Chapter 2: Methodology

Recruitment of Experts

The EPC team identified a group of 16 experts to provide input at key points during the project (see Appendix A). These experts included representatives from our partner organization, the American Academy of Family Physicians (AAFP), and other relevant professional associations, as well as clinical specialists and allied health representatives.

The EPC team involved a core group of the experts in defining the key questions (see Identifying the Specific Questions, below) and asked the entire group of experts to participate in review of the draft report (see Peer Review Process, below).

Target Population

The main targeted users of the report are clinicians, including family physicians, internists, cardiologists, and other specialists managing patients with VTE.

Identifying the Specific Questions

The AAFP generated a list of key questions to be addressed. The EPC team conducted preliminary literature searches and formulated the questions in specific terms that would focus the review process on the most relevant published studies. The team then sent the draft questions to the core experts, asking them to rank the questions in terms of *importance* and *uncertainty* about the answers. After reviewing the experts' ratings and comments, the EPC team established the final list of key questions to address in this Evidence Report. Because some of the questions have been addressed in previous systematic reviews, each question was designated to be addressed either through review of previous systematic reviews, through review of primary literature, or through a combination of the two. This strategy enabled the EPC team to address more questions than if it had relied solely on a primary review of all original studies on each question.

Key Questions

The EPC team sought to address the following key questions as they pertained to management of DVT.

Q1. What are the efficacy and safety of LMWH compared to UFH for the treatment of DVT?

Q2. What are the efficacy and safety of LMWH compared to UFH for treatment of PE?

The experts indicated that these two questions were associated with little uncertainty but remained important questions. Given that many systematic reviews had already been done on this topic, the EPC team decided to review the quality and content of the earlier *systematic reviews.*

Q3a. What are the efficacy and safety of outpatient versus inpatient treatment of DVT with LMWH or UFH?
Q3b. What is the cost-effectiveness of outpatient versus inpatient treatment of DVT with LMWH or UFH?

The experts identified these questions as a high priority. For these questions, the EPC team decided to review the *primary literature* as well as any existing meta-analyses and cost-effectiveness analyses on this topic.

Q4. What is the optimal duration of treatment for DVT and PE in patients without known thrombophilic disorders and in patients with known thrombophilic disorders?

The experts indicated that this question was important and was associated with uncertainty. The EPC team decided to review the *primary literature* to answer this question.

Q5. How accurate are clinical prediction rules used for the diagnosis of DVT or PE?

The experts generally indicated that this question was at least moderately important and was associated with considerable uncertainty. The EPC team decided to review the *primary literature* to determine the accuracy of validated clinical prediction rules for diagnosing DVT or PE.

Q6a. What are the test characteristics of ultrasonography for diagnosis of DVT?
Q6b. Are calf vein thromboses adequately identified with ultrasound?

The experts reported that use of ultrasound was an important topic that was associated with moderate uncertainty. Because this topic has been addressed in a number of systematic reviews, the EPC team decided to review the quality and content of the *systematic reviews.*

Q7a. What are the test characteristics of helical CT for diagnosis of PE relative to V/Q scanning or standard angiography?

Q7b. What are the test characteristics of MRI and MRA for diagnosis of PE relative to V/Q scanning and/or standard angiography?

The experts reported that these two questions were very important and were associated with uncertainty. There have been systematic reviews on this topic, particularly regarding CT. For these questions, the EPC team decided to review published *systematic reviews* and update these with a review of the *primary literature* that used the most appropriate reference tests.

Q8. What are the test characteristics of D-dimer for diagnosis of VTE?

The experts indicated that this question was relatively important and was associated with moderate uncertainty. Instead of reviewing the large diffuse body of literature on this topic, the EPC team decided to review previous *systematic reviews.*

Causal Pathway

To show how the key questions relate to the overall management of patients with VTE, the EPC team developed a description of a causal pathway (Figure 1)**.** The causal pathway depicts the diagnostic and treatment course for a patient with venous thrombosis and the types of outcomes that need to be considered in management decisions. The pathway also provides a conceptual framework for linking the responses to our key questions and for identifying gaps in our knowledge about management of VTE.

Literature Search Methods

The literature search consisted of several steps: identifying sources, formulating a search strategy for each source, and executing and documenting each search.

Sources

Electronic literature sources were used to identify all studies potentially relevant to the research questions and included both electronic database searching and manual searching. Preliminary searches were performed in January to March, 2002, with followup searches in April, 2002. The following databases were searched.

MEDLINE®

MEDLINE, or MEDLARS on-line, is a database of bibliographic citations and author abstracts from approximately 3,900 current biomedical journals published in the United States and 70 foreign countries, dating back to 1966. MEDLINE was accessed through PubMed, the Internet access to the database provided by the National Library of Medicine (NLM).

Cochrane

The Cochrane Database of Systematic Reviews includes full text articles reviewing the effects of healthcare. The reviews are highly structured and systematic, with evidence included or excluded on the basis of explicit quality criteria, to minimize bias.

To ensure a comprehensive literature search, the team examined the reference lists from our database of reference material previously identified through the electronic searching, queried our technical reviewers and reviewed the tables of contents from journals cited most frequently in the literature searches (see Appendix B). The team reviewed the tables of contents of these journals published between October 2001 and March 2002.

MICROMEDEX®

The Micromedex worldwide editorial team reviews and edits all information compiled from the most current sources available. The unbiased documents are thoroughly researched, evaluated, and referenced based on the world's leading literature. Healthcare and environmental safety professionals rely on Micromedex information in over 8,000 facilities in more than 90 countries.

Search Terms and Strategies

The search strategies were designed to maximize sensitivity and were developed in consultation with Johns Hopkins University Welch Medical Library staff and team members. Preliminary strategies were developed to identify key articles. Using key articles determined to be eligible for review, search strategies were developed and refined in an iterative process. A strategy was first developed for PubMed. This strategy was then modified to create separate search strategies for the Cochrane and Micromedex electronic databases (see Appendix C).

Organization and Tracking of Literature Search

The results of the searches were downloaded from electronic sources, where possible, or manually entered into a ProCite database. (ProCite, ISI Research Soft, Berkeley, CA)The duplication check in the bibliographic software was used to eliminate articles already retrieved. This ProCite database was used to store citations and track search strategies and sources. The use of this software also allowed for the tracking of the abstract review process.

Abstract Review

As a first step in the review process, two members of the study team independently reviewed the abstracts identified by the search to exclude those that did not meet our eligibility criteria. At this step we excluded citations when: the articles did not apply to a key question, the article reported only on prevention of VTE (not treatment), the articles were not written in English, the articles did not include human data, or the articles reported on a meeting only (i.e., no full article to review). In addition, for those questions for which we reviewed primary literature, we excluded articles that did not include any original data or were case reports. For our key questions relying on review of systematic reviews, we excluded articles that did not include a systematic review, meta-analysis or cost effectiveness analysis.

The EPC team used abstract review forms appropriate for the search processes (See Appendices D and E). The forms were based on those used in previous EPC reports. Each abstract was circulated to two members of the study team who independently reviewed the abstract and indicated which of the key questions the article addressed. For those articles found not eligible, the reviewers indicated a reason for exclusion. When there was no abstract or when the reviewers could not determine from the abstract whether the article met the eligibility criteria, the team obtained a full copy of the article to review. Investigators met face-to-face to adjudicate when there were disagreements between them on study eligibility. Our process emphasized arriving at agreement on which studies met our pre-established criteria.

Qualitative and Quantitative Data Abstraction

The study team developed article review forms that were pilot tested and revised before use. These included both a quality assessment and a content abstraction form. Due to the different types of questions addressed, the team had four sets of quality and content forms (see Appendices F, G, H, and I): one set addressed key questions 3a and 4, treatment questions, and one set addressed the diagnostic testing questions, questions 5 and 7. The team developed a third set of quality and content forms to address question 3b on cost-effectiveness. The review of published systematic reviews (questions 1, 2, 6, 7, and 8) required a fourth set of forms, which were created based on our review of several systems for evaluating systematic reviews.[22-27] To make sure that all articles met eligibility criteria, the study quality form began with a check of

the eligibility criteria (see Abstract Review, above). For questions 3 and 4, the team limited the review to studies with a comparison group and a minimum sample size of five.

The quality assessment forms for diagnosis and treatment studies included items about study quality in the following categories: representativeness of study population; bias and confounding; description of therapy/testing; outcomes or test interpretation; and statistical quality and interpretation. The items in these categories were derived from study quality forms used in previous EPC projects[28,29] and were modified for this project. Because of the variety of issues covered by our key questions, not all items were required for each of the key questions.

The study team responded to each question with a score of zero (criteria not met), one (criteria partially met), or two (criteria fully met). The score for each category of study quality was the percentage of the total points available in each category for that study and therefore

could range from zero to 100 percent. As there is presently no consensus on reporting quality scores, we have reported scores by category, giving each category equal weighting. Therefore the overall quality score was the average of the five categorical scores.

The quality assessment forms for cost-effectiveness studies and systematic reviews had fewer items without category scores. The overall quality score for these articles was based on the average of the scores on the individual items.

The content abstraction form for the review of the original studies included items that described the type of study, geographical location, the definition of study groups, the specific aims, the inclusion and exclusion criteria, characteristics of tests and interactions, demographic, social and clinical characteristics of subjects, and outcomes or results related to each of the key questions.

Article Review Process

The team reviewed each eligible article identified by the abstract review process. Two reviewers independently reviewed each article. One team member was responsible for completing both the quality assessment and content abstraction forms, and the second reviewed and confirmed the material abstracted. Differences between the two reviewers in either quality or content abstraction were resolved at face to face meetings. Reviewers were not masked to author or journal names because previous work has shown that masking is unlikely to make a significant difference in the results of the data abstraction.[30]

The team developed a database to collect, maintain, and analyze the quality assessment and content abstraction data. The evidence tables were built in Microsoft Access 2000 (Copyright © 1992-9 Microsoft Corporation), with a data-entry front end developed in Delphi© (Borland Delphi, Scotts Valley, CA).

Evidence Tables

For each key question, the EPC team created a set of evidence tables. Each set of tables contained basic information about study aims and eligibility criteria, assessments of study quality, selected characteristics of study participants, and results most pertinent to the key question.

For two of the questions, we abstracted data from the studies to fill in contingency tables, and from these, calculated true positive (TPR) and false positive rates (FPR). If this primary data was not presented in an article, we abstracted only the summary statistics reported, including sensitivity, specificity, positive predictive value, negative predictive value and area under the receiver operating characteristic (ROC) curve. If the data were available, we calculated test characteristics separately for each strata of pretest probability, or for each test cutoff for which data was provided. The area under the ROC curve was measured using ROCFIT©, (Chicago,

IL).

For the question regarding the utility of clinical prediction rules, we plotted the true positive rates and false positive rates from several studies to create a summary ROC curve. For this analysis, we used as a cutoff the score that separated patients with a low pretest probability of DVT from those in the moderate and high categories. In our analyses of the utility of CT and MRI, we also prepared a summary ROC curve. We specified that the TPR and FPR be from analyses that used data from all the participants in the study and be data points which represented the best test performance of cutoffs studied.

Evidence Grades

Five members of the EPC team independently graded the strength of the evidence on each key question. If the team members disagreed about an evidence grade, the final grade given was based on the majority opinion. The grading scheme was derived from the scheme used in previous EPC projects.[28,29,31] For questions 1, 2, 3, and 4 the grades were as follows:

Grade A (strong): Appropriate data available, including at least one well done randomized controlled trial; study population sufficiently large; adequate controls; data consistent across studies; intervention clearly superior, equivalent or inferior to another strategy;

Grade B (moderate): Appropriate data available; study population sufficiently large; adequate controls; data reasonably consistent across studies; intervention likely to be superior, equivalent, or inferior to another but not enough evidence to conclude definitively;

Grade C (weak): Some data available; study population reasonably large; data indicate trend supporting benefit (or no benefit) of one intervention compared to another; not enough evidence to conclude that intervention is likely to be superior, equivalent or inferior to another;

Grade I (insufficient): Appropriate data not available or insufficient number of patients studied.

For questions 5, 6, 7, and 8 the evidence grades were as follows:

Grade A (strong): Appropriate data available, including at least one high quality study; study population sufficiently large; adequate reference standard; data consistent across studies; test definitely is or is not useful;

Grade B (moderate): Appropriate data available; study population sufficiently large;

adequate reference standard; data reasonably consistent across studies; data indicate test is likely to be or is likely not to be useful but not enough evidence to conclude definitively;

Grade C (weak): Some data available; study population reasonably large; data indicates trend supporting or not supporting usefulness of the test; not enough evidence to conclude that test is or is not likely to be useful;

Grade I (insufficient): Appropriate data not available or insufficient number of patients studied.

Peer Review Process

The EPC team sent a copy of the draft report to the core experts and the peer reviewers, as listed in Appendix A. The reviewers were asked to comment on the form and content of specific sections of the report, according to their areas of expertise and interest, and were invited to comment on other parts as well. The EPC team incorporated the reviewers' comments into the final report.

Figure 1: Causal pathway for diagnosis and treatment of venous thromboembolism as it relates to our key questions