

Appendix 1 for

*Simulating and Evaluating Local Interventions
to Improve Cardiovascular Health*

Jack Homer, Bobby Milstein, Kristina Wile,
Justin Trogdon, Philip Huang, Darwin Labarthe, Diane Orenstein

Submitted October 2008 to *Preventing Chronic Disease*

Risk Calculation in the Cardiovascular Health System Dynamics Simulation Model (CVH-SD)

Prepared by
Jack Homer, PhD

In the CVH-SD model we have sought to link the individual risk factors of high blood pressure (“B”), high cholesterol (“C”), diabetes (“D”), and smoking (“S”) (hereafter referred to jointly as “BCDS”) to event rates for coronary heart disease (CHD), stroke, and total cardiovascular disease (CVD). As described in Finkelstein et al. 2006 (1), a number of risk calculators have been developed for making such a linkage, the great majority of them being based on the (largely Caucasian) Framingham, Massachusetts population for their data sample. The Framingham-based risk calculators have been found to perform reasonably well when applied to other racial and ethnic subpopulations and some other countries, including U.S. blacks and Hispanics (2, 3) as well as Australia and New Zealand (4).

The risk calculators provide event probabilities at the level of the individual, over some specified period of time (e.g., 5 years or 10 years), based on various versions of the BCDS risk factors (blood pressure may be based on systolic BP (SBP) and/or diastolic BP (DBP); cholesterol may be based on total cholesterol and/or high density lipoprotein levels) as well as sex and age. Some of these calculators treat blood pressure and cholesterol only categorically based on accepted cutpoints to distinguish “high” from “normal”, while others allow entry of the absolute values. None of the risk calculators reviewed by Finkelstein includes obesity as a separate risk factor. Only one (National Cholesterol Education Program/Adult Treatment Protocol III, NCEP/ATP) considers whether the person is on medication (BP medication, to be specific). Another one (5) also includes ECG-detected left ventricular hypertrophy (ECG-LVH) in the risk equation.

After reviewing the advantages and disadvantages of the various risk calculators, we have settled on the one described in Anderson et.al. 1991 (5). This calculator (which comes in two versions) is based on Framingham data N=5573, men and women ages 30-74 at baseline 1968-1975, who had no CVD or cancer at baseline, and were followed for 12 years. The inputs include sex, log(age), log(SBP) in version 1 and log(DBP) in version 2, log(TC divided by HDL), smoking (yes-no), diabetes (yes-no), and ECG-LVH (yes-no).

The outcome events include CHD (=MI+AP+CHD death + coronary insufficiency), stroke (including TIA), CVD (=CHD+stroke+CHF+PAD), and deaths from CHD and CVD. The equations employ a “non-proportional hazards Weibull accelerated failure time model.” They include as inputs the variables described above as well as some interaction terms, such as “log(age) x female” and “diabetes x female”.

Here’s how the Anderson risk calculator works. First, for a given outcome type (e.g., CHD event), one uses the input variables (x_i) for an individual and the corresponding beta coefficients for that event type (β_i , provided in the article) to compute:

$$\text{Mu} = \text{SUM of } [(\beta_i)(x_i)]$$

Also provided in the article are two theta coefficients per event type, used to compute:

$$\text{Sigma} = \text{EXP}(\text{Theta}_0 + (\text{Theta}_1)(\text{Mu}))$$

Now one may compute using the following Weibull equation the probability that the event in question will occur sometime within a specified time t (in years):

$$P(t) = 1 - \text{EXP}(-\text{EXP}[(\log(t) - \text{Mu}) / \text{Sigma}])$$

This equation tends toward 0 as t approaches 0 and tends toward 1 as t approaches infinity, as it should. The authors say that “Time intervals of 4 to 12 years are recommended.” I suppose the 12 year upper limit is because the data follow-up time was only 12 years, while the 4 year lower limit may indicate that the model does not provide accurate and reliable estimates of event probabilities for values of t less than 4.

Let’s take an example to see how this works. Consider the 4-year risk of a CHD event for a 50-year old female with no prior CVD or cancer, and no LVH. The calculator gives the following:

No RFs	0.48%	
B only	0.98%	(delta 0.50% vs no RFs)
C only	1.01%	(delta 0.53% vs no RFs)
D only	1.36%	(delta 0.88% vs no RFs)
S only	1.05%	(delta 0.57% vs no RFs)
BCDS	7.45%	(all 4 risk factors: delta 6.97% vs no RFs)

The calculations show that there is a strong synergy among the risk factors: The increase in risk from having all 4 risk factors is much greater than one would expect from simply summing the increases for each risk factor by itself. Separate comparisons of risk for BC and BCD also show that the amount of synergy increases with each additional risk factor. That is, there is some synergistic effect for having 2 risk factors, even more for having 3 risk factors, and much more synergy for having 4 risk factors. As noted by Magnus and Beaglehole (6): “...these factors do not act as separate...risks that can be simply added. Rather, the factors interact in various combinations to produce a total risk for CHD in a multiplicative way.”

Our current CVH-SD model does represent such multiplicative interaction, but it does not do so by portraying individuals with different combinations of risk factors. Rather, being a compartmental model, it deals only in terms of population prevalences of each risk factor. The model calculates each prevalence (B, C, D, S—imagine each of these as a circle in an intersecting Venn diagram) separately, but does not attempt to represent or calculate the prevalence of each of the 16 mutually exclusive combinations (e.g., B-but-not-CDS, BCS-but-not-D, etc.—imagine each of these as one of the 16 discrete areas in the Venn diagram.) We examined NHANES data on these various combinations and found that in some cases the extent of overlap was somewhat different from what one would expect based on a null hypothesis of pure independence. But, we ultimately decided *not* to pursue this route of modeling all of the 16 risk segments (multiplied by 6 sex-age categories), in large part because of concerns over small sample sizes in the NHANES database when it is carved into many small pieces.

We also decided not to portray all 16 risk segments because of inherent limitations in the Anderson risk calculator which suggest that it should not be taken too literally for population modeling of CVD risk. First, the Anderson calculator does not distinguish between diabetes that is controlled versus not controlled. Yet the literature shows clearly that control of elevated blood sugar, even in a diabetic, can reduce the risk for CVD, just as the control of elevated blood pressure or blood lipids can (7, 8). Second, the Anderson calculator does not include risks due to secondhand smoke (SHS) and particulate (PM 2.5) air pollution, whereas our model does. The literature indicates that these factors can contribute directly and significantly to CHD risk (9, 10)

For the reasons listed above, it was clear we would be using the Anderson calculator selectively for guidance in the CVH-SD model, rather than in a literal way. We have done this as follows:

1. Use the Anderson calculator to estimate risks of CHD event, Stroke event, CVD event, and CVD death for each of the six sex-age groups, using age 20 as the prototype for Ages 18 to 29; using age 50 as the prototype for Ages 30 to 64; and using age 75 as the prototype for Ages 65 plus.
2. In particular, calculate the risks for having (a) no risk factors, (b) each one of the 4 risk factors by itself, and (c) all 4 risk factors together. For these latter two, calculate relative risks [RR] relative to having no risk factors.
3. Calculate an exponent that describes the extent to which the RR for having all 4 risk factors differs from the multiplicative product of the four RR's for the individual risk factors. For events other than stroke, we have found that these exponents lie in the range of 0.7 to 0.9, indicating that the synergy among risk factors is somewhat less than what one would get from pure multiplication. The exponents for stroke, interestingly, are all about 1.0, indicating that the synergy among risk factors is very close to what one would get from pure multiplication.
4. For each of the 4 risk factors, calculate a multiplier based on prevalence as follows:

$$\text{Multiplier}[i] = (1 - \text{Prevalence}[i]) + \text{Prevalence}[i] * \text{RR}[i; \text{event type, sex, age}]$$

5. Calculate a combined risk multiplier by multiplying together the four individual risk multipliers, then raise that product to the exponent described above. Then multiply the adjusted combined risk multiplier by the Anderson risk for having no risk factors (#2a above) and by the Non-CVE population in 2003. Let's call this "Anderson-based 4-year event rate". Do this for each sex-age group.
6. Also calculate risk multipliers from SHS and particulate air pollution. For SHS these risk multipliers are based on relative risks estimated from the literature (9) combined with model calculations of the fraction of the non-CVD population exposed to significant SHS (by sex and age group). For particulate air pollution the risk multipliers are based on relative risks estimated from the literature (10), applied equally to all sex and age groups.
7. Multiply the "Anderson-based 4-year event rate" from #5 above with the SHS and air pollution risk multipliers in #6 to get the "Adjusted Anderson-based 4-year event rate". For each sex, sum these across the age groups, and compare this result to the data from AHA on the number of first-time events (CHD, stroke, total CVD, CVD death) by sex in 2003. Call this ratio the "event rate multiplier": it is the factor to be used for translating from the "Adjusted Anderson-based 4-year event rate" to a measured 1-year event rate based on AHA data. In this way we have found the following event multipliers:

	Male	Female
CHD event	.30	.46
Stroke event	.76	.79
CVD event	.44	.59
CVD fatality*	1.32	2.44

The CVD fatality multiplier compares AHA 1-year fatality rate with the Anderson-based 1-year fatality rate. The Anderson rate is calculated as the ratio of the risk of CVD death to the risk of a CVD event.

8. The estimated number of events (for each of the four types listed above) by sex and age may now be calculated as the Adjusted Anderson-based 4-year event rate (by sex and age group) multiplied by the corresponding event multiplier (by sex).

References

1. Finkelstein EA, Khavjou OA, Will JC, Farris RP, Prabhu M. Assessing the ability of cardiovascular risk calculators to evaluate effectiveness of trials and interventions. *Expert Reviews in Pharmacoeconomics Outcomes Research* 2006;6(4):417-24.
2. D'Agostino RB, Sr., Grundy S, Sullivan LM, Wilson P. Validation of the Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. *Journal Am Med Assoc* 2001;286(2):180-7.
3. Liao Y, McGee DL, Cooper RS, Sutkowski MB. How generalizable are coronary risk prediction models? Comparison of Framingham and two national cohorts. *Am Heart Journal* 1999;137(5):837-45.

4. Eichler K, Puhan MA, Steurer J, Bachmann LM. Prediction of first coronary events with the Framingham score: a systematic review. *Am Heart Journal* 2007;153(5):722-31, 31 e1-8.
5. Anderson KM, Odell PM, Wilson PW, Kannel WB. Cardiovascular disease risk profiles. *Am Heart Journal* 1991;121(1 Pt 2):293-8.
6. Magnus P, Beaglehole R. The real contribution of the major risk factors to the coronary epidemics: time to end the "only-50%" myth. *Arch. Intern. Med.* 2001;161(22):2657-60.
7. Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development and progression of long term complications in insulin dependent diabetes mellitus. *N Engl Journal Med* 1993; 329:977-986.
8. Prospective Diabetes Study Group. Tight blood pressure control and risk of macrovascular and microvascular complications in type 2 diabetes. *Lancet* 1998; 352:703-713.
9. United States. Public Health Service. Office of the Surgeon General. The health consequences of involuntary exposure to tobacco smoke: a report of the Surgeon General. Rockville, MD: U.S. Dept. of Health and Human Services, Public Health Service, Office of the Surgeon General; 2006.
10. Pope CA, 3rd, Burnett RT, Thurston GD, Thun MJ, Calle EE, Krewski D, et al. Cardiovascular mortality and long-term exposure to particulate air pollution: epidemiological evidence of general pathophysiological pathways of disease. *Circulation* 2004;109(1):71-7.