Lifetime Risk calculator:

Background

A recent review identified at least 70 different cardiovascular scoring systems [ref NICE review]. Each requires a specification of three main components

1. What events are being predicted (CHD, CVD, CVD death etc)?
2. What factors are used to make predictions?
3. What measure of risk is used?

Here we examine various choices that have been made about the final component, and describe the measures incorporated in the JBS risk calculator. An Appendix contains the technical details of our implementation.

Relative and absolute risks

Additive scoring systems are constructed on the basis of standard regression models, for example logistic models for fixed-time events or Cox regression models for survival analysis. These estimate the *relative risks* associated with factors: specifically, if a factor x_1 multiplies a risk by r_1, then a series of factors $x_1,\ldots, x_k$ will produce an overall relative risk $r_1 r_2 \ldots r_k$. Taking logarithms shows the effect on the log(relative risk) to be $\log r_1 + \log r_k$, which can then be interpreted as an additive scoring system. The scores can therefore be interpreted as log(relative risks). Technically, for logistic regression models these relative risks are actually odds ratios, for Cox regression models they are hazard ratios.

However relative risks are not enough for decision-making – in order to decide whether interventions are worthwhile we need to know absolute risks of events, and the effect on those absolute risks of different interventions. Total scores are therefore transformed to absolute risks by assuming some baseline risk appropriate to the population of interest. In general, this has generated absolute risks of CVD events up to a fixed period of time in the future, say 5 or 10 years.

It is important to note that no allowance for 'competing risks' is made in these short-term prediction models, *i.e.* it is implicitly assumed that the individual will not die of non-CVD causes. As people get older, this becomes an increasingly inappropriate assumption.

10-year and 5-year risks

This has been the standard measure of CV risk since the early development of the Framingham score based on a logistic regression. The latest revision - Framingham General Cardiovascular Risk Score (1) is based on a Cox survival analysis but still only considers a 10-year horizon with no allowance for competing risks.

The Framingham score has been adapted by many other groups, for example JBS2 (2) uses a modified 10-year risk model, and the The New Zealand system (3) uses 5-year risks based on

Framingham. Other groups have used their own data to produce 10-year risk estimates, including QRISK2 (4), ASSIGN (5), HeartScore (6) and so on.

However, for reasons outlined elsewhere [crossref to discussion of limitations of 10-year risk] , there has been increasing attention to longer-term outcomes.

Absolute risks for longer horizons, allowing for competing risks

The Framingham system has been adapted to provide 30-year risk estimates (7). A number of different techniques were investigated, including simply putting together 3 10-year risk scores. Their recommended procedure was a full survival analysis that allows for *competing risks* from other causes of deaths.

Without allowance for other causes of death, the apparent cumulative risk of CVD will increase to implausibly high values over a long period, due to the implicit assumption that the individual cannot die from anything else. By including competing risks, a more realistic assessment of the risks of CVD is obtained, even though some apparently paradoxical behaviours can result. For example, someone who stops smoking may increase their risk of a CVD event, since the chance of their dying of non-CVD causes is substantially reduced, thus giving them a greater opportunity of a CVD event.

The mathematical details of the competing risk approach are outlined in the Appendix.

Heart-age / vascular age

These terms have been introduced as a way of visualising your current risk profile as a characteristic of your current state, rather than relating it to something that may or may not happen to you in the future. In smoking cessation research, telling patient's lung age has been found to have a beneficial effect on the smoking quit rate [need to add Parkes, not sure what is going wrong with Zotero] .

Heart-age and *vascular-age* are used interchangeably in the revised Framingham (1) score, each meaning the age of someone with the same risk but with all other risk factors set at 'normal' levels. This has been particularly emphasized in the New Zealand http://www.knowyournumbers.co.nz/heart-age-forecast.aspx risk calculator '*know your numbers'*, which shows an individual's 5-year risk, and how this is projected to increase during their lifetime. It compares this risk with an 'ideal' profile [non-smoker, TC/HDL ratio 4, BP 120/80] and so obtains a *heart-age* – the age of an 'ideal' person with the same current risks as the patient.

Lifetime risk profile

A lifetime risk perspective looks at the development of CV risk throughout an individual's life, using a variety of summary measures. Different summaries include

- *Cumulative risk of CVD*: the accumulating risk of a CVD event occurring before each age, retaining individuals who die from other causes and hence are not available to suffer an event. Risk of a CV event up to any specified age can be read off from this.
- *Survival without CVD*: this is a survival curve in which people are withdrawn as 'failures' if they suffer either a CVD event or death from another cause. It therefore displays the chance of being alive and without having suffered a CV event.

- *Mean (median) age at first CVD event*: the median is simply read off from the survival curve – the mean is the area under the survival curve, added to the current age.

The CardioRisk (9) analysis carried out a discrete-time competing risks analysis in 5-year bands, using Framingham for the 5-year CV risks and Office of National Statistics mortality data for non-CV risks. The Framingham data has been used to create lifetime risk (10) assessments, with allowance for competing risks, for an age range between 50 and 95. However these models do not appear to be available for general use. The QRISK lifetime risks model (11) uses a competing risks analysis, producing both summary CVD risk up to age 95, and a curve showing the cumulative risk of a CVD event.

Which summary risk measure is it best to use?

The academic literature is clear that there is no 'correct' way to communicate risks (12) – preferences and understanding can depend on an individual's educational level and in particular their numeracy, and also their personality, particularly their optimism. Selecting among the range of potential measures therefore presents a problem: while some individuals may be motivated by short-term risk, others may have a stronger attachment to longer-term outcomes such as maximising their chance to reach an age, say 70, where they might expect to know their grandchildren. Our use of multiple forms of presentation therefore reflects there is no 'best' way of communicating to people their future risks – different summaries may have different salience for individuals.

Methods in JBS3 Risk calculator

Use of QRISK

As noted previously, many different cardiovascular risk systems have been developed. QRISK Lifetime has been chosen as the basis for the JBS3 risk calculator as it provides lifetime risk calculations and is based on a UK population. However, all risk assessment systems have limitations in the factors that can be taken into account and the broad assumptions made, and so any numerical risk calculation must necessarily be interpreted in the light of the specific circumstances.

The data collected is based on the requirements of QRISK® lifetime http://www.qrisk.org/lifetime/

Item	***Coding***	***Adjusted hazard ratio (female)***	***Adjusted hazard ratio (male)***
Gender	Male Female	-	-
Date of birth	Age (years)	-	-
Smoking status	non-smoker former smoker light smoker (<10/day) moderate smoker (10–19/day) heavy smoker (≥20 cigs/day)	1 1.17 1.39 1.57 1.84	1 1.18 1.38 1.55 1.79

Self-assigned ethnicity	white (or not recorded) Indian Pakistani Bangladeshi other Asian, black African black Caribbean Chinese Other (including mixed)	1 1.42 2.04 1.61 1.14 1.03 0.69 0.77 0.99	1 1.50 2.05 2.14 1.32 0.71 0.70 0.79 0.90
Total cholesterol (mmol/l) HDL cholesterol (mmol/l)	Ratio of total cholesterol to HDL cholesterol (per unit increase)	1.17	1.18
Height (m) Weight (kg)	Body mass index: weight/height^2 (transformed to $\sqrt{}$(BMI/10) for females, log(BMI/10) for males)	1.32	1.54
Systolic blood pressure	per 20 mm Hg	1.13	1.11
Family history of coronary heart disease in first degree relative aged <60 years	No Yes	1 1.67	1 1.84
Townsend deprivation score - output area level 2001 census data	-3.95 (10%ile) -2.57 (30%ile) -0.81 (50%ile) 1.66 (70%ile) 5.34 (90%ile)	0.93 0.96 1 1.06 1.16	0.96 0.98 1 1.03 1.08
Treated hypertension (diagnosis of hypertension and at least one current prescription of at least one antihypertensive agent	No Yes	1 1.33	1 1.37
Rheumatoid arthritis	No Yes	1 1.43	1 1.37
Atrial fibrillation	No Yes	1 1.89	1 1.63
Type 2 diabetes	No Yes	1 1.67	1 1.60
Chronic renal disease	No Yes	1 1.67	1 1.59

Applying transformations and the hazard ratios as described in QRISK paper, we obtain hazard ratios for CVD events and non-CVD deaths. Applying these to a baseline age-specific risk of death (see Appendix) provides the probability of being alive and CVD-free at each age, and cumulative risk of CVD.

In QRISK the baseline ('average') individuals are assumed to have the following characteristics:

> *Baseline female*: Non-smoker, white, no clinical conditions, and with $\sqrt{}$ (BMI/10) = 1.605 (BMI = 25.8), SBP = 129.8, cholesterol ratio = 3.71 and Townsend = -0.30 (where the covariate is in terms of 5 units, and so corresponds to an original Townsend score of -1.51.

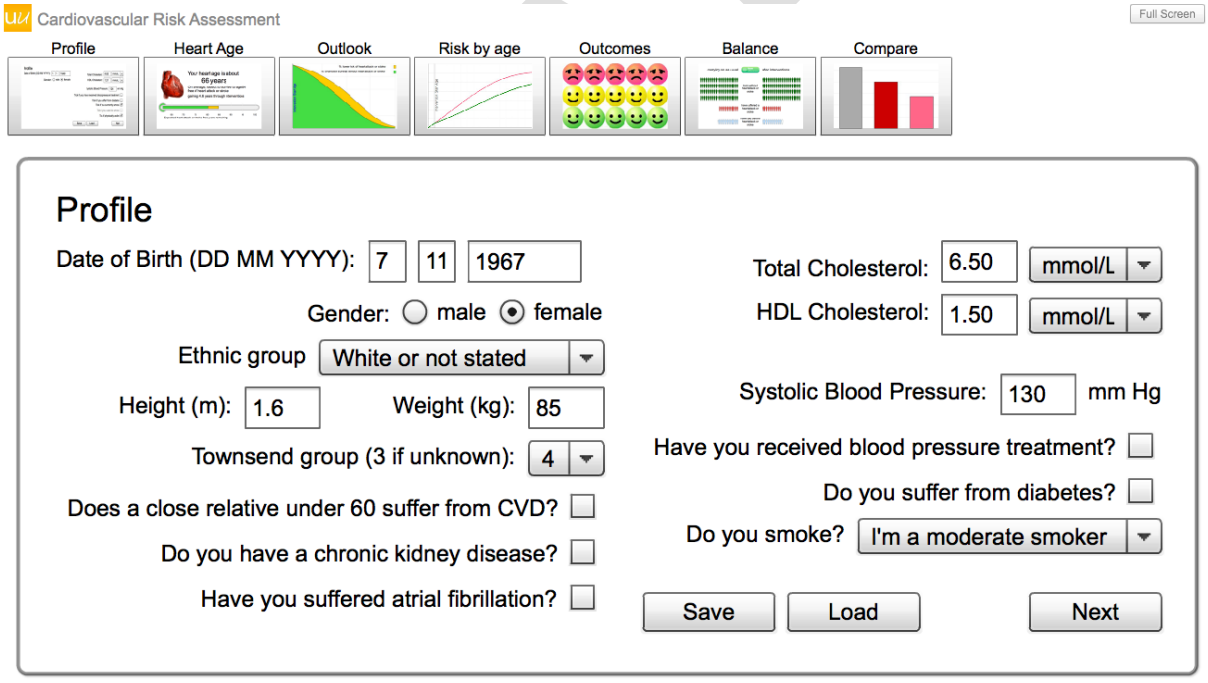
Baseline male: Non-smoker, white, no clinical conditions, and with log(BMI/10) = 0.968 (BMI = 26.3), SBP = 133.3, cholesterol ratio = 4.44 and Townsend = -0.165 (where the covariate is in terms of 5 units, and so corresponds to an original Townsend score of -0.83.

Output from JBS3 Risk calculator

Reflecting the findings that 'one size does not fit all' in risk communication, a range of different graphical presentations are implemented in the risk calculator. A deliberate choice is provided and it is intended that users will prefer some images and summaries to others. Below we summarise the available outputs, explaining their content and the reasoning behind the choices. Each screen allows observation of the effect of changing risk factors.

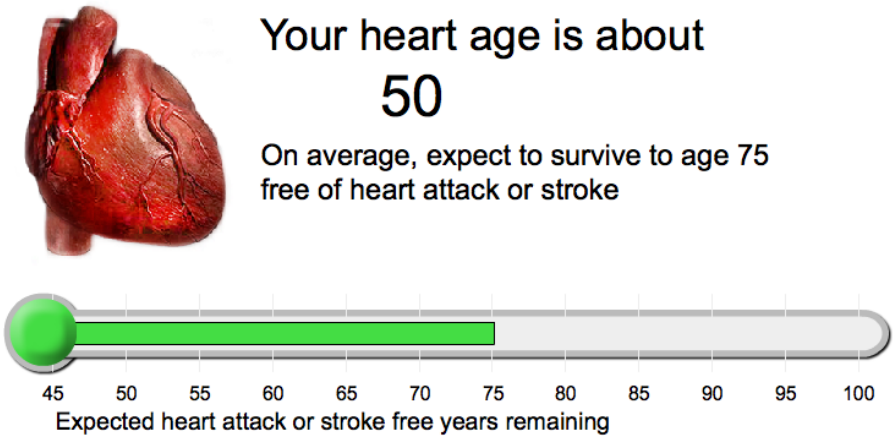
Running example: 44 year old woman, moderate smoker (15 a day), total cholesterol 6.5 mmol, HDL cholesterol 1.5 mmol, blood pressure 130/80, not receiving any medication, height 1.6m, weight 85 kg, not diabetic, no symptoms of CVD, no family history of CVD under 60, no kidney disease, no atrial fibrillation, Townsend quintile 4 (above average deprivation).

[NEED TO EDIT THIS AND SCREENS]



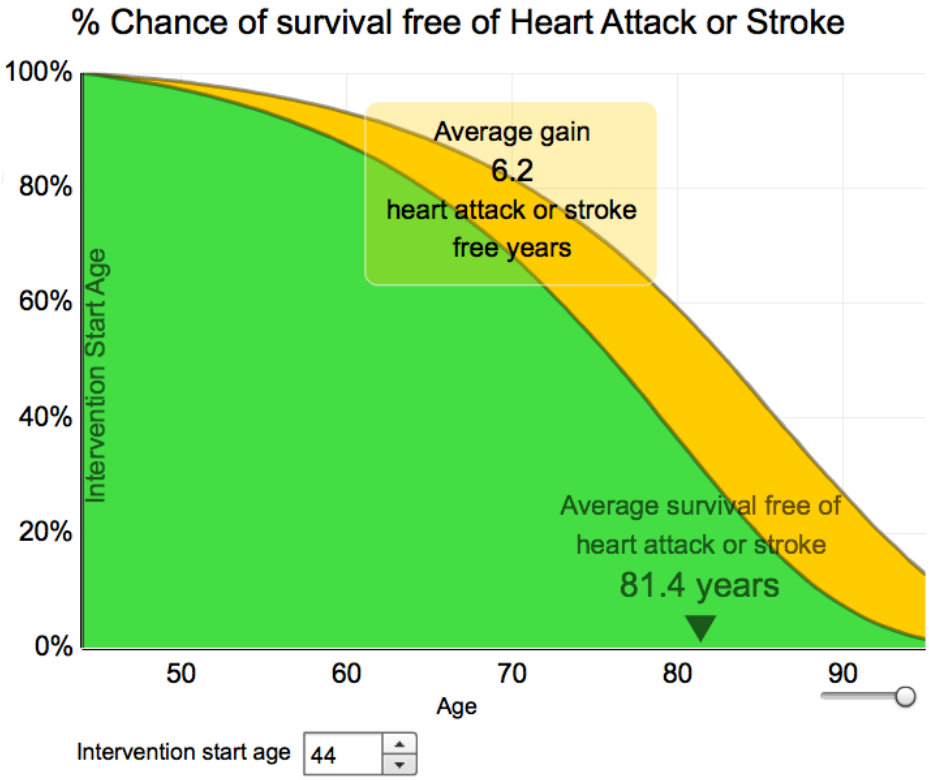
NOTE: NUMBERS ARE PROVISIONAL AND WILL CHANGE! ALSO LABELS AND 'EXPLANATION' WILL BE ADDED

1. **Heart age** . Heart/vascular age is the age of an individual with the same risk-profile in terms of lifetime-CVD risk, and the same gender and ethnicity, but otherwise a baseline patient. The expected age at the first CVD event is also provided.



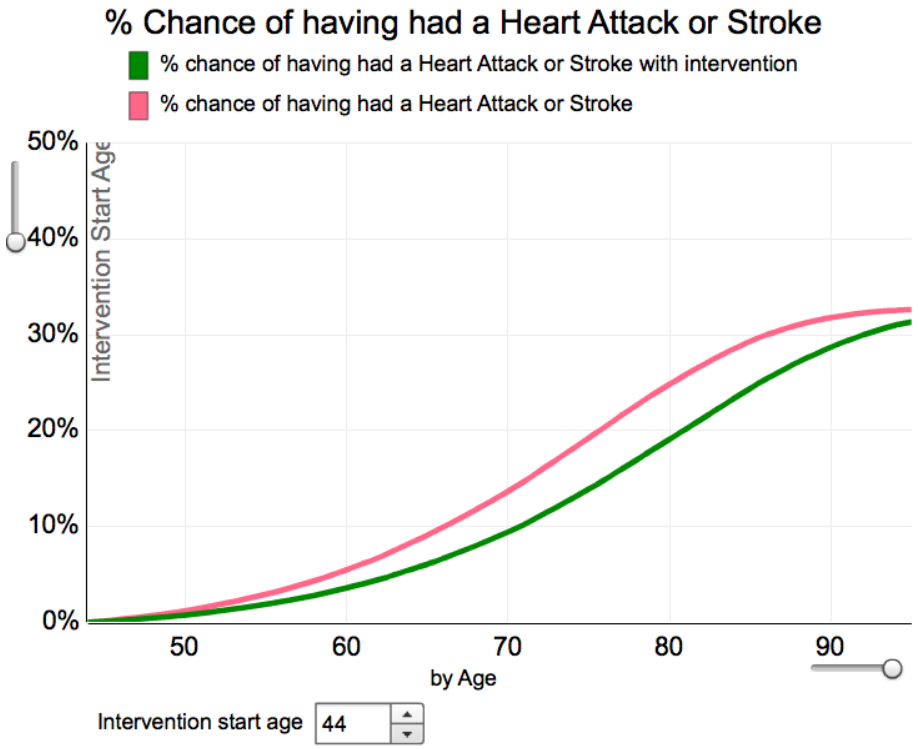
Comment: This is a simple introductory screen with no uncertainty.

2. **Outlook**. CVD-free survival, mean age at first event (gain in CVD-free life expectancy) – Effect of (possibly delayed) intervention effects,



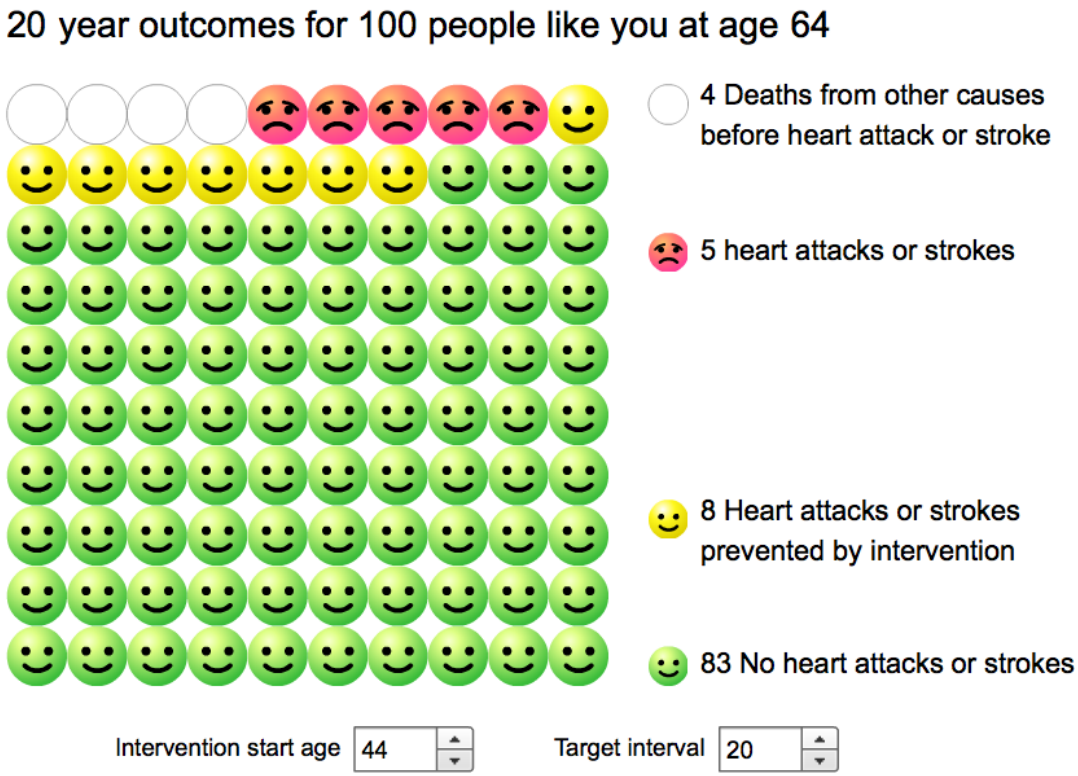
Comment: This is a survival curve and will require some additional explanation, although research suggests that many people can understand survival curves provided the instruction and labelling is clear. The curve moves up when considering an intervention, revealing a yellow area, which is intended to vividly portray the benefit of change. It is attractive that the total yellow area, which dominates the image, is in fact the mathematical expected gain in life-years until the first CVD-event, which is an appropriate quantity to communicate,

3. **Risk by age**. Cumulative risk of CVD up to 95.



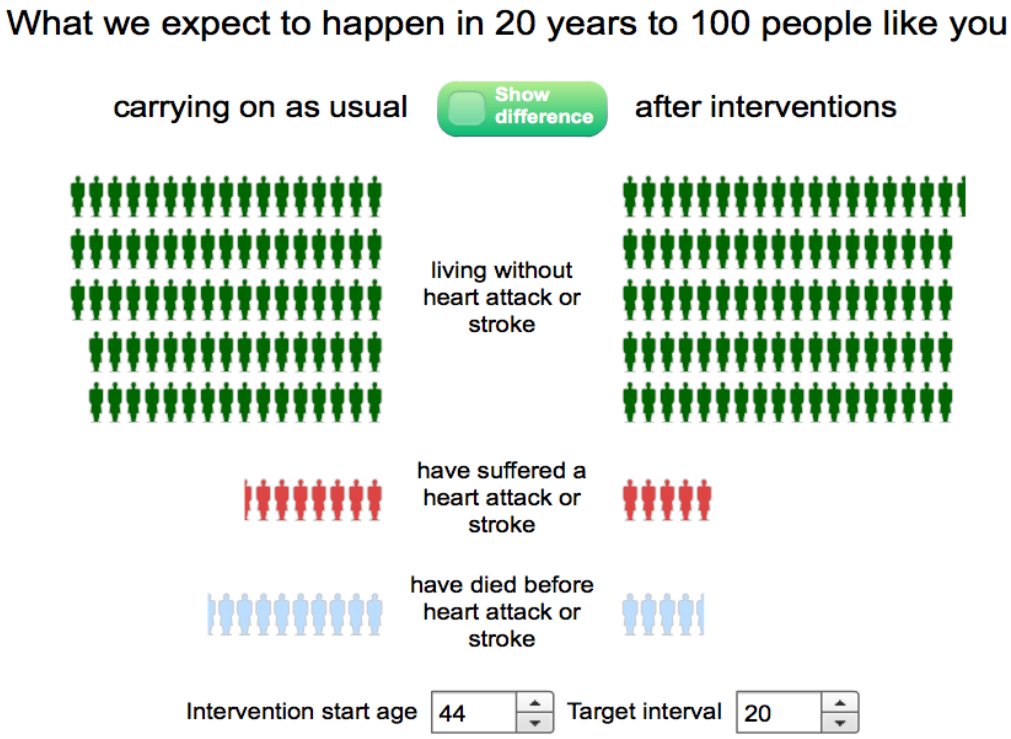
Comment: This is the standard output from QRISK lifetime. It can display some paradoxical behaviour due to the computing risk model: eg stopping smoking can increase the lifetime risk of CVD due to reducing the chances of dying from other causes.

4. **Outcomes.** 10-year (or any other period) risks including competing risks. Displayed as 100 'smilies' divided into three groups.

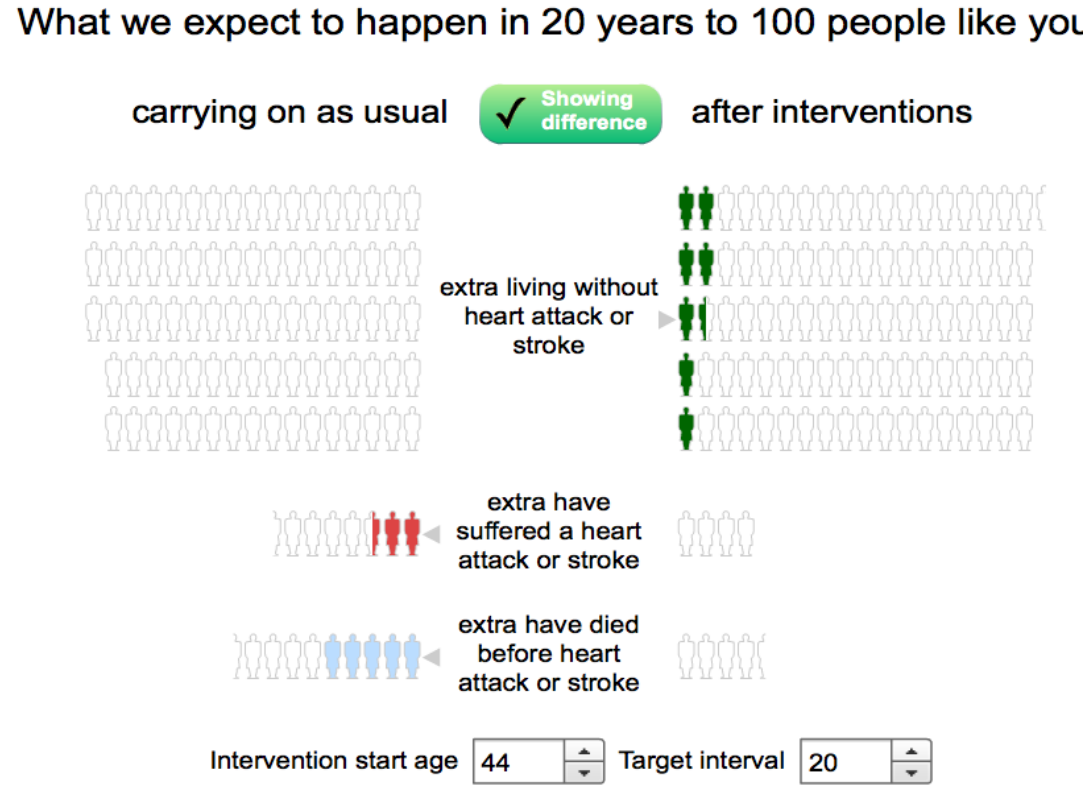


Comment: This is a fairly complex output featuring the 'smilie' grid used in other CVD risk applications. The deaths from other causes are downplayed in the graphic to avoid distraction. The crucial message is that out of 100 people like you who stop smoking, 8 will have a heart attack or stroke prevented over the next 20 years. The future interval can be easily changed to say 10 or 30 years.

5. **Risk Balance**. 10-year (or any other period) risks including competing risks. Displayed as stacked human icons, with and without intervention. –

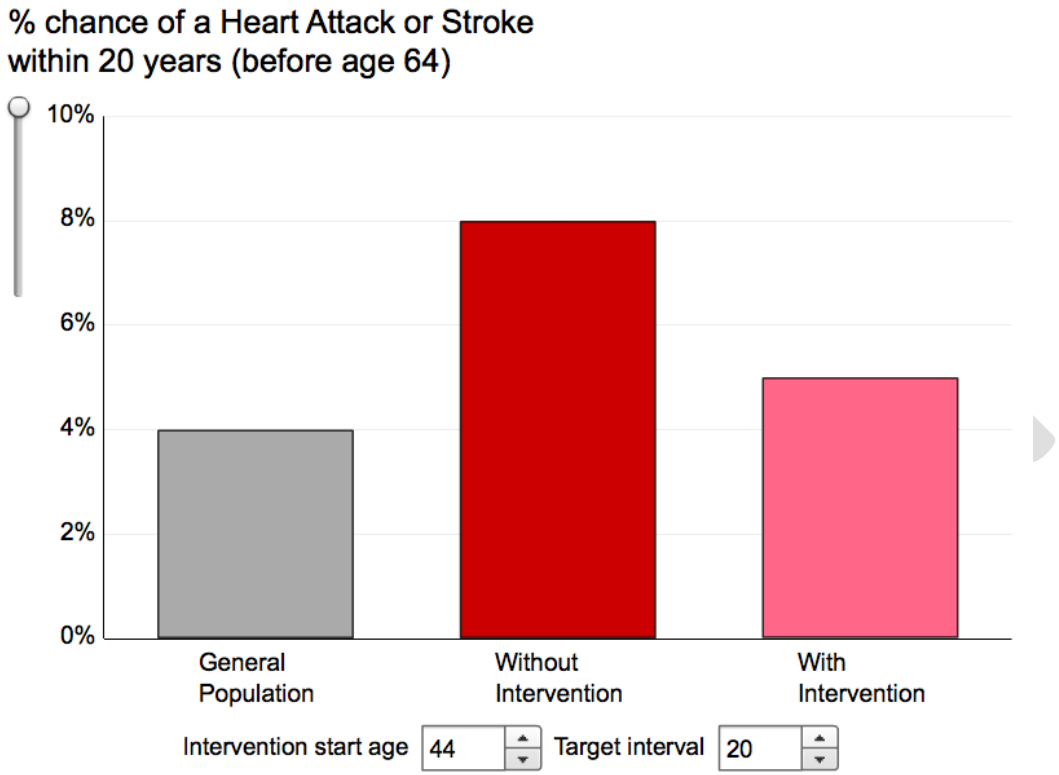


Comment: This features the same data as in the 'smiley' plot but arranged as stacked icon arrays, a format that has been shown to be attractive and comprehensible. But a comparison between the numbers with and without intervention is not straightforward - clicking on the Green 'Showing difference' button reveals the 'risk balance' plot below.



Comment: This is intended to make clear that the 100 people who have stopped smoking have prevented around 5 deaths from other causes and 3 CVD events over 20 years.

6. **Compare**. Comparison with population risk, ie same gender, age and ethnicity, but same as 'baseline' individual for other risk factors.



Comment: This focussed solely on the chance of a heart-attack or stroke, and clearly shows that stopping smoking will being the risk down to near an average person's.

Further examples?

Intervention assumptions

The table above provides an 'epidemiological' hazard ratio for CVD events, which is the ratio of the daily risks of two different people who differ in this risk factor. Changing an individuals profile will change the risks. However, when assessing the effect of intervening on an individual, we need to decide the hazard ratio when an intervention changes the risk factor, ie the ratio of the daily risks before and after the intervention *in the same person*. This may be less than the epidemiological hazard ratio, but also could be greater.

Direct evidence of the effect of changing behaviour or intervening to change physiological measures comes from clinical trials, or it is possible to make indirect inferences from observing cohort studies. However evidence is limited and some assumptions are inevitable.

Blood pressure

In a recent review of trial data, Law *et al* (13) estimated around 50% reduction in CVD events per 20mmHg reduction in SBP, similar to that expected from the epidemiological evidence for CVD mortality from the Prospective Studies Collaboration (14). This is substantially larger than the association found in the QRISK lifetime-risk formula in Table XXX, which uses a hazard ratio of 1.13 (female) and 1.11 (male) per 20 mm Hg increase in SBP, ie around 11% reduction in event rate per 20 mmHg reduction in SBP. This is a substantial difference, perhaps due to the many other correlated factors included in the risk formula. We adopt the Law *et al* estimate, which corresponds to a hazard ratio of 0.966 per mmHg reduction in SBP.

Law *et al* find no impact of blood-pressure intervention on non-vascular causes of death, and so intervening on blood pressure is not assumed to have any effect on non-CVD mortality.

Cholesterol

A recent meta-analysis estimates a 22% reduction in vascular event rate per 1 mmol/L reduction in LDL (15), which is fairly robust over a wide range of baseline conditions.

Since TC and HDL are the most familiar to general practitioners, we adopt non-HDL = TC – HDL as our measure of intervention: assuming that non-HDL is approximately 1.24 x LDL [**add robinson ref**] leads to a hazard ratio of $(0.78)^{1/1.24}$ = 0.82 per I mmol/L reduction in non-HDL cholesterol.

No impact of cholesterol interventions on non-vascular causes of death is assumed.

Smoking

Since 'former smoker' is a category in the QRISK lifetime score, we assume that an intervention to stop smoking leads to the epidemiological risk associated with 'former smoker'.

There is also a clear impact on non-vascular causes of death, and again the 'former smoker' risk is assumed.

Weight

We do not have reliable evidence on the effect of weight reduction in its own right, and so currently interventions to change weight does not change the lifetime risk calculations.

Caveats

It is crucial to acknowledge that any numerical summary that comes from a risk calculator is not *the* risk of an individual. Many assumptions are used in the calculation, and limited information on each

individual is included so that other evidence, such as genetic information, would almost certainly change the assessed risk. It is perhaps best to think of the quoted risk as being constructed on the basis of available evidence from populations and limited information on the individual: it does not *exist* as an objective state of the world and hence it is unreasonable to seek great precision.

Nevertheless it is reasonable to seek some *discrimination*, in that different people do get assigned a range of risks, and *calibration*, in that if we quote a risk of 70% for an event, then it will happen in around 70% of cases. See QRISK lifetime publication (11) for validation details.

List of QRISK caveats – who should use it etc.

Additional caveats?

Validation tools

Field trials (KB to cover)

Appendix: technical details of risk calculator

Lifetime risk model

Age is taken as the time axis in this model, and so age does not appear as a risk factor in the model. Rather the risk factors adjust the baseline age-specific risks for an 'average' individual. The term 'hazard' is used for the risk of an event occurring in a small interval, given CVD-free survival until the start of that interval. Baseline hazards are required for the whole age range, both for CVD-events and non-CVD deaths.

Hazard ratios

The following steps are carried out independently for each gender and each cause of 'failure', ie CVD-events and non-CVD deaths.

Each individual provides a set of risk factors x (possibly transformed to measure deviation from 'average'). A set of coefficients b is applied to the risk factors to create a linear score $S = \Sigma_i \, b_i \, x_i$ for each of CVD-events and non-CVD death.

With a proportional hazard assumption this is a log(hazard ratio). Taking the exponent e^S gives the hazard ratio to be applied to the baseline hazard for each of the two causes of failure. For example,

if the baseline risk of a CVD-event in an age-interval t is assessed to be h_t, then the risk for the specific individual is $h_t e^S$. Within each age-interval t, we therefore obtain the 2 crucial quantities:

- a_t : the risk of a CVD-event in interval t, given CVD-free survival up to the start of interval t
- b_t : the risk of a non-CVD death in interval t, given CVD-free survival up to the start of interval t

From these two quantities all else follows.

The 'life-table' below follows Ulrich et al [ref] although shows arbitrary age intervals (these were 5 years in Ulrich et al, but in QRISK the table is at a very fine scale, in our algorithm we use XXXXX). The life table begins at the current age of the individual.

	Interval 0 - starting at current age	*Interval t-1*	*Interval t*
Risk of CVD event in age interval	*0*	a_{t-1}	a_t
Risk of non-CVD death in age interval	*0*	b_{t-1}	b_t
Proportion of original cohort having non-CVD death in interval	*0*	c_{t-1}	$c_t = e_{t-1} \times b_t$
Proportion of original cohort having CVD event in interval	*0*	d_{t-1}	$d_t = e_{t-1} \times a_t$
Proportion surviving free of CVD at end of interval	*1*	e_{t-1}	$e_t = e_{t-1} - c_t - d_t$
Cumulative proportion with CVD events by end of interval	*0*	f_{t-1}	$f_t = e_{t-1} + d_t$
Cumulative proportion dying from non-CVD causes by end of interval	*0*	m_{t-1}	$m_t = m_{t-1} + c_t$

Note that $e_t + f_t + m_t = 1$ for all t.

Changing the risk factors changes the hazard terms a_t and b_t to a_t^* and b_t^* : repeating the life table calculations leads to a revised CVD-free surviving proportion e_t^* and revised cumulative CVD proportion f_t^*

From these quantities we obtain the information for the risk calculator as follows (assuming we are examining 10-year risks).

- **Heart/vascular age**: For the individual with risk factors x, we find a single summary risk measure $f(x,age)$, such as expected age at first CVD event, or current 1-year CVD risk. We then find the 'heart-age' where $f(x,age) = f(x_A,heart\text{-}age)$, where x_A is a 'normal' person. We assume that 'normal' is as in comparing with a population: the same age, ethnicity, and gender, but otherwise as baseline patient. Currently we use the 1-year CVD risk as a matching criterion. This is calculated iteratively.
- **CVD-survival curve**: e_t : Probability of surviving free of CVD at end of age interval t
- **Median age at first event**: the age at which $e_t = 0.5$
- **Expected age at first event**: $\Sigma_t\, e_t$ x (width of interval t)
- **CVD-free years gained by changing risk factors**: $\Sigma_t\, (e_t^* - e_t)$ x (width of interval t)

- **Cumulative CVD risk curve**: $f_{t:}$ Cumulative proportion with CVD events by end of age interval *t*
- **10-year risks**: for interval *t* corresponding to 10-years older, e_t *(*green: alive and CVD -free*)* f_t (red: CVD event*)* m_t *(clear:* death from non-CVD cause*)*
- **Comparison with population**: Based on 10-year risks for someone who has the same age, ethnicity, gender, but other factors match the QRISK baseline patient.

Implementation of QRISK in JBS calculator:

Source of code

The JBS3Risk calculator is a Flash program which runs inside desktop internet browsers that support the Flash v10.2 player. The JBS3Risk model uses a port of the QRISK®-lifetime-2011 C source to actionscript 3. Further details of the QRISK®-lifetime-2011 code are available at http://www.qrisk.org/lifetime/index.php. The QRISK®-lifetime-2011 lifetime code is licensed under the GNU Lesser General Public License v3.0 held at http://www.gnu.org/licenses/lgpl-3.0-standalone.html. The JBS3Risk calculator contains a modified version of that code and is licensed under the GNU General Public License version 3, held at http://www.gnu.org/copyleft/gpl.html. Refer to the calculator front matter for access to source code and license. The JBS3Risk model reads in data from verbatim copies of the QRISK®-lifetime-2011 data tables which are published under the Creative Commons Attribution-NoDerivs 2.0 UK: England & Wales License available at http://creativecommons.org/licenses/by-nd/2.0/uk/.

Validation against online version

In order to speed up the calculations, the JBS3Risk calculator accumulates values from the QRISK® data tables in bins of size 0.02 years before applying the model. A suitable bin size was determined in randomised parameter tests: 99% of the random parameters generate a score within 0.2% of the native score, with worst case less than 1% error. This is for both *n*-year risk and lifetime risks.

Townsend

Rather than using postcodes as a predictor of the Townsend social deprivation score, the JBS3Model uses a scale based on the quintiles of the Townsend distribution, -3.94, -2.57, -0.8, +1.66 and 5.34. Quintiles were calculated from 2001 census data, published at http://www.apho.org.uk/resource/item.aspx?RID=47504.

Heart Age

The JBS3Risk calculator heart age model does not use QRISK®HeartAge code. The JBS3Risk heart age is the age at which a baseline patient with the same gender, ethnic group and age would match the calculated lifetime risk.

Conditions, legal stuff

(There'll be some other credits that we probably only need to hang off the front matter within the program)

REFERENCES