T. D. Koepsell
5/10/10

Writing Your Thesis

I. Preparing to write

A. Review UW Graduate School's *Format Guidelines for Theses and Dissertations* at:
http://www.grad.washington.edu/students/thesis-dissertation/format-guidelines/index.shtml

1. Contains specific requirements and recommendations concerning margins, spacing, font size, page order, etc.
2. Significantly revised in December, 2009, mostly allowing more flexibility, so don't use old theses as a format model
3. Leaves considerable latitude to authors as to:
   a. Length of the thesis
   b. Organization of thesis body (e.g., Introduction / Methods / Results / Discussion)
   c. Number and placement of figures and tables
   d. Citation format
4. Can exploit that latitude by tailoring thesis body to format requirements of a scientific journal

B. Should you plan to publish your thesis?

1. Yes.
2. Serves everyone's interests:
   a. The field: gives others the benefit of your work
   b. You
      i. Provides satisfaction of making a contribution to knowledge on a problem you're interested in
      ii. Builds your reputation and curriculum vitae
   c. UW: increases awareness of the quality and content of the program in which you trained
3. Many departments encourage students to write their thesis in format of a journal article anyway, so can take advantage of the opportunity

C. Choose a target journal early

1. Which journal?
   a. Look again at articles you plan to cite, to see where previous related research has been published
   b. Get suggestions from thesis committee
   c. Avoid journals not listed in PubMed

d. Among realistic possibilities, opt for journal with more prestige, higher readership. (See [1] for discussion of *journal impact factors* and partial list.)

2. Read target journal's instructions for authors
   a. Usually specify a word limit for whole article and sometimes for sections
   b. May provide guidelines for internal format
   c. May require adherence to specific guidelines for certain types of articles—e.g., CONSORT for randomized trials
3. Look over recent articles published there
   a. Suggests normal range for length, level of detail, number of tables and figures usually allowed
   b. May offer good examples on internal organization of Methods section

D. Also review *Uniform Requirements for Manuscripts Submitted to Biomedical Journals* (www.icmje.org)

1. Largely adopted by most major health journals
2. Discuss authorship and many generic issues of manuscript format
3. Help make manuscript portable to another journal if first choice doesn't take it
4. Provide links to guidelines and checklists for certain types of articles, which some journals require

Guidelines	Apply to:
CONSORT	Randomized trials
STARD	Diagnostic test evaluations
QUOROM	Systematic reviews, meta-analyses
STROBE	Observational studies in epidemiology
MOOSE	Meta-analyses of observational studies in epidemiology

II. Writing

A. General advice [2, 3, 12, 13, 15]

1. Don't re-invent the wheel on format: 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 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 a lay reader or a clinician without research training; readers who cannot understand or appreciate the importance of methodological features will skip over them anyway.

3. Introduction and Methods can often be drafted before results are available, drawing heavily on thesis proposal
4. In Methods section, try to give enough information to allow a knowledgeable researcher to replicate the study
5. 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*.

B. Suggestions on subsections of Methods section

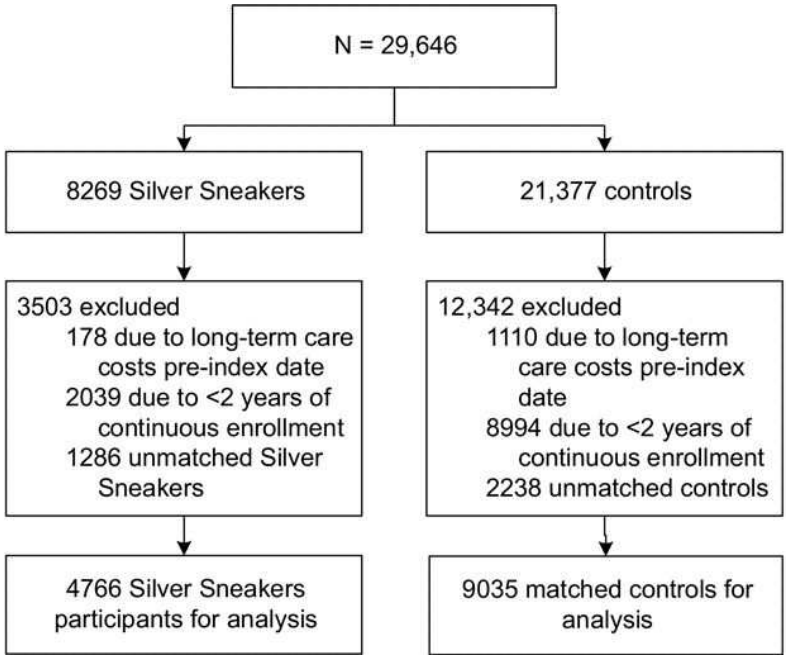
1. Study design type
   a. Usually very short, to orient reader as to generic type of study design used. May be just one sentence if study followed a standard design—e.g., randomized trial, case-control study, case series
   b. Not all studies are straightforward examples of a standard study design. 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 cross-sectional analysis to assess appointment timeliness between clinics that did and did not accept Medicaid."

2. 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 co-payment 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, which affect representativeness and generalizability of the results
    b. Eligibility criteria (including, e.g., case definition, if appropriate)
    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. Example of flow diagram from a study of depression in relation to participation in an exercise program for older adult enrollees in an HMO [6]:



4. Intervention, if any
    a. Describes what was done in ways that let readers judge intervention's potency, resource requirements, amenability to implementation elsewhere
    b. Example from a randomized trial of whether a clinic-based intervention in pediatricians' offices could increase enrollment in Head Start [7]:

> "…Families of all children in the control and intervention groups 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 collection (or Data sources, if pre-existing data used)
    a. Mention only data actually used in present paper
    b. Method(s) of data gathering: e.g., structured questionnaires, focus groups, medical record abstracts, laboratory data
    c. If a previously used instrument or scale, references to source and information about validity or reliability
    d. Example of description of data sources from a case-control study of whether screening for diabetes is effective at preventing end-organ complications [8]:

> "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."

e. Steps taken within the study to check and promote data quality
f. Example from a qualitative research study involving interviews with physicians about physician-assisted suicide [9]:

> "Trustworthiness, a qualitative research concept akin to reliability and validity, was ensured by several steps. These included: (1) ongoing review of interview technique and data with a medical anthropologist (L.A.R.), (2) review of a sample of interview transcripts by 2 of us (L.A.R. and R.A.P.) to confirm the generated themes, (3) review of the work for accuracy by 50% of the sample physicians, and (4) professional peer review in different settings, such as work-in-progress sessions and hospital seminars."

6. Analysis
   a. How important variables were defined and operationalized: e.g., categories used for analysis
   b. Statistical techniques used to obtain reported estimates, p-values, confidence limits
   c. Example from a study of physician experience with treating AIDS patients and survivorship with AIDS [10]:

> "To control for improved survival due to advances in the treatment of AIDS, we grouped the dates on which patients were given diagnoses of AIDS-defining 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 Kaposi's 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, 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."

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

III. Getting feedback and making revisions

A. Make thesis drafts reviewer-friendly

1. Use 12-point or larger font
2. Use at least 1-inch margins
3. Run text through a spell-checker, but also…
4. Proof it yourself to correct obvious errors
5. If provided to thesis committee members in *electronic* form
    a. Word doc-files most convenient, to allow comments and suggestions using the Track Changes feature
    b. If a PDF file provided instead, number the pages
6. If provided to thesis committee members in *paper* form
    a. Double-space (to allow room for mark-ups by those who wish to do so on a paper copy)

b. Number the pages

B. Allow a reasonable amount of time for readers to review

C. Wait for comments from all recipients of one draft before preparing another draft

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. Group meeting may provide chance to resolve conflicting advice

IV. Publishing

A. Authorship

1. You should be first author
2. Guidelines on authorship from the International Committee of Medical Journal Editors (http://www.icmje.org):

> "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."

3. Thesis committee members would usually meet above criteria and qualify as co-authors, but not automatic
4. Other people may qualify as co-authors even if not on thesis committee
5. First author usually decides on who qualifies as a co-author and order in which they are listed, which should reflect overall scientific contribution to paper

B. Maximizing chances of acceptance for publication

1. Methodology matters [11]
2. Hard-nosed internal review among authors or by colleagues before submission

C. Journal peer review process

1. Manuscripts can only be under consideration by one journal at a time. Most journals insist on author's written assurance on this point.
2. Submitted manuscripts usually pre-screened by editor or an associate editor
   a. Checks whether subject matter appropriate for journal
   b. Looks for obvious major problems of methodology
   c. Considers who might be suitable peer reviewers
   d. Views manuscript in relation to its "competition"

    e. Can return manuscript promptly (thus saving everyone's time) if significant problem(s) found

3. If manuscript passes editor's pre-screening, it is sent out to peer reviewers (usually 2-3). Candidate peer reviewers often identified by:
    a. Editor's knowledge of who works in subject area
    b. Authors of related past articles in same journal
    c. Researchers whose work is cited in manuscript's bibliography
    d. Mix of perspectives: subject-matter expert, methodologist, etc.
    e. Suggestions from authors themselves, if journal's instructions for authors invites authors to nominate their own reviewers
4. Turn-around time varies substantially among journals, ranging from ∼2–4 weeks to ∼3 months. Quicker is not necessarily better.
5. Many journals provide a website where authors can monitor status of their manuscript, but OK to contact journal's editorial office if review process appears to be stalled

D. Dealing with results of journal review

1. If accepted: celebrate! (safely, of course)
2. If invited to revise and resubmit
    a. Usually wise to regard as a "foot in the door" and revise for that journal
    b. 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 suggestion was based on misinterpretation, try to improve and clarify wording anyway to prevent this happening to other readers
        iii. If following a reviewer's advice would do damage, use cover letter to provide a careful, respectful explanation of why
3. If rejected
    a. Hardly ever worthwhile or successful to challenge journal's verdict
    b. Don't take it personally—you're in good company
    c. Scavenge for useful suggestions in any reviewers' comments provided by journal
    d. Try to submit promptly to another journal (which you might have in mind already)
    e. 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. Nguyen HQ, Koepsell T, Unutzer J, Larson E, LoGerfo JP. Depression and use of a health plan-sponsored physical activity program by older adults. Am J Prev Med 2008; 35:111–117.

7. 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.

8. 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.

9. Kohlwes RJ, Koepsell TD, Rhodes LA, Pearlman RA. Physicians' responses to patients' requests for physician-assisted suicide. Arch Intern Med 2001; 161:657–663.

10. 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.

11. Lee KP, Boyd EA, Holroyd-Leduc JM, Bacchetti P, Bero LA. Predictors of publication: characteristics of submitted manuscripts associated with acceptance at major biomedical journals. Med J Aust 2006; 184:621–626.

12. Browner WS. Publishing and presenting clinical research. Baltimore, MD: Lippincott Williams & Wilkins, 1999.

13. Day RA. How to write and publish a scientific paper. Phoenix, AZ: Oryx Press, 1998.

14. Lang TA, Secic M. How to report statistics in medicine. Annotated guidelines for authors, editors, and reviewers. Philadelphia, PA: American College of Physicians, 1997.

15. Rivara FP, Cummings P. Writing for publication in *Archives of Pediatrics & Adolescent Medicine*. Arch Pediatr Adolesc Med 2001; 155:1090–2.

16. Binswanger IA, Stern MF, Deyo RA, Heagerty PJ, Cheadle A, Elmore JG, et al. Release from prison–a high risk of death for former inmates. N Engl J Med 2007; 356:157–165.

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

I. Introduction

   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

   A. Type of study design used (e.g., cohort study, focused ethnographic interviews)
   B. Setting
   C. Study subjects
      1. Source of subjects
      2. Sampling method
      3. Criteria for eligibility
      4. Number, response rates
   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

   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

   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