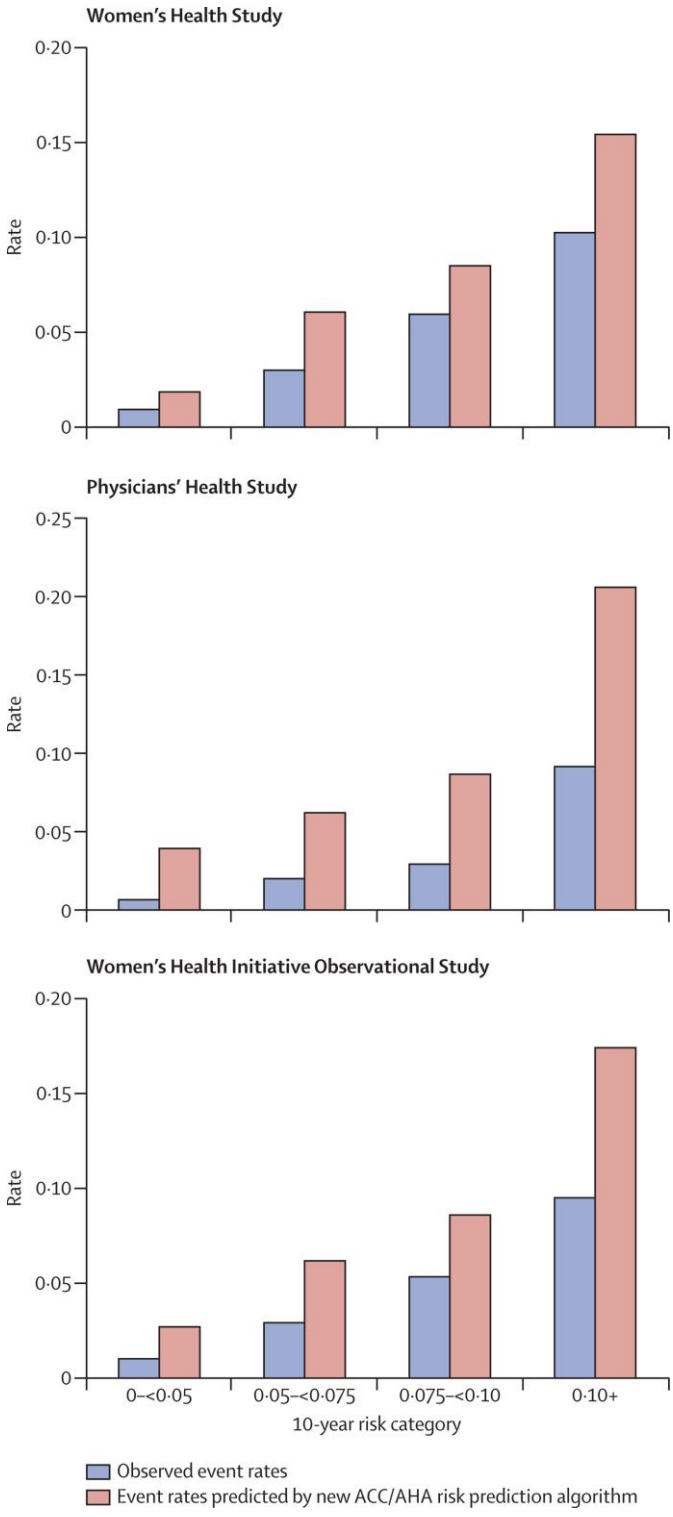
6.4.A Pooled Cohort Equations for estimating risk of cardiovascular events

For many preventive interventions, the balance of benefits and harms depends on the absolute risk of the event(s) to be prevented. Thus, guidelines for statin and aspirin treatment to prevent cardiovascular disease are based on the 10-year risk of heart disease or stroke, estimated using an online calculator (available at http://www.cvriskcalculator.com/).

However, Ridker and Cook(Ridker and Cook 2013, Cook and Ridker 2016, Ridker and Cook 2016) have found that the risk estimated from the pooled cohort equations is substantially higher than that observed in more recent cohorts. (Three examples are shown in Figure 1, from (Ridker and Cook 2013)).



A. Is this a problem with discrimination or calibration? Explain.

Calibration. Poor calibration means that the probability estimates are off – too high or too low. Poor discrimination would mean that predicted event rates in those who died were not much higher than in those who survived.

Although the figure is not a typical calibration plot, it contains the same information: a comparison of observed and expected mortality rates in different risk groups.

Figure: Comparison of observed event rates with event rates predicted by new ACC/AHA risk prediction algorithm in three external validation primary prevention cohorts: the Women's Health Study, the Physicians' Health Study, and the Women's Health Initiative Observational Study. Reprinted with permission. Reprinted from the Lancet, Vol. 382 Paul Ridker And Nancy Cook: Statins: new American guidelines for prevention of cardiovascular disease. Pages 1762-5, Copyright 2013 with permission from Elsevier.

B. The guidelines recommend estimating each subject's risk using a calcualtor, then managing based on whether the predicted 10-year risk is <5%, 5-7.4%, 7.5-9.9%, or ≥ 10%. Based on the description above, do the risk groupings in the figure represent quartiles of risk?

No. We don't know what proportion of the population would be classified as having a 10-year risk of <5%, 5 to 7.4%, etc., but there is no reason why each category would include 25% of the population, which is what quartiles of risk would require.

C. Explain briefly, step by step, how the numbers needed produce figures like the bar graphs above would be obtained.

1. Find an existing cohort (or assemble a new one) cohort to obtain the data
2. Use the values of the subjects at baseline with the risk calculator to predict the 10-year risk in each subject.
3. Group the subjects by predicted risk into 4 groups: e.g., 0-5%; 5-7.5%;7.5%-10%; >10% as was done here. This will be the X-coordinate.
4. Use the follow-up of the cohort you assembled to obtain the observed proportions with events in each risk group.
5. Compare the mean *predicted* risk in each group versus the observed proportion with events over 10 years. This could be plotted as the authors did or with a more traditional calibration plot.

You can see that if you already have a cohort study with values of baseline variables and follow-up, this would be easy to do.

D. In which cohort was the calculator most poorly calibrated? Explain your answer including any assumptions you had to make given your answer to (b) above.

The Physician's Health Study shows observed event rates farther below the predicted rates than the other two cohorts for all 4 points, so it is probably the worst calibrated. However, to know for sure, we would need to assume roughly similar distribution of the cohorts between the 4 risk groups. If the Physician's Health Study had a much higher proportion of patients in the low risk groups, it could be better calibrated if the metric for evaluating calibration was the average absolute difference between observed and predicted proportions.

e) As already mentioned, treatment recommendations are based on a patient's risk group as determined by the calculator. If we assume that, in fact, the risk calculator is overestimating risk, what more do we need to know about the recommended treatment thresholds to conclude that these overestimated risks will lead to excessive treatment? Explain.

We need to know whether the treatment thresholds in the guideline are too high. If they are too high, then the overestimated risks might actually lead to optimal treatment, because more people who would benefit from treatment would receive it. In this case, the error in calibration could cancel out the error in threshold determination.

f. Ridker and Cook(Ridker and Cook 2016) have pointed out that American Heart Association/American College of Cardiology (AHA/ACC) risk calculator was based on pooled cohort equations derived from cohorts that enrolled subjects from 1968 to 1990, whereas the contemporary external validation cohorts in which risk was found to be overestimated enrolled subjects 20-30 years later. During that time, death rates from cardiovascular disease (CVD) and coronary heart disease (CHD) were declining (figure).

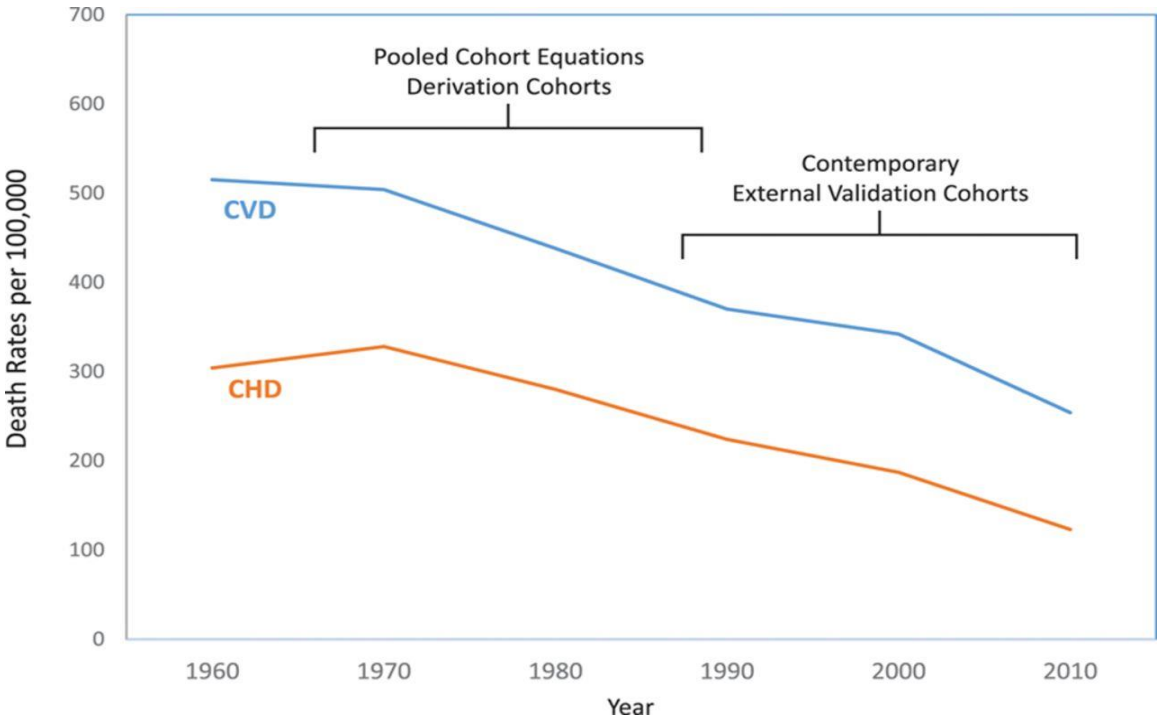


Figure. US death rates per 100 000 from cardiovascular disease (CVD) and coronary heart disease (CHD). From Ridker and Cook(Ridker and Cook 2016). (Open access article; figure reprinted with permission from the author.)

They wrote that data from these older cohorts "do not reflect the lower current rates of cardiovascular disease that largely result from secular shifts in smoking, diet, exercise, and blood pressure control." The calculator's inputs include current smoking (yes or no), and levels of total cholesterol, HDL-cholesterol and systolic and diastolic blood pressure.

f) If secular shifts in cardiovascular risk factors are responsible for poor calibration, which of the above risk factors do you think are the most likely to be responsible?

Answer: Of the risk factors listed, exercise and (to a lower extent) diet seem the most plausible explanations for poor calibration, because they are not included in the calculator. This requires the reasonable assumption that both exercise and diet have beneficial effects on CVD risk not entirely captured by their effects on total or HDL-cholesterol or blood pressure.

Although smoking is included in the calculator, it is only as a dichotomous variable for current smoking. If smokers in recent cohorts smoke significantly fewer cigarettes per day than the smokers in the derivation cohorts, this could also explain overestimation of risk by the pooled cohort equations.

We would not expect secular shifts levels of risk factors included in the calculator as continuous variables to explain poor calibration. Thus, lower blood pressures and cholesterol levels should lead to lower predicted risk, not poor calibration.

g.) The secular decrease in CHD-death rates shown in the figure could also be partly due to widespread use of statins in later years. If you wish to use the calculator to help decide whether to start taking a statin, all else being equal, would it be better to have it be well calibrated for cohorts not taking statins, or cohorts in which statin use was common?

I want to use the calculator to estimate what my risk would be if I did not take a statin, so I'd prefer to have it be derived from cohorts not using statins (all else being equal).