Appendix: Detailed Methods

Overall approach

Web Figure 1 shows the analytic framework which was used to guide the review. For questions 1, the focus was on the systematic collection of individual family history information, and communication of family history-based personal risk of one or more of the conditions of interest, in populations considered representative of primary care populations. For question 1, appropriate health related outcomes identified from previous studies included patients' screening intention,(88) uptake of and adherence to screening tests and procedures(88-97) preventive health behavior,(12;93;98) and prophylactic preventive treatment and surgery.(82) For question 2, the outcomes of interest included perceived risk,(82;90;92;96;99;100) and perceived vulnerability and worry.(13-15;82;92;100) In order to examine the direct evidence, observational studies were included, and the review focused on published intervention studies only (randomized and non-randomized controlled trials, and uncontrolled before-after studies).

For question 3, noting the clinical decision-making context, the general approach was to work out which family history elements (see below) contributed most to discriminatory accuracy.(86) Discriminatory accuracy is the extent to which a test result (in this case positive or negative family history) accurately classifies or predicts actual disease occurrence in individuals. The principal metrics of interest were sensitivity and specificity. We considered the following to be categories of family history 'elements': specific ancestry; specific relative of interest (e.g. mother); number of affected relatives; degree of relationship of relative of interest to informant (e.g. first degree); age of onset of condition in affected relative(s); and lineage of affected relative(s) (i.e., informant’s maternal or paternal line). We treated any definition of 'positive family history' reported in an article as a combination of these elements. For each report, we noted the elements incorporated in every positive family history definition used. We then calculated the sensitivity and specificity of each definition for predicting/detecting the disease of interest, treating ‘fulfilling the specific definition’ as the ‘test’, and ‘not fulfilling the specific definition’ as the ‘reference’. For each condition, we compared the sensitivities and specificities for each definition, i.e. for each combination of one or more elements.

The initial strategy was to assess the independent contribution of each family history element to discriminatory accuracy for each condition, preferably by conducting a meta-regression analysis.

Data Sources

We searched the following bibliographic databases, for the years 1995 to second week of August 2008 inclusive: MEDLINE®, EMBASE®, CINAHL®, Cochrane Controlled Trials Register (CCTR)® and PsycINFO. We supplemented this by searching a limited number of grey literature sites, including National Coalition for Health Professionals Education in Genetics (NCPEG) and the Centers for Disease Control and Prevention (CDC). We supplemented this by searching the citation lists from eligible studies to identify studies that were not captured in our search. We restricted our search to studies published in English. The eligibility criteria varied by research question (below), but for all questions we excluded publications that were editorials, letters, conference papers, comments, opinions, or abstract only.

Questions 1 and 2:

We included studies with samples drawn from general populations (seen in any setting), participants in screening programs not based on family history, and patients seen in health care settings with primary care providers. We excluded studies where participants were seen by specialists (e.g., geneticists, cancer surgeons, oncologists, cardiologists) or where participants were recruited on the basis of genetic testing (irrespective of test result).

We included studies where the intervention was systematic collection or use of family history, alone or as part of multiple risk assessment. We excluded studies where actual collection of family history was not part of the intervention, and/or where family history was used as a participant selection criterion for the study.
We included primary studies with the following study designs: randomized controlled trials, non-randomized controlled trials, uncontrolled before-after studies. For controlled studies, we included studies where the comparator group received no intervention or received preventive advice without inclusion of family history-based information. We excluded observational cohort studies, case-control studies, case series and case reports.

Studies were included only if they reported at least one of the following outcomes: disease-specific mortality or morbidity (Q1); uptake of target health-related behavior (Q1); quality of life (Q2); family functioning (Q2); social functioning (Q2); psychological impact (Q2); risk perception (Q2).

Question 3:

We included studies with samples drawn from general populations (seen in any setting), participants in screening programs not based on family history, and patients seen in health care settings with primary care providers. We excluded studies where participants were seen by specialists (e.g., geneticists, cancer surgeons, oncologists, cardiologists) or where participants were recruited on the basis of genetic testing (irrespective of test result), or because of elevated risk of disease.

We included studies where the exposure was family history of the disease of interest (Full report online at www.ahrq.gov Table 1), collected in any manner. We excluded studies where family history information was not collected. We included prospective and retrospective cohort studies and cross-sectional analyses where family history of one of the conditions of interest was analyzed in relation to risk of disease. We excluded case control studies, case reports, and non-quantitative designs.

We included studies reporting occurrence, either as incidence (longitudinal studies) or prevalence (cross-sectional studies) of at least one of the disease conditions as defined in Full report online at www.ahrq.gov Table 1, and either clinically evident or ascertained by routinely available screening tests. We excluded studies reporting only pre-disease states as outcomes, or where non-clinically evident disease was ascertained through research-based technologies not available for general clinical use.

Question 4:

We included studies with samples drawn from general populations (seen in any setting), participants in screening programs not based on family history, patients seen in health care settings with primary care providers, and from specialist settings; we excluded where participants were recruited on the basis of genetic testing (irrespective of test result).

We included studies where the index test was family history, collected in any manner, and where the reference standard was verification of the disease status in relatives, ascertained according to any of the following methods: death registries; disease registries; medical records; direct contact with relatives; confirmation by relatives’ physicians. We excluded studies where family history was not examined, or where verification was not done according to one of the specified methods, or where patient report of family history was itself the reference standard.

We included studies with any quantitative design, comparative or non-comparative, and excluded case reports.

We included studies which reported one or more of the following: metrics of study accuracy (sensitivity, specificity, positive predictive value, negative predictive value); measures of completeness of family history collection; metrics of agreement (e.g. percent agreement, Kappa score). We excluded studies where true positives only were reported, with no additional information to calculate sensitivity or specificity; studies that evaluated test-retest reliability alone; and studies that did not report any outcomes listed above.

**Study Selection**
A team of study assistants was trained to apply the eligibility criteria for screening the title and abstract lists and the full text papers. All levels of screening were done in web-based Systematic Review Software (SRS) (TrialStat Corporation, Ottawa, Ontario Canada). Standardized forms and a training manual explaining the criteria were developed by the authors and reviewed with the screeners (see www.ahrq.gov). Two reviewers evaluated each citation for eligibility on the basis of title and abstract, and articles were retrieved if either one of the reviewers judged it eligible or there was insufficient information to determine eligibility. For screening of full text articles, two screeners came to consensus on the identification, selection, and abstraction of information. By discussion, authors resolved disagreements that remained and reviewed all reports where eligibility was unclear. Figure 1 shows flow of studies through initial identification to final inclusion.

Data Extraction and Quality Assessment

Pre-specified data were abstracted from all studies retained after full text screening, using data collection forms developed for use in SRS. One data extractor transferred the data onto these forms, and another checked the answers for accuracy before they were entered into SRS. All data entries were verified by at least one author.

Quality assessment was applied separately for each research question. (The details of the quality criteria used for all questions in this review can be found at www.ahrq.gov.) For questions 1 and 2, the Jadad scale(9) was used to evaluate internal validity of controlled trials. For reports using uncontrolled before and after design, no formal scale was available therefore critical appraisal was undertaken for the risk of selection and outcome biases. For question 3, the studies were grouped into longitudinal and cross-sectional analyses. Cohort studies where family history was assessed at the same time point as disease outcome were treated as cross-sectional for the purpose of this review. Finding no suitable existing scale, we developed a checklist of questions designed to assess likelihood of selection and information biases and applied this directly. For question 3, the Quality Assessment of Diagnostic Accuracy Assessment (QUADAS) was selected and all but four of the 14 criteria (items 3, 4, 12, 13) were not considered applicable to patient reporting of family history as an “index test”. We made the assumption that the index test (family history collection) and the reference test were equivalent across studies.

Data Synthesis and Analysis

A qualitative, descriptive approach was used to summarize study characteristics and outcomes for all research questions. Where applicable, multiple publications from the same study were grouped together and treated as a single study, with the most current data reported for presentation of summary results. Standardized summary tables explaining important study population and population characteristics and study results were created.

For question 3, we reviewed longitudinal and cross-sectional studies separately. We judged longitudinal studies as providing evidence on family history in predicting future disease in unaffected individuals, and cross-sectional studies as providing evidence on family history as a way of detecting current disease. Noting the absence of studies which permitted meta-regression analysis, we simplified the approach to a descriptive comparison of the different discrete, specific definitions of ‘positive family history’ (see ‘Overall approach’ above). For each study, we extracted the actual numbers of true and false positive and true and false negative results (TP, FP, TN, and FN), or estimated these numbers on the basis of reported proportions, for each family history definition used. We calculated sensitivities, and specificities with the accompanying 95 percent confidence intervals (CI). We also extracted data on the crude disease frequency in the sample and the proportion of participants meeting each definition of positive family history.

Meta-analysis was not appropriate for any of the research questions: for questions 1 and 2, there were an insufficient number of studies; for question 3, there was significant clinical heterogeneity across studies, and many observations were compared within studies meaning that they could not be considered independent; for question 4, there was significant clinical heterogeneity across studies, too few studies for some disease categories, and/or insufficient data (no measures of variance).