Optimising Cost-Effectiveness in Risk-Based Screening Programmes: A Modelling Approach

HESG/CES 2011

Chris Sampson, University of Nottingham
Room B21 Sir Colin Campbell Building, Jubilee Campus, Nottingham, NG7 2TU, UK
chris.sampson@nottingham.ac.uk

Prof. Marilyn James, University of Nottingham
Room B12a Sir Colin Campbell Building, Jubilee Campus, Nottingham, NG7 2TU, UK
marilyn.james@nottingham.ac.uk

Prof. David Whynes, University of Nottingham
Room B32 Sir Clive Granger Building, University Park, Nottingham, NG7 2RD, UK
david.whynes@nottingham.ac.uk
Abstract

The need to ensure the cost-effectiveness of interventions has lead to an increased scrutiny of screening programmes. As a result there has been some argument in favour of individualised screening interventions, with risk-based screening being evaluated against fixed-interval programmes. However, the level of risk at which individuals should or should not be screened is not currently guided by the expected cost-effectiveness of screening at a given risk level.

We present a discussion of the potential for optimisation of screening based on an individual's estimated level of risk. We consider the needs and objectives of existing screening programmes and the ways in which modelling methods may be capable of estimating an optimal level of cost-effectiveness. We highlight the key questions that a generalisable modelling methodology must be able to answer in order to be widely applicable to a variety of screening programmes.

This study will inform the development of a generalisable modelling methodology to be tested in the future. We hope that the findings of this study will enable us to develop a comprehensive decision analytic model, able to present decision makers with the most cost-effective screening programme, based on individual risk.
Introduction

Decision analytic modelling is commonplace in the evaluation of health care interventions. As a methodology it offers great flexibility by enabling the adept modeller to use data from various sources, and estimate the uncertainty around parameters, in evaluating an intervention. Using these methods it is possible to simulate complex processes and estimate outcomes without the high cost of clinical trials. Modelling methods have been used extensively and effectively in the evaluation of screening programmes. Recently there has been some discussion around the implementation of risk-based screening (Bradbury & Olopade, 2006), including some evaluations of proposed programmes (Aus, Damber, Khatami, Lilja, Stranne, & Hugosson, 2005).

Currently screening programmes implicitly consider risk by offering screens to a limited section of the population. The nature of screening is usually heterogeneous for all participants, regardless of their differentiating characteristics. Risk-based screening seeks to alter this. At its most basic level it involves defining groups of high risk and low risk individuals and offering them different screening interventions accordingly. The purpose of this is to increase the overall cost-effectiveness of screening by reducing the number of ineffective screens and increasing the number of more effective screens, potentially with a reduced overall cost. As discussed below, existing screening programmes consider risk in a very limited way.

Support for risk-based screening, and its limited uptake, represents an acknowledgement of a relationship between risk and cost-effectiveness. However, definitions of high and low risk, and the corresponding adjustments to screening programmes, are currently defined by expert opinion. It is for this reason that they are restricted to defining risk categorically. In reality, estimations of individuals’ underlying risk are continuous on a probability scale, implying an indefinite number of possible programmes each with differing levels of cost-effectiveness. It would be prohibitively expensive to evaluate an indefinite number of differently defined screening programmes, in a trial setting, in order to find the most cost-effective programme. Methods of decision-analytic modelling have the potential to address this problem. In this paper we investigate the feasibility and capability of modelling methods to do this. We discuss the background to risk-based screening and the
theory and assumptions that underpin our notions of optimisation. We review current relevant methodology and existing screening programmes in order to further develop our knowledge. This study will inform our future work; our intention being to construct a modelling methodology able to optimise the cost-effectiveness of screening programmes based on individual risk.

**Theoretical optimisation of screening**

In economics, optimisation is paramount. This applies to economic evaluations and particularly to decision-analytic modelling. In the case of economic evaluations in health care it is usually cost-effectiveness that we seek to optimise. Often in the form of optimal net benefit based on willingness to pay for QALYs. Most commonly this methodology is used in the evaluation of 2 or more health care decisions. Health care interventions are usually well defined in terms of their administration; evaluated drugs are usually of a fixed dosage and therapeutic interventions tend to be fixed-term. Economists are therefore presented with a finite, and usually small, number of possible interventions for evaluation. Based on the costs and effects of these different interventions an economist seeks to elicit the optimal choice for a given population. In the case of risk-based screening there is not a definitive set of interventions established for evaluation. In addition to the screening procedure itself, screening interventions can be varied in two ways; the risk level at which an individual is screened and the frequency with which screening should take place. This dynamic is a feature of all screening programmes. Time and risk are both continuous variables, infinitely divisible, meaning that we are presented with an infinite number of possible interventions ranging from constant screening for the whole population to no screening for anybody. One of these infinite possibilities represents the most cost-effective intervention, dependent on an individual’s risk level. At present there is not a generally accepted way of estimating this. For it to be possible to define this optimal point, based on an individual’s risk level, there are a number of relationships that must be defined.

**The relationship between risk and value**

Health care services in the UK and elsewhere collect many data on individuals’ health and personal characteristics. Through the analysis of these data health care professionals are, in many cases, able to identify individuals at risk of developing a
given disease and therefore refer them on to screening. Not all individuals are
referred to screening because health care professionals understand that there is no
value in screening an individual with very low or no risk. At present, as discussed
below, screening programmes tend not to be risk-based. This means that this
relationship between risk and value is largely implicit in that referral to screening is
based on an individual’s characteristics and not a calculated risk level. However,
there are tools available that enable us to estimate an individual’s risk level using the
best data and knowledge available.

In their most basic form, screening outcomes are binary. An individual can either
screen positive or negative. Issues regarding false positives and false negatives are
addressed later in the paper. Figure 1 and Table 1 highlight our assumptions around
screening programmes and their evaluation at different levels of individual risk. The
implications of relaxing these assumptions are discussed later in the paper. Figure 1
demonstrates a basic screening model. In this model an individual has an underlying
risk which determines the probability \(r\) that they develop the disease in a given
period of time. Table 1 demonstrates how an individual’s level of risk might
determine the expected costs and outcomes and, therefore, the cost-effectiveness of screening
them.

Health care providers, such as the NHS, do not screen the whole population for
every possible illness on a daily basis. This is because most people, at any given
time, have very little or no risk of developing a given disease. There is no value in
screening an individual who has no risk, while there may be great value in screening
an individual at high risk. As such, the relationship between an individual’s risk of
developing a disease and the expected cost-effectiveness of screening them will tend
to be positive. If we assume that there is a fixed effect on outcomes of a positive
screen \(Q_s\) across risk levels, and assume there to be no effect of a negative screen,
then the relationship between an individual’s risk and the expected incremental
benefit of screening will be positive and linear, as shown in Figure 2. The implications
of relaxing these assumptions are discussed below.

Furthermore, if we assume a fixed cost of screening \(C_s\) across levels of risk, we will
find a relationship between risk and cost-effectiveness similar to that shown in Figure
3. The higher the risk level the more likely it is that an individual will screen positive
Work in progress. Do not cite.

and therefore benefit as intended from the screening programme. This means that the expected cost per outcome for an individual with a high risk is very low, while for those at a low risk it is very high.

Figure 3 highlights that, for a given level of willingness to pay per expected gain, there will be a corresponding minimum level of risk at and above which individuals should be screened, and below which the expenditure would not be deemed worthwhile. Current practice does not estimate this figure, but we believe modelling methods to be capable of doing so.

A number of factors will influence the relationship between value and risk. One can imagine that high screening costs ($C_s$) may mean that those at low risk should not be screened. Alternatively, if screening positive can have a substantial impact ($Q_s$) on an individual’s well-being it might be cost-effective to screen even those at a very low risk of developing a disease. Figure 3 demonstrates that the cost-effectiveness of screening at various risk levels could vary dramatically. Current practice does not consider the impact of risk on cost-effectiveness. We hope that modelling methods will enable us to do so by finding the optimal point of this relationship.

The relationship between screening frequency and value

Figure 3 demonstrates the potential for optimality in screening based on risk. However, as explained above, screening can also be varied in relation to time. In many cases screening programmes involve recurrent screens. One can consider that the optimally cost-effective screening interval for an individual is greater the lower their risk, because more regular screening is associated with higher costs and a lower average probability of a positive screen. This suggests that screening those with a lower risk less often could increase the cost-effectiveness of their screening. In implementing a standardised programme it is very unlikely that screening will be optimal in terms of cost-effectiveness. Furthermore, this remains true in risk-based screening programmes where individuals are defined as either high or low risk.

The consideration of recurrent screening affects our model in a number of ways. If we assume that the cost of each screen is constant in frequency, then the relationship between frequency of screen and expected cost will be positive. That is to say that two screens per year will cost twice as much as one screen per year. Using modelling methods it may be possible to simulate a variety of screening
programmes for individuals of a given risk level. In doing this we may have the potential to define the optimal screening interval for individuals based upon their own level of risk.

**Optimisation of screening in practice**

In designing a modelling methodology it is important to understand the nature and purpose of the processes being modelled. As stated above, screening programmes that are explicitly based on individual risk do not currently exist. Here we review existing programmes, considering their similarities and differences, in order to inform the development of our model.

A number of screening programmes have been established in England. These programmes differ in terms of their target populations, screening intervals and the underlying purpose of the programme itself, as summarised in Table 2. None of the current programmes are aimed at the whole population. This is because there is, at the least, some basic understanding about individual risk and capacity to benefit from screening. All of the screening programmes identify a target population and define particular screening intervals. Breast, abdominal aortic aneurysm, cervical and bowel cancer screening programmes are all limited by age. Cervical and breast screening is offered only to women while abdominal aortic aneurysm screening is only offered to men. The screening population for diabetic retinopathy is defined by a diagnosis of diabetes. Screening intervals for these programmes vary between 1 and 5 years and some programmes, including antenatal and newborn screening, involve a one-off screen. The programmes also differ in the nature of the screening procedure itself. For example, diabetic retinopathy screening involves a two stage screening process.

Something that all of the programmes have in common is an underlying, though implicit, consideration of risk in their definition. The screening programmes tend to have well-defined aims, with most programmes aiming to reduce mortality rates. The purpose of cancer screening programmes is to identify the presence of cancerous cells at an early stage in order to improve the probability that they can be removed successfully, avoiding premature death. Diabetic retinopathy screening has a similarly defined primary goal of preventing blindness. These outcomes lend themselves well to techniques of decision modelling.
**Existing methodologies**

A review of the literature highlights a number of methodologies that may be used to estimate the relationships discussed above. These methods can be divided into two categories; models for the elicitation of risk and models to estimate cost-effectiveness. Karnon et al (2007) review the use of modelling in the design of screening programmes. The authors discuss the advantages and disadvantages of a variety of mathematical and statistical modelling methods in relation to their use in the evaluation of screening programmes. These methods include Markov models, decision tree analyses, Monte Carlo simulations and discrete event simulations. Furthermore, the authors address the use of survival models in their application to time-dependent Markov states and transition probabilities.

**Elicitation of risk**

A key requirement of our model will be to calculate the relationship between an individual’s risk level and the value of screening them at a given interval. We must therefore be able to calculate the likelihood that an individual will develop the undesirable health outcome within a period of time. To do this we can employ statistical methods of risk elicitation, based on individual characteristics. Karnon et al (2007) discuss the use of these methods in existing cost-effectiveness studies in screening. A semi-parametric model such as a Cox proportional hazard model could be used, while non-parametric models are likely to be of less use. In our case it is likely that we would choose to adopt a standard parametric survival model. Aspelund et al (2011) use a Weibull proportional hazard model to estimate factors of individual risk in diabetic retinopathy. Weibull models consist of two parts. Firstly, they provide us with the cumulative probability of developing a disease and secondly they estimate the relative risk associated with our defined factors. Weibull models are regularly used in the elicitation of time-dependency of outcomes in decision models, and may be ideal for our purposes. However, under our assumptions, individual risk is a parameter to our cost-effectiveness model and can be calculated separately. As such the most appropriate model may vary dependent on the population.

**Cost-effectiveness analysis**

Karnon et al (2007) comprehensively review the use of modelling methods in the evaluation of screening programmes and we do not intend to reproduce this
information here. The implications of the study’s findings inform our own conclusions below.

Generally, models can carry out simulations of either patient cohorts or individuals. Both cohort and micro-simulations tend to involve the use of Markov models and/or decision tree analysis. Decision trees are arguably too limited in their scope for evaluating complex interventions and outcomes. Markov models, on the other hand, require fewer assumptions and can be adapted and augmented to fit the purpose of a model. Karnon et al (2007) find that;

“The cohort Markov model is the standard technique used to model the economic impact of healthcare interventions over time, and is also the most common approach identified in the review”.

However, the best models often adopt multiple methods, including decision tree processes either at the beginning or end of Markov simulations or individual cycles (Khandler, Dulski, Kilpatrick, Ellis, & Mitchell, 2000). Markov models can take a number of forms, employing transition probabilities or being based on the amount of time an individual spends in a state (Neilson & Whynes, 1995). Markov assumptions can be relaxed by employing methods such as ‘tunnel states’, which can allow for time-dependent detection rates. Such methods could also be used to model changes in risk factors post-baseline. It is also possible to sample transition rates from a distribution, allowing for the quantification of uncertainty around such parameters (Parmigiani, 2002). By combining Markov modelling methods with those of risk elicitation it is possible to, for example, calculate transition probabilities that are dependent on variables such as age.

The complexity of a model such as the one we are trying to create lends itself to micro-simulation. Micro-simulation, in contrast to cohort simulation, is very flexible in that individual characteristics can be accounted for. Discrete event simulation, for example, is based on defining the next event given an individual’s current state, and the time to that event. However, while micro-simulation is more flexible, it also demands more time for the building and parameterising of the model and is dependent on an ability to carry out a far greater number of simulations. Discrete event simulation, for example, can only be carried out using first-order Monte Carlo simulation, which increases the analysis time dramatically.
Figure 1 shows a decision tree for a basic screening intervention. We believe this structure to be generalisable across screening programmes and clinical areas. In practice this model would be followed by a Markov simulation of disease progression, specific to the clinical area of interest.

Implications
Using the information gathered and discussed above we can identify some of the limitations of our proposed model, and the implications of relaxing particular assumptions. We can also highlight some of the challenges we face in this work. Karnon et al (2007) devise a seven-point checklist of features to consider in an assessment of the quality of modelling studies in the evaluation of screening programmes:

1. Research question
2. General modelling approach
3. Model structure
4. Modelling technique
5. Model population
6. Validation and calibration
7. Issues specific to [clinical area].

This checklist addresses the requirements of a modelling methodology and as such we use it to discuss the requirements of our model and the challenges we face in relation to each point, including the relaxation of particular assumptions.

Research question
Our model will seek to define the most cost-effective screening interval for an individual of any given risk level. It should be capable of doing so in terms of lifetime costs and QALYs. In order to optimise the screening programme in terms of screening interval it will involve the simulation of many possible risk levels and screening intervals.

As economists, our primary outcome tends to be QALYs. However, as discussed above, different screening programmes have differing objectives. As such our model will need to consider non-QALY outcomes. This may mean that our model will have
to be able to incorporate willingness to pay thresholds for different outcomes of screening. For example, we may need to consider willingness to pay for screening that avoids blindness or may have to adopt a ‘mortality reduction’ approach.

**General modelling approach**

Our model must be capable of considering disease progression, for which a Markov model may be most appropriate. Our basic model will assume that disease progression is not dependent on an individual’s level of risk. This may be an unreasonable assumption as those at a higher risk level may be likely to progress more quickly through the disease or vice versa. However, it seems more likely that factors that define risk may also define disease progression and as such individual characteristics could be accounted for in transition probabilities, for example, rather than in varying models based on risk. Our theory is also based on an assumption that the value of screening increases linearly with risk. While it seems unlikely that this relationship is not positive it is possible that returns to screening are a function of risk. It may be the case that returns to screening are diminishing in risk. While this relationship would not undermine our model it represents a potentially important relationship that could affect results dramatically. We have also assumed that screening has no negative effect. It is possible that screening, particularly if an individual receives a false positive result, can have a negative impact on an individual’s well-being. This assumption could be relaxed by testing this relationship and incorporating it into our model.

We also intend for the structure of our model to be generalisable. From the review above it is clear that there are a number of differences between screening programmes, and we hope to be able to accommodate these.

**Model structure**

The structure of our cost-effectiveness model will be defined in the same was as a standardised non-risk based model. Disease progression is likely to differ for those who are and those who are not screened. This difference will have to be incorporated in to our model. However, it might also be the case that disease progression differs on the basis of an individual’s risk, in which case a relationship between these parameters will have to be established.
Existing screening programmes have differential features and involve specific techniques that our model structure must be able to incorporate, such as multi-stage screening for diabetic retinopathy.

An important feature in the evaluation of screening programmes is rates of attendance. It is important that our model structure is able to address this issue and its relation to risk. Based on our assumptions, for a given level of risk, non-attendance will lead to the same outcomes and costs as not being screened. As such, if we can estimate the rate of attendance, it will be possible to calculate the associated costs and outcomes. However, it might be the case that the rate of attendance varies at different levels of risk. In this case this would represent another relationship that we must define and include as a parameter in our model. Furthermore, disease progression may vary in attendance and, if attendance varies based on risk level, this means that disease progression may vary, indirectly, based on risk. If the rate of disease progression is not random across risk levels then this relationship must also be considered.

Modelling technique
Karnon et al (2007) conclude that studies often fail to justify the use of complex mathematical techniques. This is because complex methods, while more flexible, can be computationally demanding. The purpose of our model is to evaluate screening programmes at varying levels of risk and willingness to pay, in the case of micro-simulation models we would require hundreds of millions of simulations in order to adequately consider uncertainty around our estimates. This would make value of information analysis, and the estimation of uncertainty, using our model very difficult. While methods exist that may address this issue, such as Gaussian process modelling (Stevenson, Oakley, & Chilcott, 2004), it is likely that our model would benefit from a more basic process of cohort simulation. Similarly we must be able to justify our choice of risk elicitation model. Various models should be trialled to establish which provides the best estimates.

Model population
Screening is defined as the identification of unrecognised disease in an asymptomatic population (Wilson & Jungner, 1968). The implication of this is that the majority of a given population is potentially eligible for screening. In some cases a
large proportion of a population can be defined as having no risk, such as in the case of diabetic retinopathy. However, in most cases our model will be considering the effects of screening for the entire asymptomatic portion of a population. Individual risk levels should be elicited from a sample that is representative of this population in order to adequately estimate individual risk factors.

In our discussion of optimisation, we assume sensitivity and specificity rates to be constant across levels of risk. If this is not the case then the relationship between sensitivity and specificity rates and individual risk will affect our results. In the development of our model we must consider the possibility of defining sensitivity and specificity rates based on risk.

We have also assumed that the cost of screening is constant across our population. In some cases there may be variable costs of screening that are dependent on individuals’ characteristics. Our model should be capable of incorporating these.

**Validation and calibration**
We intend for our model to be comparable against other non-risk based models. It should be possible to evaluate risk-based screening, defined by our model, against standardised screening programmes. As such the disease progression models used should be guided and informed by existing, validated models.

**Issues specific to [clinical area]**
As highlighted by the review of screening programmes above, there are a variety of screening characteristics and purposes that must be accommodated by our model. It is possible that, in some cases, screening may have unintended benefits. For example, screening for one disease may lead to the diagnosis of another. Similarly, those who attend screening may receive enhanced care generally and be more able to manage illness. Multi-stage screening interventions will also need to be modelled appropriately where different stages may be associated with differing rates of false positives and false negatives. It may be possible that, in some cases, screening can have a negative impact, meaning that there would be a detrimental effect of screening for false positives. This could and should be incorporated as it may have implications for the optimal screening interval.
Conclusion

Cost-effectiveness is a vital input to screening policy decisions, alongside other factors such as patient preferences. Existing screening programmes in England are designed with an implicit recognition of a relationship between individual risk and cost-effectiveness. We believe that this relationship can and should be measured explicitly and used to optimise the cost-effectiveness of screening programmes.

We identify a number of modelling methods, with appropriate modifications, that may enable such a relationship to be defined. In practice, risk factors for a given population could be calculated based on routinely collected data. By estimating the relationship between risk and cost-effectiveness, physicians would be able to ensure that they offer patients the most cost-effective screening programme possible. We foresee that our work may also have implications for other preventative interventions; the cost-effectiveness of which is likely to depend upon the risk of an individual developing a given health problem.
References


FIGURE 1: SCREENING PROCESS
### Screening Pathway

<table>
<thead>
<tr>
<th>Screening pathway</th>
<th>Probability</th>
<th>Cost</th>
<th>Expected cost</th>
<th>Outcome</th>
<th>Expected outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>$r(1-\alpha)$</td>
<td>$C_S + C_T$</td>
<td>$(r(1-\alpha))(C_S + C_T)$</td>
<td>$Q - Q_D + Q_S$</td>
<td>$(r(1-\alpha))(Q - Q_D + Q_S)$</td>
</tr>
<tr>
<td>B</td>
<td>$r(\alpha)$</td>
<td>$C_S$</td>
<td>$(r(\alpha))C_S$</td>
<td>$Q$</td>
<td>$(r(\alpha))Q$</td>
</tr>
<tr>
<td>C</td>
<td>$(1-r)(1-\beta)$</td>
<td>$C_S$</td>
<td>$((1-r)(1-\beta))(C_S)$</td>
<td>$Q$</td>
<td>$((1-r)(1-\beta))(Q)$</td>
</tr>
<tr>
<td>D</td>
<td>$(1-r)(\beta)$</td>
<td>$C_S + C_T$</td>
<td>$((1-r)(\beta))(C_S + C_T)$</td>
<td>$Q - Q_D$</td>
<td>$((1-r)(\beta)(Q - Q_D)$</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>1</td>
<td>$C_1$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### No screening

| E                 | $r$ | $C_T$ | $(r)(C_T)$ | $Q - Q_D$ | $(r)(Q - Q_D)$ |
| F                 | $(1-r)$ | - | - | $Q$ | $(1-r)Q$ |
| **Total**         | 1 | $C_0$ | - | $Q_0$ | |

**Table 1: Screening Costs and Outcomes**

- $r$ = individual’s probability of developing a disease within a given period of time
- $C_S$ = cost of screening
- $C_T$ = cost of treating an individual if they develop the disease
- $Q$ = outcome for individual who does not develop disease
- $Q_D$ = detrimental effect on outcome of developing disease
- $Q_S$ = reduction in $Q_D$ from positive screen
- $C_1$ = expected cost of screening pathway for individual of risk $r$
- $C_0$ = expected cost of no screening pathway for individual of risk $r$
- $Q_1$ = expected outcome of screening pathway for individual of risk $r$
- $Q_0$ = expected outcome of no screening pathway for individual of risk $r$

- $\alpha$ = false positive rate
- $\beta$ = false negative rate
FIGURE 2: THE RELATIONSHIP BETWEEN RISK AND THE EXPECTED BENEFIT OF SCREENING

FIGURE 3: THE RELATIONSHIP BETWEEN RISK AND THE COST-EFFECTIVENESS OF SCREENING
<table>
<thead>
<tr>
<th>Programme</th>
<th>Population</th>
<th>Screening interval</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diabetic retinopathy (ENSPDR)</td>
<td>Diagnosed diabetes, over 12</td>
<td>1 year</td>
<td>Reduce sight loss</td>
</tr>
<tr>
<td>Bowel cancer (NHS BCSP)</td>
<td>Over 60</td>
<td>2 years</td>
<td>Prevent death</td>
</tr>
<tr>
<td>Cervical cancer (NHS CSP)</td>
<td>Women, 25-64</td>
<td>3 years (25-50) / 5 years (50-64)</td>
<td>Prevent death</td>
</tr>
<tr>
<td>Breast screening</td>
<td>Women, over 50</td>
<td>3 years</td>
<td>Prevent death</td>
</tr>
<tr>
<td>Abdominal aortic aneurysm (AAA)</td>
<td>Men, 64</td>
<td>One screen</td>
<td>Prevent death</td>
</tr>
<tr>
<td>Newborn (various)</td>
<td>All newborns</td>
<td>One screen</td>
<td>Improved management</td>
</tr>
<tr>
<td>Antenatal (various)</td>
<td>Pregnant women</td>
<td>One screen</td>
<td>Improved management</td>
</tr>
</tbody>
</table>

**TABLE 2: CHARACTERISTICS OF ENGLISH NATIONAL SCREENING PROGRAMMES**