Summary
The overall purpose of this thesis has been to contribute to the systematic review evidence base for CAM interventions. As discussed in the General Introduction Chapter (Chapter 1), this is important because in this era of evidence-based medicine, physicians and policy-makers require evidence from systematic reviews to inform decisions about the value of acupuncture and other CAM interventions for treating specific health conditions. In addition, CAM interventions are widely used by the public at considerable expense, so evidence to inform the public’s decisions about the benefits and risks of specific CAM interventions is valuable. Although the chapters that comprise this thesis all relate to the broad topic of systematic reviews and CAM, these chapters can be aggregated into two Parts. The key objective of Part 1 (Chapters 2-4) of this thesis has been to develop and analyze CAM systematic review-related databases. The key objectives of Part 2 (Chapters 5-9) have been to prepare systematic reviews to evaluate the effects of acupuncture for three specific health conditions, as well as to conduct research with a view to improve the methods of future acupuncture RCTs and systematic reviews.

The aim of Chapter 2 has been to develop a convenient way for users of The Cochrane Library (researchers, clinicians, consumers) to identify Cochrane reviews that are CAM-related, and to easily find Cochrane reviews on specific CAM interventions. To address this aim, a Field ‘topics list’ was developed, which compiled and categorized all 396 CAM-related Cochrane reviews (as of The Cochrane Library, Issue 4, 2009). As described in Chapter 2, the first step in developing this topics list was developing a standardized “operational” definition, to provide an objective, reproducible, and systematic method for defining and classifying CAM therapies. Chapter 2 also discusses the challenges in developing the topics list, including developing the operational definition of CAM, deciding which reviews should be included within the CAM Field’s scope, developing the structured list of CAM Field-specific topics, and determining where in the topics list the reviews should be placed. Chapter 2 concludes with a discussion of remaining challenges and opportunities related to this topics list, for instance identification of gaps in Cochrane systematic review coverage of CAM therapies.

To facilitate and ease the identification of CAM trials by systematic reviewers, a specialized register of citations to CAM-related controlled trials has been developed. The first objective of Chapter 3 was to describe the sources and methods for developing the specialized register. The second objective was to analyze the trial citations included in the register, on the following characteristics: publication dates, languages, publication journals, presence in MEDLINE, the type of CAM intervention addressed, and type of medical condition addressed. At the time of analysis [March 2012], the CAM Field trials register included 44,840 citations of CAM-related controlled trials. Most citations (60%) were from 2000 or later, and the majority (71%) were reported in English; the next most common language was Chinese (23%). The journals with the greatest number of citations were CAM journals published in Chinese, and non CAM nutrition journals published in English. More than one-third of the register citations (36%) were not indexed in MEDLINE. The most common CAM intervention in the register was non-vitamin, non-mineral natural products (34%), followed by Chinese herbal medicine (27%). The implications for CAM systematic reviewers of the improved accessibility of these difficult to locate CAM trials, many of which are of low methodological quality and may overestimate treatment effects, are reviewed in the Discussion section of Chapter 3.

Chapter 4 describes a review overview of all Cochrane reviews on the topic of traditional Chinese medicine (TCM). In order to identify reviews focusing on TCM, the titles and abstracts of all reviews in Issue 4, 2008 of the Cochrane Database of Systematic Reviews were searched. For each review, data on the number of trials included and the total number of participants were extracted. An indication of the strength of the review findings was provided by assessing the ‘Authors’ Conclusions’ section of the systematic review abstract. This assessment of the Authors’ Conclusions
was supplemented with a listing of the comparisons and outcomes showing statistically significant meta-analyses results. Seventy Cochrane reviews of TCM were identified, primarily acupuncture (n=26) and Chinese herbal medicine (n=42). The Authors’ Conclusions of 7/26 acupuncture reviews and 20/42 herbal medicine reviews indicated a suggestion of benefit, which was qualified by a caveat about the poor quality and quantity of studies. The remainder of the reviews were inconclusive, due to the poor methodology and heterogeneity of the studies reviewed. The topics of the reviews that indicated a suggestion of benefit might be prioritized for future rigorous randomized controlled trials. As further discussed in the General Discussion (Chapter 10), classifying the bottom line findings of these systematic reviews was a challenging and complex endeavor that will require much additional research. Although the classification methods used in Chapter 4 may have limitations, the methods used may potentially serve to inform future research efforts to classify the overall findings of systematic reviews.

Chapter 5 is a book chapter that summarizes methodological research related to the preparation of CAM-related systematic reviews, and also suggests approaches for addressing some of the challenges in systematically reviewing CAM interventions. Chapter 5 begins by explaining how systematic reviews ensure rigorous quality standards and maintain objectivity during each phase of review preparation, including: (1) identifying relevant RCTs; (2) assessing the quality of the RCTs; and (3) combining the data from the RCTs. The discussion of assessing quality of RCTs (section 2) is related to thesis Chapter 6, and section 2 discusses the advantages and limitations of contacting RCT authors to request information about randomization and blinding procedures not described in RCT publications. The discussion of combining data from RCTs (section 3) uses the acupuncture for low back pain review (Chapter 7) as a case study to discuss the decisions that systematic reviewers need to make in terms of whether and how to statistically pool the results of similar but separate RCTs. The remainder of Chapter 5 focuses on the advantages and drawbacks of different methods for assessing treatment adequacy and practitioner adequacy in CAM trials. Systematic review methodology is evolving, and advice for addressing methodological challenges can change in light of new research findings, or in light of additional experience gained in preparing systematic reviews. For example, in Chapter 5, a holistic approach to assessing treatment adequacy had been advocated. However, as described in the General Discussion (Chapter 10), the experience in preparing the three systematic reviews comprising this thesis led to a revision of this earlier view about the optimal approach for assessing treatment adequacy. Namely, the additional experience gained has led to the conclusion that an objective method of assessing individual treatment adequacy-related components has advantages over a holistic, composite approach of assessing treatment adequacy.

Chapter 6 describes a survey undertaken to evaluate whether the description of randomization and blinding in published articles is an accurate and complete reflection of the study procedures used. This question is important to address because systematic reviews generally evaluate RCTs based on the published articles. A major disadvantage of quality assessment based solely on the published articles is that this relies on a critical assumption that what is written in the article reflects actual procedure and that “if it is not reported, it probably was not done”. For this survey, the authors of 51 RCTs included in a systematic review of acupuncture for chronic pain were sent a brief questionnaire that included questions related to the following three study quality dimensions: (1) generation of allocation sequence, (2) allocation concealment, and (3) blinding of outcomes assessor. Among the 51 questionnaires sent out, 35 responses were received. Of 35 studies described as randomized in the published articles, associated survey responses indicated that four actually used quasi-randomized methods (i.e., methods of assignment not considered strictly random, such as assignment by alternation, date of birth, or medical record number). Among published articles with missing information on allocation concealment and outcomes assessor blinding, 13 of 34
studies used adequate allocation concealment and two of 10 studies were blinded, according to survey responses. With the exception of the four quasi-randomized trials described as randomized in the RCT publications, survey responses generally confirmed information about randomization and blinding already described in the investigators’ RCT publications. Thus, this survey suggested that contacting trial authors may result in obtaining previously unpublished information about methodological quality. However, the potential gains of obtaining the missing information need to be weighed against the reporting bias such efforts may introduce, particularly when asking investigators to select from a list of response choices, as was done in this survey.

Chapter 7 describes a systematic review with the objective of assessing acupuncture’s effectiveness for treating low back pain. For this systematic review, computerized databases were searched to find RCTs that assigned patients with low back pain to receive either acupuncture or one of the following comparison treatments: sham acupuncture, sham transcutaneous electrical nerve stimulation (TENS), no low back pain treatment, or traditional treatments for low back pain. The 33 RCTs that met inclusion criteria were subgrouped according to acute or chronic pain, style of acupuncture, and type of control group used. The primary outcome of this review was a quantitative synthesis of the short-term effectiveness of acupuncture on pain in each subgroup. For the quantitative synthesis, the RCTs’ standardized mean differences (SMDs) were used. These SMDs were calculated as differences in post-treatment scores between groups. For the primary outcome of short-term relief of chronic pain, the meta-analyses showed that traditional Chinese acupuncture was significantly more effective than sham acupuncture, sham TENS, and no additional treatment. For patients with acute low back pain, data were sparse and inconclusive. Data were also insufficient for drawing conclusions about acupuncture’s short-term effectiveness compared with most other therapies. The overall conclusion that acupuncture may be more effective than a sham acupuncture control for providing short-term relief of chronic low back pain was considered preliminary. This was because the comparison of acupuncture versus sham acupuncture was based on only four small to moderately sized trials, and also because there were several large and recently completed but not yet published trials that would strongly influence these findings.

Chapter 8 describes a Cochrane review with the objective of assessing acupuncture’s effectiveness for treating peripheral joint osteoarthritis. For this review, computerized databases were searched to find RCTs that compared needle acupuncture with a sham, another active treatment, or a waiting list control group in people with osteoarthritis of the knee, hip, or hand. Authors of all included RCTs were contacted to obtain additional information about their RCTs’ methods and results. For the primary outcomes of short-term improvements in pain and functioning, the effects of acupuncture relative to control were calculated as SMDs. These SMDs were calculated using the differences in improvement between groups. Sixteen RCTs involving 3,498 people were included in the systematic review. Twelve of the RCTs included only people with osteoarthritis of the knee, three only osteoarthritis of the hip, and one a mix of people with osteoarthritis of the hip and/or knee. In comparison with a sham acupuncture control, true acupuncture resulted in statistically significant, short-term improvements in osteoarthritis pain and function. However, these pooled short-term improvements did not meet our predefined thresholds for clinically relevant improvements, and furthermore, the trials that were pooled showed substantial statistical heterogeneity. This statistical heterogeneity among the pooled trials was largely explained by the adequacy of participant blinding variable, which showed a statistically significant subgroup effect. When restricting to only those sham-controlled trials that were judged to have adequate participant blinding, the pooled improvements were no longer statistically significant. Trials with a waiting list control showed statistically significant and clinically relevant benefits. Based on these findings, it was concluded that the use of different control comparators
Summary

explains the variability in the results of acupuncture for osteoarthritis trials, and that placebo effects due to patient expectations probably accounts for most of the observed benefits.

Chapter 9 describes a systematic review with the objectives of assessing adjuvant acupuncture’s overall pooled effect in increasing in vitro fertilization (IVF) pregnancy success rates, and also assessing whether study design-, treatment-, and population-related factors influence these effect estimates. For this review, computerized databases were searched to find RCTs that compared needle acupuncture administered within one day of embryo transfer, versus sham acupuncture or no adjuvant treatment. The investigators of all RCTs provided additional methodological details and outcome data not included in their original RCT publications. The primary outcome was clinical pregnancy rates. Eleven subgroup variables (five clinical and six methodological) were pre-specified to investigate sources of heterogeneity, using single covariate meta-regressions. Sixteen trials (4,038 participants) were included in the meta-analyses. There was no statistically significant difference between acupuncture and control when combining all trials, or when restricting to sham-controlled or no adjuvant treatment-controlled trials. The type of control used did not significantly explain the statistical heterogeneity. However, the no adjuvant treatment-controlled trials had a somewhat larger pooled effect than the sham-controlled trials, and the pooled effect of the no adjuvant treatment-controlled trials was statistically significant if an outlying trial was removed. Of the eleven variables pre-specified for meta-regression subgroup analyses, only the baseline (i.e., control group) pregnancy rate variable showed a statistically significant subgroup effect, and this variable explained most of the heterogeneity in results for the primary outcome measure across all trials. The subgroup finding of a benefit in trials with lower, but not higher, baseline pregnancy rates was the only statistically significant subgroup finding in our earlier review, was confirmed in this update, and was not explained by any confounding variables evaluated. However, we concluded that this subgroup benefit finding still requires a cautious interpretation because of the multiple subgroup tests conducted, the risk of unidentified confounders, the multiple different factors that determine baseline rates, and the possibility of publication bias.

This thesis concludes with a General Discussion (Chapter 10), which reflects on the findings and discusses the methodological issues that were identified while working on this thesis, and sets the agenda for future research after the completion of this thesis. Part 1 of Chapter 10 discusses the challenges of developing eligibility criteria for including studies in the databases of CAM-related reviews and trials described in Chapters 2-4. Part 1 also discusses the challenges of developing criteria for further classifying the trials and reviews selected for inclusion in these databases. Part 2 discusses Chapters 5-9, the systematic reviews that were conducted to evaluate the effects of acupuncture for three specific conditions, and methodological studies conducted in the context of these evaluations. A primary objective in preparing these three reviews (or indeed any systematic reviews), was to determine whether acupuncture ‘works’ for treating the given conditions. This determination was made based not only on the meta-analysis estimates from the pooled trials, but also the various biases of the trials, and the possibility of different effects across different subgroups of trials. Part 2 discusses the primary issues that were encountered in attempting to draw conclusions about the effects of acupuncture relative to a given control for the three conditions studied, as well as methodological problems of existing trials which impede the ability to draw conclusions about whether acupuncture ‘works’ relative to a given control. Some issues discussed in Part 2 include the calculation, interpretation, and reporting of subgroup analyses (particularly as this relates to the use of sham acupuncture versus usual care as the control intervention) and sensitivity analyses, and the methodology of the pre-requisite step for these subgroup and sensitivity analyses, which is contacting the trial authors to obtain the necessary data. Other issues discussed in Part 2 of Chapter 10 include the appraisal of the effects of the unpublished or not yet published, trials, and the interpretation of the clinical relevance of the treatment effects for real
world practice. Part 2 concludes with a brief, general overview of implications for future research of systematic reviews of acupuncture. This section largely focuses on the advantages and drawbacks of sham-controlled efficacy trials versus comparative effectiveness pragmatic trials in investigating the effects of acupuncture.