1.a. Full Title: The Effect of Using Framingham Risk Calculators with Different Predicted Outcomes

b. Abbreviated Title (Length 26 characters): CHD Risk Outcomes

2. Writing Group:
Writing group members: Stacey L. Sheridan, MD, MPH, Gerardo Heiss, MD, PhD, Ross Simpson, MD, PhD, Couper, David, PhD

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal. ___SLS__ [please confirm with your initials electronically or in writing]

First author: Stacey L. Sheridan, MD, MPH
Address: Division of General Medicine and Clinical Epidemiology, 5039 Old Clinic Building, CB 7110 Chapel Hill, NC

Phone: 919-966-2276 Fax: 919-966-2274 E-mail: sls593@med.unc.edu

Corresponding/senior author (if different from first author correspondence will be sent to both the first author & the corresponding author):
Address:

Phone: Fax:
E-mail:

3. Timeline: December 2005 through December 2006

4. Rationale:
Available tools to calculate global CHD risk employ different predicted outcomes. Most tools provide information on the combined risk of stable and unstable angina, myocardial infarction, and CHD death (total CHD events). Some tools, however, report only the risk of myocardial infarction and CHD death (hard events); these tools produce smaller numeric estimates of risk than tools that also include angina. Tool developers have provided little justification for their choice of outcomes and current risk-based prevention guidelines are similarly silent on these issues in their support of various tools. Thus, the most appropriate tool for CHD risk calculation is unknown to the average clinician.

Using different predicted outcomes may have important consequences in decision-making for CHD prevention. For instance, clinicians who use a risk calculator that predicts total events might be expected to recommend aspirin more often than clinicians who use a calculator that predicts hard events, because the CHD event rate (and the absolute benefit of treatment) appears to be greater. Similarly, patients who receive information about their risk of total events may be more likely to accept treatment, particularly if specific information on outcome is not provided. These differences in recommendations and acceptance of treatment are important to the extent that they alter the rates of beneficial and harmful consequences of CHD prevention.

To avoid unintended consequences, predicted outcomes of tools would ideally be aligned with evidence regarding treatment benefit. For instance, a clinician would ideally use a tool that predicts the risk of hard events when deciding about aspirin because primary prevention studies of aspirin tested only hard outcomes. The problem is that not all primary CHD prevention studies used the same predicted outcomes (1-3) and relevant outcomes may differ among subgroups (i.e. angina is a more prevalent presentation in women). Thus, centralizing decision-making around the concept of global risk will likely require compromises in choosing the most appropriate predicted outcome.

The optimal outcome to drive decision-making about CHD prevention is unclear. Thus, we undertake this study to help clarify which outcomes should be used in the process of global CHD risk calculation.

5. Main Hypothesis/Study Questions:

Our overarching hypothesis is that calculators that predict total CHD events (e.g. MI, CHD death, and stable and unstable angina) result in the decision to treat significantly more individuals with preventive therapy than calculators that predict hard CHD events (e.g. MI and CHD death). Whether this excess treatment is justified depends on 1) whether calculators predict total events better than hard events in important population subgroups and 2) whether primary CHD prevention studies have shown reductions in the outcomes predicted.

To determine the optimal outcome to be used for CHD risk calculation for primary prevention, we propose the following aims:

1) Perform a secondary data analysis of the Atherosclerosis Risk in Communities (ARIC) database to:
   a. Determine the population impact of using risk calculation tools with different predicted CHD outcomes

Figure 1: Predicted Outcomes in Available Framingham Risk Calculators

Calculators that predict total† CHD events:
Framingham Risk Tables from Wilson, 1998
Modified Sheffield Tables
Canadian Nomogram
Stat Cardiac Risk
Fram Plus
Heart-to-Heart (H2H)

Calculators that predict hard* events:
Joint British Societies Risk Prediction Charts
Joint European Societies Coronary Risk Chart
National Cholesterol Education Program (NCEP)
American Heart Association

† total events = stable and unstable angina, MI, and death
* hard events = MI and death
b. Determine whether risk calculation tools with different predictive outcomes have different predictive abilities among important population subgroups

2) Systematically review the medical literature to determine what outcomes have been used in primary prevention studies.

6. Data (variables, time window, source, inclusions/exclusions):

**Study Population:** For our analysis, we plan to use data from patients previously enrolled in the ARIC study. (4) ARIC is a NHLBI-funded prospective epidemiologic study conducted in four diverse U.S. communities (Forsyth County NC; Jackson, MI; Minneapolis, MN; Washington Co, MD). It is designed to investigate the etiology and natural history of atherosclerosis, the etiology of clinical atherosclerotic diseases, and variation in cardiovascular risk factors, medical care and disease by race, gender, location, and date. ARIC includes two parts—the Cohort Component and the Community Surveillance Component. For our analysis, we plan to focus on the cohort component, which includes 15,792 individuals aged 45-64 who were randomly selected and recruited from the four ARIC communities.

**Sample Selection:** We will include all individuals in the ARIC database in our sample unless they have a prior history of cardiovascular disease using standard ARIC definitions.

**Measurements:** In the sections below, we detail our planned measurements.

All measurements are taken or derived from data in the ARIC database and use current ARIC definitions (4) unless otherwise indicated below. ARIC cohort participants received extensive interviews and assessments, including full assessment of CHD risk factors, primary prevention treatments, and coronary outcomes. Participants were examined every three years with the first screen (baseline) occurring in 1987-89, the second in 1990-92, the third in 1993-95, and the fourth and last exam in 1996-98. Follow-up occurred yearly by telephone to maintain contact with participants and to assess the health status of the cohort.

**Clinical Data for CHD Risk Calculation:** CHD risk calculation requires input about several clinical variables. We will use the following ARIC measurements for our assessment of CHD risk: age, gender, smoking history (yes/no), fasting glucose (with glucose >126 on fasting check considered to represent diabetes), systolic or diastolic blood pressure levels, and concentrations of total and HDL-cholesterol.

**Global CHD Risk:** We will calculate CHD risk using two different CHD risk calculators: the National Cholesterol Education Program calculator (NCEP; [www.nhlbi.nih.gov](http://www.nhlbi.nih.gov)) and the Heart-to-Heart calculator (H2H; [www.med-decisions.com](http://www.med-decisions.com)). As described in our original proposal, Framingham risk calculators calculate CHD risk based on prospectively collected data from men and women aged 30-74, are well-validated, and provide an absolute measure of the risk of coronary disease over a defined period of time.

The two calculators we plan to use for this study (e.g. NCEP and H2H) are unique in two aspects. First, they require input about the same clinical variables (age, gender, smoking history, presence or absence of diabetes—with NCEP forcing a “no” answer; systolic or diastolic blood pressure levels, and concentrations of total and HDL-cholesterol). Second, they report the predicted outcomes we wish to compare: MI and death (hard outcomes; NCEP calculator) and stable and unstable angina, MI, and death (total outcomes: H2H calculator). We will use identical demographic and risk factor information from the baseline screen of the ARIC cohort to calculate global CHD risk using these calculators.

**The Potential Impact of Calculators with Different Predicted Outcomes:**
We will estimate the potential impact of calculators with different predicted outcomes on CHD decision-making in two ways: 1) by comparing the proportion of participants assigned to clinically meaningful risk categories when using different risk calculators, and 2) by estimating the proportion of patients eligible for preventive therapy when using different risk calculators.
Proportion of participants in various risk categories
To estimate differences in the proportion of participants assigned to clinically meaningful risk categories when using different risk calculators, we will first categorize participant risk scores into low (0-5%), moderate (6-9%), high (10-19%), and very high (>=20%) risk categories. We consider these categories as clinically meaningful based on recommendations in recent consensus and evidence-based guidelines that stratify the aggressiveness of LDL treatment by <10%, 10-20%, and >20% risk categories (5) and stratify aspirin recommendations by 0-5%, 6-9%, >=10% categories. (3;6)

Proportion of participants eligible for preventive therapy
To estimate differences in the proportion of participants who are eligible for preventive therapy when using different risk calculators, we will categorize participants as eligible or ineligible for preventive therapy using the primary prevention guidelines shown in table 2. Because risk-based recommendations are not currently available for hypertension treatment or smoking cessation, our focus will be on participant eligibility for cholesterol treatment and aspirin.

Table 2. Eligibility for Preventive Therapy by Current Primary Prevention Guidelines

<table>
<thead>
<tr>
<th>Primary Prevention Guideline</th>
<th>CHD Risk Level Over 10 years</th>
<th>Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cholesterol Guidelines (NCEP)</td>
<td>CHD risk &gt;20% or diabetes</td>
<td>Treat if LDL &gt;100</td>
</tr>
<tr>
<td></td>
<td>CHD risk 10-20%</td>
<td>Treat if LDL &gt;130</td>
</tr>
<tr>
<td></td>
<td>CHD risk &lt;10%</td>
<td>Treat if LDL &gt;160</td>
</tr>
<tr>
<td>Aspirin Guidelines (AHA)</td>
<td>CHD risk &gt;10%</td>
<td>Treat with aspirin</td>
</tr>
</tbody>
</table>

The Predictive Ability of Calculators with Different Outcomes in Important Subgroups:
We will examine the predictive ability of calculators with different outcomes in important subgroups in two ways.

First, we will examine the discriminative ability of each calculator using Receiver Operator (ROC) curves. ROC curves are a graphical method of depicting the trade-offs between the true-positive predictions and false positive predictions of risk calculators and are constructed by plotting the sensitivity versus 1-specificity of calculations for CHD outcomes at various thresholds of CHD risk. The area under the ROC curve represents the ability of a model to correctly predict who will have a CHD event and who will not. We will quantify the area under the ROC curve for each of our risk calculators using the c-statistic.

Second, we will examine the calibration of the model by comparing predicted risk from the calculators with the actual outcomes incurred over 10 year follow-up of the ARIC cohort. This degree to which the predicted risk approximates the actual outcomes will be measured using the Hosmer-Lemeshow chi-square statistic.

Actual CHD Outcomes
The ARIC study records the occurrence of a number of CHD outcomes, including 1) hospitalized MI and stroke, 2) death from CHD, stroke and all-causes, and 3) a number of non-hospitalized non-fatal events, such as angina pectoris. Cohort events are identified through review of hospital discharge indexes, death certificates, and information elicited during the annual follow-up interviews. Both non-fatal and non-hospitalized fatal events are then investigated by review of hospital records and death certificates using well-defined protocols. Events are classified by computerized algorithm and then reviewed by the Morbidity and Mortality Classification Committee for final classification.

For our study, we will focus on the combined outcomes of 1) MI and CHD death, and 2) MI, CHD death, and angina measured at the fourth cohort exam (1996-1998), a 10-year interval from baseline risk assessment. We chose these outcomes because they are the outcomes predicted by the Framingham calculators we are studying. The hard outcomes are different than the current
incident CHD events defined by ARIC, however, which also include procedural markers of CHD such as angioplasty and coronary artery bypass. We will use sensitivity analysis to determine how well the NCEP Framingham calculator predicts ARIC-defined incident events.

In analyses including the outcome of MI, we will include both recognized and unrecognized MI. For recognized MI, we will include all hospitalized events classified by ARIC as “definite MI”, “probable MI”, and “suspected MI.” We will perform sensitivity analyses to determine the effect of excluding probable and suspected MI on study outcomes. We will additionally perform sensitivity analyses to estimate the impact of changes in enzyme diagnosis techniques (i.e. use of troponin vs. CK-MB) on study results. Unrecognized MI will be determined by ARIC follow-up examination EKG meeting one of the following criteria: 1) a major Q-wave, 2) a minor Q-wave with ischemic ST-T changes, or 3) an MI by computerized NOVA-CODE criteria confirmed by side-by-side visual comparison of baseline and follow-up EKGs.

In analyses including the outcome of CHD death, we will include all events classified by ARIC as “definite fatal MI”, “definite fatal CHD”, and “possible fatal CHD.” We will perform sensitivity analyses to determine the effect of excluding possible fatal CHD on study results.

In analyses including the outcome of stable angina, we will include events classified in three ways. First, we will include events classified as definite angina by the Rose questionnaire (e.g. brought on by exertion, situated in central or left anterior chest, forced the subject to slow down, and was relieved by rest within 10 minutes). Rose angina has a sensitivity of 26-44% and a specificity of 55-79% compared to a gold standard of a exercise thallium testing (7; 8) and 44% and 86% compared to a gold standard of clinician diagnosis (7) such as that employed in the Framingham study. It additionally is a comparable predictor of cardiovascular outcomes when compared to physician diagnosis (9) and predicts subclinical disease. (10) Improved sensitivity for the diagnosis of angina can be achieved with the Rose questionnaire by considering nitroglycerin responsive chest pain (sensitivity 60%/specificity 63%)(7), therefore, we will secondarily perform analyses incorporating nitroglycerin responsiveness. This may mitigate against gender differences in the test characteristics of the Rose Questionnaire that stem from atypical anginal presentations in women. Finally, we will perform analyses using patient-reported physician diagnosis of angina. This measure has been reported to have a sensitivity of 80% when compared with medical record review. (11)

For analyses including the outcome of unstable angina, we will include chest pain events that subjects report are occurring in an escalating pattern (e.g. twice as often as before, more severe than before, longer than before, requiring more nitroglycerin than before for relief, or occurring with less exertion or even at rest).

Subgroups
Subgroups of interest in this analysis are those in which rates of studied outcomes are expected to vary: these subgroups include different absolute risk groups (low versus moderate versus high risk), different racial groups (white versus African-Americans), different gender groups (women versus men), and different community sites (Forsyth County NC; Jackson, MI; Minneapolis, MN; Washington Co, MD).

Statistical Methods: The primary purpose of this study is to determine optimal risk calculator to use in our planned CHD intervention. Our primary outcome will be the proportion of participants eligible for aspirin and cholesterol therapy when calculating CHD risk using different risk calculators. We will secondarily assess differences in the proportion of participants assigned to clinically meaningful risk categories when using different risk calculators and determine whether risk calculation tools with different predictive outcomes have different predictive abilities among important population subgroups. Specifics of our analyses are described below.

- We will use chi square tests to determine differences in the proportion of participants who are eligible to receive preventive therapy when using different risk calculators. We will consider a 10% difference in eligibility for preventive therapy significant, accepting a two-sided alpha of 0.01.
• We will use chi square tests to determine significant differences in the proportion of participants assigned to various CHD risk categories when using different risk calculators. Again, we will consider a 10% difference in assignment significant, accepting a two-sided alpha of 0.01.

• We will compare receiver operating characteristic (ROC) curves and corresponding c-statistics to determine whether risk calculation tools with different predictive outcomes have different predictive abilities in important subgroups. Because we will be working with censored data, we will use Kaplan-Meier-like methods to calculate the relevant probabilities of events by 10 years in the ARIC population. We will then compare these probabilities with those estimated by the different risk calculators, calculating sensitivity and 1-specificity values across a range of risk cut-offs and creating an ROC curve. We will use C-statistics, which represent the area under the ROC curve, as a measure of discriminating ability of the prediction. If the risk calculation is unrelated to the outcome of interest the c-statistic would be 0.5; if the risk calculation is perfectly related to the outcome of interest, the c-statistic would be 1. We will consider a 10% difference in the c-statistics for calculations among important subgroups clinically meaningful difference. We will additionally use a bootstrap method to estimate differences in ROC curves because c-statistics are not directly applicable to curves derived from censored data.

• We will compare the number of predicted CHD events with the number of actual CHD events (calculated through Kaplan-Meier-like methods) using Hosmer-Lemeshow chi square tests. We will consider small values (<20) to indicate good calibration (p<0.01).

Because the risk calculators we’ll use in this study handle patients with diabetes differently (e.g. the NCEP calculator forces use of heuristics (treat all) whereas H2H allows risk-based decision for all), we will stratify all analyses by diabetic status.

**Aim 2: Systematic Review of Primary Prevention Outcomes**

To determine which predicted outcomes are best aligned with the outcomes studied in primary CHD prevention trials, we will perform a systematic review of the medical literature.

**Search Strategy**

We will use U.S. Preventive Services Task Force Guidelines and searches of the COCHRANE Collaboration Library to identify the most recent systematic evidence reviews on primary CHD prevention therapy. We will use identified reviews to summarize the CHD outcomes used in primary prevention studies and search MEDLINE for any more recent articles examining the benefits (compared to no intervention or placebo) of treating hypertension, treating abnormal cholesterol, stopping smoking, or taking aspirin to prevent primary CHD events. To update information from recent systematic evidence reviews, we will identify English-language articles published between the search date of the most recent systematic review and the current date, and use manual searches of relevant articles, personal libraries, and peer review to ensure that we included all appropriate articles.

**Inclusion Criteria for Admissible Evidence**

We will include recent systematic reviews and subsequently dated individual observational studies or randomized controlled trials (RCTs) that examine the following topics in primary prevention populations: treatment of hypertension, treatment of cholesterol, smoking cessation interventions, and aspirin use. Outcomes of interest will include the composite endpoints of MI and death (hard outcomes) or stable and unstable angina, MI, and death (total outcomes) to be consistent with outcomes predicted by various risk calculation tools.
Data Extraction and Synthesis
Two authors (SS and her primary mentor, RS) will review abstracts of potentially relevant articles to
determine if they should be included. When the reviewers disagree, they will obtain the full articles and
resolve the disagreements by consensus.

From the most recent fair to good quality systematic review of each CHD prevention topic, a single
reviewer will extract data only about the CHD outcomes used in primary prevention studies. To update
information from these systematic reviews, the same reviewer will extract outcome data from more recent
studies. To allow quality assessment of newer studies, the reviewer will also create evidence tables
detailing the study design, intervention and control groups, relevant outcome measurements, and results of
each newer study and use guidelines developed for the USPSTF reviews to evaluate their internal and
external validity. Experts will review all work for accuracy and completeness.

Data synthesis will focus on identifying the primary outcomes used in fair and good quality studies for each
primary preventive therapy. These outcomes will then be considered in the context of data from our
secondary data analysis to determine the optimal calculator for use in CHD prevention.

7.a. Will the data be used for non-CVD analysis in this manuscript? ____ Yes
     ____ No

   b. If Yes, is the author aware that the file ICTDER02 must be used to exclude
      persons with a value RES_OTH = “CVD Research” for non-DNA analysis, and
      for DNA analysis RES_DNA = “CVD Research” would be used? ____ Yes  ____ No
      (This file ICTDER02 has been distributed to ARIC PIs, and contains
      the responses to consent updates related to stored sample use for research.)

8.a. Will the DNA data be used in this manuscript? ____ Yes
     ____ No

8.b. If yes, is the author aware that either DNA data distributed by the
     Coordinating Center must be used, or the file ICTDER02 must be used to
     exclude those with value RES_DNA = “No use/storage DNA”?  ____ Yes  ____ No

9. The lead author of this manuscript proposal has reviewed the list of existing ARIC
Study manuscript proposals and has found no overlap between this proposal and
previously approved manuscript proposals either published or still in active status.
ARIC Investigators have access to the publications lists under the Study Members Area
of the web site at: http://www.csec.unc.edu/ARIC/search.php

     ____ X__ Yes     _______ No
10. What are the most related manuscript proposals in ARIC (authors are encouraged to contact lead authors of these proposals for comments on the new proposal or collaboration)?


11. a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data?  
   ____ Yes   ___X__ No

11.b. If yes, is the proposal
   ____ A. primarily the result of an ancillary study (list number* _________)
   ____ B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)* _________ _________ _________)

*ancillary studies are listed by number at http://www.cscc.unc.edu/aric/forms/

12. Manuscript preparation is expected to be completed in one to three years. If a manuscript is not submitted for ARIC review at the end of the 3-years from the date of the approval, the manuscript proposal will expire.
Reference List


(4) ARIC investigators. Atherosclerosis Risk in Communities Study; Manual 1: General Description and Study Management. 2005.


