a. "Too broad" is often more of a problem than "too narrow"
    b. Well-focused question greatly facilitates planning, conducting, and writing up the thesis (or any) research
    c. Primary goal of thesis is educational—be willing to acknowledge that this work most likely will not be your magnum opus, since you will be working with limited resources and (usually) under a limited time frame

4. Can arrange with a faculty member to earn Independent Study (EPI 600) credits to look for a thesis topic. See next section about how to do so.

5. Can arrange to take Thesis (EPI 700) with your thesis chair to do background reading or pilot work on a project you might want to use for your thesis

    a. Negotiate specifics of what is required for the credits with the faculty member
    b. Aim for a tangible deliverable that your chair can evaluate (e.g., a literature review that would serve as background for the thesis proposal)
    c. Guidelines on level of effort: 1 credit hour = 3 hours per week for one quarter

C. Human Subjects Requirements

1. You must discuss the human subjects requirements for your thesis with your Thesis Committee Chair. He or she must sign the Graduate School form: Use of Human/Animal Subjects for Theses/Dissertations. The signed form must be provided to the Epidemiology Student Academic Services Office before you may register for EPI 700 Thesis credits.

2. Human Subjects Approval
    a. The official UW human subjects regulations are on the UW Human Subjects Division (HSD) website.
    b. If you will access to information that personally identifies study subjects, the project will probably require Internal Review Board (IRB) review and approval. The type of approval required varies.
        1) If it is a new project, the project will probably require full review, which can take months.
        2) If the project already has approval, you will probably have to apply for modification to have your name added to the approval.
        3) Some projects are considered "minimal risk" and only involve expedited review.
        4) If you are not sure whether approval is needed or what type, the HSD has information to guide you.
    c. Which IRB must review the project depends on where the data come from.
        1) If the data are from a UW study, then UW Human Subjects Division must review it.
        2) It the data is from an outside source, you must apply for approval through the appropriate human subjects office for that agency (e.g. Fred Hutchinson Cancer Research Center, etc.)
        3) If your data are from another country, you must obtain approval from the appropriate agency in that country.
        4) The UW has cooperative agreements with certain IRB's in the west that authorize human subjects' research without additional UW approval. However, do not assume that a cooperative agreement exists with a particular agency. The list of cooperative agreements is on the HSD website.

5) If the UW does not have a cooperative agreement with an agency (e.g. industry, a foreign ministry of health, etc.), you must apply for approval through UW as well.

D. What is the role of the thesis committee?

1. Ensures that SPH, Graduate School and Department requirements (see Epidemiology MS or MPH Student Handbook) are met

2. Usually comprises 2 people (occasionally 3, but rarely 4, the maximum allowed by the Graduate School) with different areas of expertise relevant to your proposed research. See the Epidemiology MS or MPH Student Handbook for specific requirements.

3. Advises on:

   a. Designing the project
   b. Collecting (if necessary) and analyzing the data
   c. Interpreting and writing up the results

4. Need not meet as a group; you may well be able to interact with members individually and conduct considerable business through e-mail

5. Select members before you undertake your research

   a. Assures that faculty advice and concerns are known to you in time to incorporate them into the project
   b. It is discourteous to ask a faculty member to sign off on a completed thesis when he or she had little or no chance to influence the process

6. When circulating drafts of your thesis proposal (and thesis) to your committee, it is important to:

   a. Double-space (to make reading and commenting easier if a faculty member wants to print it out)
   b. Number the pages (provides easy way for readers to identify text referred to in a comment)
   c. Use 1 inch margins
   d. Proofread carefully to correct obvious mistakes
   e. Provide an electronic, editable version (e.g., a Microsoft WORD document); many faculty use a feature such as “track changes” to make suggested revisions and add comments directly in the document
   f. Indicate a date by which you would like to receive comments; a date 3-4 weeks after distribution is usually sufficient for a first draft. More time may be

necessary during the summer or other "holiday" seasons. Final drafts may require only a few days for faculty review, depending on how many previous versions have been circulated. Do not assume, however, that your committee members can complete their review in a time frame that works for you.

g. Set up a time to meet (or talk by phone) with each committee member to discuss the indicated changes.

E. What's the role of the thesis advisor (thesis committee chair)?

1. Serves as the primary person from whom you will get advice about how to conduct and write-up your thesis research

2. Convenes and chairs meetings of the committee, if necessary

3. Must meet Departmental requirements (See the Epi Student Handbook for your degree program.

4. Usually assumes role of academic advisor as well

5. Helps trouble-shoot issues among committee members, or between you and committee members

6. Usually has the final say as to whether, at each stage of your thesis research (proposal, final written version), you are "good to go"

F. Must you write up and submit a formal thesis proposal? Yes, for several reasons:

1. Forces you to be explicit (and ideally, focused) about your plans

2. Parts of it can be used, with some modification, for your thesis (e.g., background, methods)

3. Facilitates getting good feedback and approval from your faculty advisors before you begin work
4. It's good practice to write proposals since it is a requirement for obtaining funding for research that is the "bread and butter" of an epidemiologist's career

5. The Department requires it and has published specific proposal requirements

III. Preparing to write your thesis

A. You have considerable latitude as to:
   1. Length of the thesis
   2. Organization of thesis body (e.g., Introduction / Methods / Results / Discussion)
   3. Number and placement of figures and tables
   4. Citation format

You can exploit that latitude by tailoring thesis body as much as possible to the formatting requirements of a scientific journal

B. Should you plan to publish your thesis?
   1. Yes

   2. Serves everyone's interests:

      a. Yours
         i. Satisfaction of making a contribution to knowledge on a problem you're interested in
         ii. Builds your reputation and curriculum vita (CV)
      b. The field's—gives others the benefit of your work
      c. UW's—advertises the quality and content of the program in which you trained

C. Choose a target journal early
   1. Look again at articles you plan to cite, to see where previous related research has been published

   2. Get suggestions from your thesis committee

   3. Avoid journals not listed in PubMed

   4. Among realistic possibilities, opt for journal with more prestige, higher readership. (See [1] for discussion of *journal impact factors* and a partial list of target journals.)
   5. Read target journal's instructions for authors

      a. Usually gives word limit for whole article and sometimes for sections
      b. May give guidelines for internal format
      c. May require adherence to specific guidelines for certain types of articles (See D.3., below)

   6. Look over recent articles published in target journals

      a. Suggests normal range for length, level of detail
      b. May offer good examples on internal organization of Methods section

D. Review *Uniform Requirements for Manuscripts Submitted to Biomedical Journals* at icmje.org

1. Discusses generic issues of manuscript format

2. Helps make manuscript portable to another journal if your first choice doesn't take it

3. Provides links to guidelines and checklists for certain types of articles

Guidelines	Apply to:
CONSORT	Randomized trials
STARD	Diagnostic test evaluations
PRISMA	Systematic reviews, meta-analyses
STROBE	Observational studies in epidemiology
SAMPL	Basic statistical reporting for biomedical journals

IV. Writing your thesis

A. General advice [2, 3]

1. Don't re-invent the wheel on format: you can usually follow a generic outline that applies to most medical journal articles (see Appendix), including subsection headings

   a. Reminds you what should be included
   b. Puts pieces into the order that readers expect
   c. Breaks the overall job of writing into smaller, more manageable tasks

2. Keep your audience(s) in mind
   a. Thesis committee members*
   b. Journal reviewers*
   c. Other researchers on the topic*
   d. Rank and file journal readers

   *Will be keenly attentive to details. Don't "dumb down" your work to make it accessible to someone without research training; readers who cannot understand or appreciate the importance of methodological features will skip over them anyway.

   e. Try to give enough information to allow a knowledgeable researcher to replicate the study
   f. When rationale for some aspect of methods may not be obvious, or a compromise had to be made, don't just describe what was done; *explain why*.

3. Introduction and Methods can often be drafted before results are available, drawing heavily on your thesis proposal

4. Wise to have a final or near-final set of tables and figures in hand before starting to write the Results and Discussion sections

B. Specific suggestions on subsections of Methods section

1. Study design
    a. May be just one sentence if study followed a standard design—e.g., randomized trial, case-control study
    b. Not all studies are straightforward examples of a standard study design. If yours is not, consider falling back to a more generic design category (e.g., "observational study") in which it clearly fits, and briefly describe what was done—formation of comparison groups, basic observation sequence—without trying to assign an ill-fitting design label.
    c. Example from a study of whether timeliness of clinic appointments for newborn follow-up depended on the insurance status of a simulated "client" [4]:

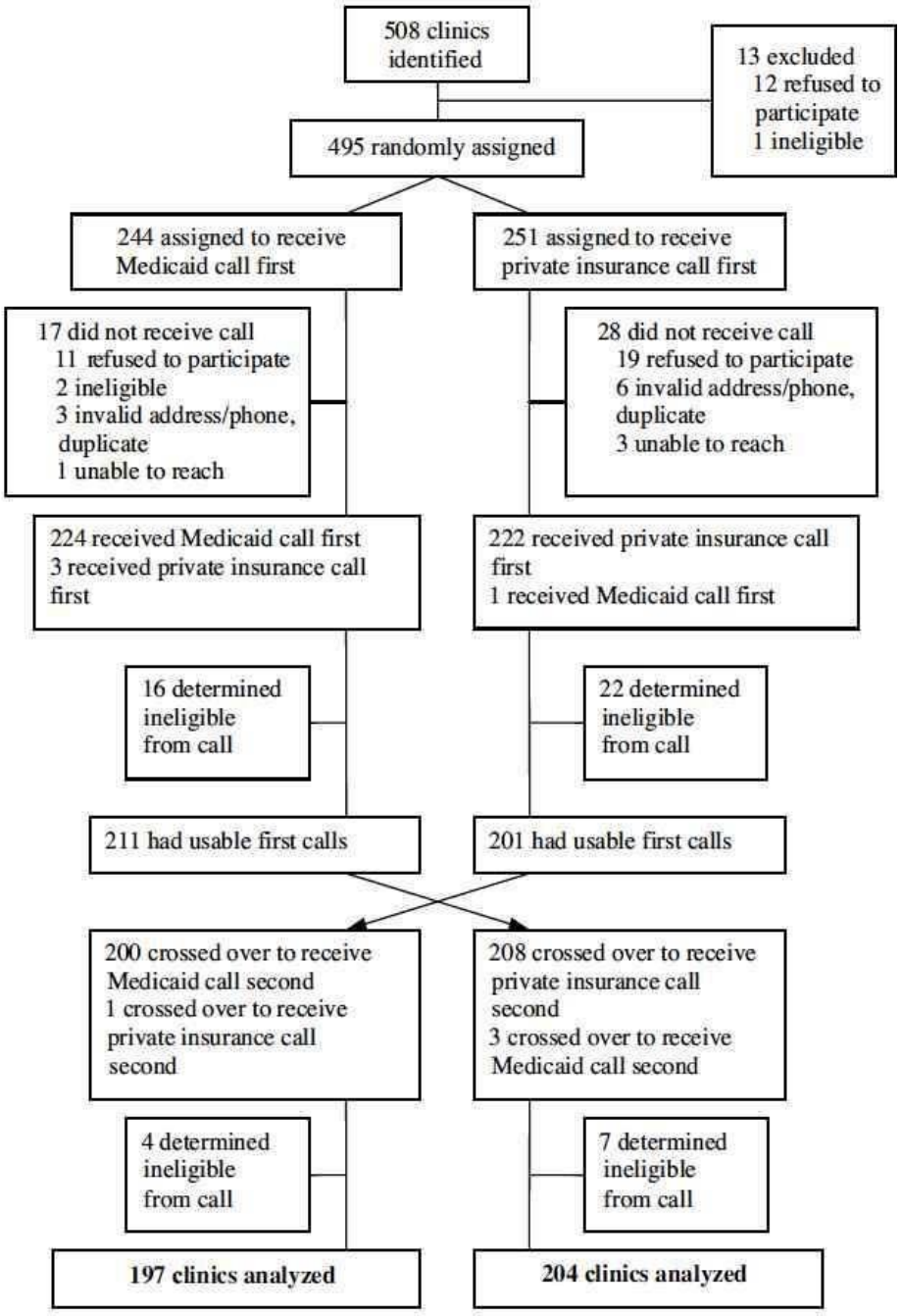
> "We used a randomized crossover study design to assess the effect of insurance status on appointment timeliness within clinics. We used a crosssectional analysis to assess appointment timeliness between clinics that did and did not accept Medicaid."

2. Study setting
    a. Provides a context for the research, which affects scope of generalizability
    b. May mention special opportunities or constraints that affected conduct of the research
    c. Example from a study of effects of emergency-visit copayment on promptness of care-seeking in patients with chest pain [5]:

> "The study was conducted at Group Health Cooperative of Puget Sound, a prepaid health plan in western Washington State with more than 500,000 enrollees who receive medical care from salaried providers. Most subscribers are insured under contracts with employers or government agencies that specify the amount of their copayment. For over 90 percent of enrollees, the emergency department copayment is not a matter of subscribers' choice."

3. Study subjects
    a. Source(s) and sampling method—affect validity and generalizability of the results
    b. Eligibility criteria

c. Formation of final analytic sample—sometimes best described with a flow diagram, a la CONSORT. (Some journals ask that this material be the initial topic in Results rather than part of Methods.)
d. Be sure to include participation proportions
e. Example of flow diagram from study noted above about timeliness of clinic appointments for newborns [4]:



4. Intervention, if any
   a. Describes what was done in ways that let readers judge its potency, resource requirements, amenability to implementation elsewhere

b. Example from a study of whether a clinic-based intervention in pediatricians' offices could increase enrollment in Head Start [6]:

"...Families of all children in the control and intervention groups 4 were given a language-appropriate telephone contact list of all Head Start agencies in the metropolitan Seattle area. For intervention children, a referral packet was also generated by computer and mailed directly to Head Start by study personnel; the packet contained a physician referral letter, including information for Head Start to contact the family; a physical examination form; and the child's immunization record. The second and third items were included only if available. Every Head Start agency in the target area participated in the project. None altered its established enrollment criteria to prioritize children from the study, and all signed a memorandum of understanding prior to study participation."

5. Data sources (or data collection)
   a. Mention only data actually used in your thesis
   b. Method(s) of measurement

      i. If a previously used instrument or scale, include references to source and information about validity or reliability
      ii. Steps taken within the study to check and promote data quality

   c. Example from a case-control study of whether screening for diabetes is effective at preventing end-organ complications [7]:

"Reviewers were not blinded to whether individuals were case subjects or control subjects. After eligibility as a study subject was verified, chart reviewers recorded every blood glucose test performed during the 10-year review period (the 10 years before the reference date). For each test, the following information was recorded: date, type of test (fasting or random, oral glucose tolerance test, hemoglobin A1c), result, whether any abnormal result had clinical follow-up, and the clinical intent or indication for the test. This last item was judged by examination of clinicians' notes and was categorized using the following guidelines:

"1. Tests for symptoms referable to diabetes. These tests occurred in the setting of classic symptoms of diabetes (e.g. polyuria, polydipsia, or polyphagia) or in the course of investigations of diseases or symptoms where diabetes might be have been a cause or underlying factor (Table 2).

"2. Tests without symptoms of diabetes. There were two subtypes of these tests. The first subtype comprised so-called 'population screening' tests, where the clinical intent was deliberately to screen for diabetes per se. The second subtype comprised so-called 'opportunistic screening' or 'case-finding' tests, where the measurement of glucose was incidental to other

clinical investigations (e.g. evaluations of acute gastrointestinal illness or follow-up of chronic diseases such as hypertension) and was not driven by concerns about diabetes.

"3. Unknown. If after medical record review the clinical intent of the test could not be categorized using the rules enumerated in (1) and (2) or was otherwise ambiguous, this classification was used.

"Data were gathered from the medical record on each subject about factors that were possible confounders. These included possible risk factors for diabetic microvascular complications that might also be associated with increased screening activity, such as body mass index (BMI) at or near to the reference date, family history of diabetes (as indicated in the medical record), and number of preventive or health maintenance visits over the review period. The presence or absence in the medical record of three comorbid states (hypertension, coronary artery disease, and hyperlipidemia) during the review period was also recorded. These conditions might also be related to the likelihood of being screened, and, at least in the case of hypertension, may be related to the likelihood of having a microvascular complication."

6. Analysis
    a. Identification of key variables and their analytic roles
    b. How key variables were defined and operationalized: e.g., categories used for analysis
    c. Statistical techniques used to obtain reported estimates, p-values, confidence limits
    d. Example from a study of physician experience with treating AIDS patients and survivorship with AIDS [8]:

"To control for improved survival due to advances in the treatment of AIDS, we grouped the dates on which patients were given diagnoses of AIDSdefining illnesses into three calendar-year periods. The first period, 1984 to 1986, preceded the availability of zidovudine and chemoprophylaxis against *Pneumocystis carinii* pneumonia, which became period, 1989 to 1994, both drug regimens were in general use and zidovudine was recommended for patients with CD4 cell counts below 500 per cubic millimeter. Previous cohort studies of HIV-infected homosexual and bisexual men have found increases in survival from the earliest to the latest of these periods.

"Severity of illness at entry into the study was determined according to a three-stage classification of AIDS-defining diagnoses developed by Turner and colleagues. Conditions such as Kaposis sarcoma are included in the category of least severe illness, moderately severe illness is defined as *P. carinii* pneumonia, and the category of most severe illness includes

diagnoses such as disseminated infection with Mycobacterium avium complex. CD4 cell counts at the time of the diagnosis of AIDS were available for 244 of the 278 patients in whom first AIDS-defining illnesses were diagnosed from 1989 to 1994 (88 percent) and were classified into four levels: 0 to 49, 50 to 99, 100 to 199, and 200 or more per cubic millimeter.

"We estimated median survival and survival curves from the time of the diagnosis of AIDS according to the patients age, the calendar period of the diagnosis, the severity of illness, the CD4 cell count at diagnosis, 6 and physician-experience category, using KaplanMeier survival analysis. Statistical significance was evaluated with the logrank test. Unadjusted and adjusted relative risks of death according to physician-experience category, the calendar period of the diagnosis, the severity of illness, and the CD4 cell count at diagnosis were estimated with Cox proportional-hazards analysis. Statistical significance for the relative risks was evaluated with the likelihood-ratio test. The test for trend in proportions was used to examine the relation between a physicians use of prophylaxis against *P. carinii* pneumonia, measurement of CD4 cells, and use of antiretroviral therapy and that physicians level of experience with AIDS. The association between the use of prophylaxis against
P. carinii pneumonia and the occurrence of *P. carinii* pneumonia as a patients AIDS-defining illness was evaluated with the chi-square test. We used generalized estimating equations to evaluate the robustness of the results with respect to the assumption of statistical independence among patients. We also examined physician-experience category as a time-dependent covariate to take into account the experience gained during the care of an individual patient with AIDS. Two-tailed P values of 0.05 or less were considered to indicate significance in all statistical tests."

V. Getting feedback and making revisions

A. Make thesis drafts "reviewer-friendly" (See II.D.6, above)

B. Share results of interim analyses with your committee members—preferably in the form of draft thesis tables, not raw computer output

C. Wait for comments on one draft before preparing another

D. If possible, discuss comments on drafts in person with thesis committee members, individually or in a group

1. Meeting date sets a deadline for completion of review
2. Better learning experience when reviewer can elaborate on the rationale behind comments
3. May provide chance to resolve conflicting advice

E. The more that each committee member has had input into the design of the study and the analysis and interpretation of the data, the easier it should be for your committee to provide feedback and the less work for you to address their concerns as you write the thesis

F. Time frame for completion

1. Communicate your planned completion date to your committee well in advance
2. Three months (i.e., one quarter) advance notice is preferred
3. Allow at least 2 months between date you plan to turn your thesis into the Graduate School and date you send your first draft to your committee; earlier is better.

VI. Preparing to Submit your Thesis

A. You must electronically submit your thesis to the Graduate School.
B. The process has many steps, so start early in the quarter (set up your account, familiarize yourself with the process, selection web publication options, etc.)
C. Please note the Graduate School mandates some specific sections and formatting

VII. Publishing

A. Guidelines on authorship from the International Committee of Medical Journal Editors:

1. "Authorship credit should be based on: 1) substantial contributions to conception and design, or acquisition of data, or analysis and interpretation of data; 2) drafting the article or revising it critically for important intellectual content; and 3) final approval of the version to be published. Authors should meet conditions 1, 2, and 3."
2. Under most circumstances, committee members would qualify under these guidelines to be co-authors, and it is collegial to invite them to do so even if the project involves other co-authors who are not part of your committee

B. Dealing with results of journal review

1. If accepted: celebrate! (safely, of course)

2. If invited to revise and resubmit

a. Treat as a "foot in the door" and revise for that journal
b. Will be expected to include a possibly lengthy cover letter with the revision, containing a point-by-point response to reviewers' comments

i. Unless following a reviewer suggestion would actually do damage, try to honor it
ii. If a reviewer's comment was based on misinterpretation, try to improve and clarify wording to prevent this happening to other readers

iii. If following a reviewer's advice would do damage, provide a careful explanation of why
iv. Always use a respectful tone when writing your cover letter

3. If rejected

a. Don't take it personally—you're in good company
b. Scavenge for useful suggestions in any reviewers' comments provided by journal
c. Try to submit promptly to another journal (which you might have in mind already)
d. Persistence pays

References

1. Garfield E. The history and meaning of the journal impact factor. JAMA 2006; 295:90–3.
2. Welch HG, Froehlich GW. Strategies in writing for a physician audience. J Gen Intern Med 1996; 11:50–5.
3. Welch HG. Preparing manuscripts for submission to medical journals: the paper trail. Eff Clin Pract 1999; 2:131–7.
4. Galbraith AA, Grossman DC, Koepsell TD, Heagerty PJ, Christakis DA. Medicaid acceptance and availability of timely follow-up for newborns with Medicaid. Pediatrics 2005; 116:1148–54.
5. Magid DJ, Koepsell TD, Every NR, Martin JS, Siscovick DS, Wagner EH, et al. Absence of association between insurance copayments and delays in seeking emergency care among patients with myocardial infarction. N Engl J Med 1997; 336:1722–9.
6. Silverstein M, Mack C, Reavis N, Koepsell TD, Gross GS, Grossman DC. Effect of a clinic-based referral system to head start: a randomized controlled trial. JAMA 2004; 292:968–71.
7. Schellhase KG, Koepsell TD, Weiss NS, Wagner EH, Reiber GE. Glucose screening and the risk of complications in Type 2 diabetes mellitus. J Clin Epidemiol 2003; 56:75–80. 8. Kitahata MM, Koepsell TD, Deyo RA, Maxwell CL, Dodge WT, Wagner EH. Physicians' experience with the acquired immunodeficiency syndrome as a factor in patients' survival. N Engl J Med 1996; 334:701–6.
9. Browner WS. Publishing and presenting clinical research. Baltimore, MD: Lippincott Williams & Wilkins, 1999.
10. Day RA. How to write and publish a scientific paper. Phoenix, AZ: Oryx Press, 1998.
11. Lang TA, Secic M. How to report statistics in medicine. Annotated guidelines for authors, editors, and reviewers. Philadelphia, PA: American College of Physicians, 1997.
12. Rivara FP, Cummings P. Writing for publication in Archives of Pediatrics & Adolescent Medicine. Arch Pediatr Adolesc Med 2001; 155:1090–2.

Appendix: Generic Thesis Outline
(Not all headings relevant to all theses)

I. Introduction (usually 1-2 pages)

A. Why is the problem important?
B. What important knowledge gaps remain despite previous work?
C. What specific research question(s) did your project address?

II. Methods (2-4 pages)

A. Study design
B. Setting
C. Study subjects
    1 Source of subjects
    2 Sampling method
    3 Criteria for eligibility
    4 Number, response proportions

D. (Description of intervention, if any)
E. Data collection (or data sources)
    1 Sources: questionnaire, interview, medical record review, vital records, etc.
    2 Protocol for a typical subject
    3 Steps taken to assess and assure data quality

F. Analysis
    1 Definition of key analytic variables, if not obvious
    2 Statistical methods used
    3 Statistical basis for sample size, if appropriate

III. Results (1-3 pages)

A. Description of study sample
B. Table(s) or figure(s) addressing each research question
C. Text used to highlight (not to repeat verbatim) results shown in tables and figures

IV. Discussion (3-5 pages)

A. Brief recap of key result(s)
B. Study strengths and limitations
C. How key results compare or contrast with previous work
D. Implications
    1 For theory
    2 For public health practice or clinical practice
    3 For future research